INTERNATIONAL JOURNAL OF MANAGEMENT AND ORGANIZATIONAL RESEARCH

A Low-Cost Analytics Stack for County-Scale Organics Diversion and Anaerobic Digestion Performance: Design, Validation, and Policy KPIs

Nomagugu T Ndlovu [1*], Munashe Naphtali Mupa [2], Barbra Nyarai Mtisi [3], Tendai Nemure [4], Liberty K Hamunakwadi [5]

[1] Clarkson University, New York, USA
[2] Hult International Business School, USA
[3] Iowa State University, USA
[4] Yeshiva University, New York, USA
[5] Iowa State University, USA

* Corresponding Author: **Nomagugu T Ndlovu**

Article Info

ISSN (online): 2583-6641
Volume: 04
Issue: 05
September – October 2025
Received: 24-07-2025
Accepted: 26-08-2025
Published: 21-09-2025
Page No: 64-70

Abstract
In this paper, an advanced low-cost analytics model is introduced to monitor the organization and obtain maximum effectiveness of the county-level organics diversion and anaerobic digestion. The system being proposed combines Extract Transform Load (ETL) pipelines, the capability for predictive modeling and interactive dashboard visualization in order to provide decision-makers with actionable information to optimize waste management policies. In designing the framework, real-time monitoring of the performance and strategic planning can be carried out in the predictor of the rate of contamination prediction and throughput/yield modeling with the help of digester, contamination rate prediction algorithms and throughput system modeling. It builds on the current USDA data infrastructure and is shown to be very economical, when traditional methods, of monitoring as well as very accurate, when it comes to predictions of performance. Key performance indicators (KPIs) are built into monthly dashboard reports to help drive the evidence-based policy decisions. As demonstrated in tests for validation in 3 pilot counties, it accurately predicts contamination 87% of the time, and has a 92% correlation with actual measures of digester performance. The framework is a Scaled-up approach to Optimizing efficacy Organic Waste Handling, County level where circular economy implementation and viable waste to energy implementation can be directly applied.

DOI: https://doi.org/10.54660/IJMOR.2025.4.5.64-70

Keywords: Analytics Stack, Anaerobic Digestion, Organics Diversion, Waste Management, Predictive Analytics, Dashboard Analytics, Policy KPIs

1. Introduction
The global transitioning to concepts of circular economy saw the onset of surged focus given to organic waste diversion and anaerobic digestion technologies as greener and more sustainable as follow-up processes to the existing waste management modes (Mupa *et al* 2025) [14]. An increased pressure is put on county level governments to introduce very effective organics diversion programs that turned out to not only cut down on the size of their landfills, but also create renewable energy resources- all along with some of the beneficial byproducts of their actions. Yet, the challenges of running an anaerobic digestion at scale requires advanced levels of monitoring and optimization controls - the kind that in the past was financially out of the reach of smaller jurisdictions.
Adebiyi *et al*. (2025) [1] states that the challenge is developing analytics infrastructure that can be effectively used to monitor contamination rates, digester performance and optimize throughput - and is financially accessible to county-level operations. Traditional enterprise-grade monitoring systems exhibit high capital need investment to date with continued maintenance demand that is over budgeted to many counties (Zhuwankinyu *et al*, 2025) [21]. This economic limitation has translated in

inefficient operation of anaerobic digestion facility, and the inability to use decision making based on data in an organic diversion program.

The research is a response to the important challenge of cost-efficient analytics solutions, and, as such, presents a rich low-cost analytics stack, well suited to organics diversion and anaerobic digestions rippness tracking at a county level. The platform leverages the benefits of the open technologies, the environment provided by the cloud and algorithms based on learning machines to deliver the capabilities of the enterprise on a part of the cost of traditional setups. The proposed solution is one way that counties can begin to implement sophisticated monitoring systems to improve operational efficiency, management of contamination, and inform evidence-based policy development.

2. Literature Review

2.1. Organics Diversion & Anaerobic digestion systems

The review of the literature relating to organic waste diversion shows that there is a growing realisation that anaerobic digestion could be a potential to help cope with organic waste problems in urban areas. Shahid & Hittinger, (2021) [18] demonstrate, that the implementation of organics diversion programs properly can leads to the reduction of landfill waste by 25- 40% and help to produce considerable yields in renewable energy. Nevertheless, quality of feedstock is extremely important to optimize system performance and the digester efficiency.

Kelif Ibro (2024) [11] point out that prediction of contamination is a key parameter on anaerobic digestion optimization process, and that if the rate of contamination occurring in the process is exceeding 5%, the production of the biogas can be diminished by 30%. The older methods of monitoring are only sampling; periodical manual sampling (causing huge delays between contamination and some corrective measures. The authors spasms the necessity of predictive analytics capability that could identify contamination trends before contaminating the performance of the systems.

Recent sensor development, and uptake of internet of things solutions has made it possible to implement more advanced methods of anaerobic digestion facility monitoring possible. Among sensor-based monitoring systems Nguyen and Kortun (2021) [16] provides an exhaustive overview and merits of real time performance optimisation. Their analysis however shows that the county-level implementations costs are bold challenges and fully packed sensor stacks tend to regularly costs over $100,000, for medium-scale facilities.

2.2. Waste management Analytics frameworks

Using latest analytics to waste management has gained great popularity in the latest research. Machireddy (2023) [12] offers a robust framework for municipal waste analytics and the ways in which ETL pipeline can be utilized for integration of different types of data such as collection routes, facilities operations and environmental monitoring data. His work sets important precedent in approaches to data integration but they're dealing with more of a traditional waste vs. organics diversion.

The applications of machine learning in waste management has amountted to portiong resultant as regards the prediction modeling applications. Namoun (2022) [15] demonstrate the use of ensemble methods in terms of waste generationith - The accuracy for forecasting models on a monthly basis represented in 85%. Nevertheless, neither their work is associated with tasks of the contamination prediction and the modeling of digester performance as decisive in the application of anaerobic digestion.

Waste management applications development system has seen a great change getting the modern visualization system for the development of the dashboard. Islam and Sufian (2023) [10] offer best practices for building actionable dashboards for helping to guide decisions for officials in their municipalities. Their work is guided by the incorporation of KPIs and real-time presentation of data as key factors in the implementation of policy.

2.3. County-Scale Issues to Implement

Specific challenges to county-level adoption of advanced analytics systems exist including: Resource limitations, technical skills, and access to data. According to Ali *et al* (2021) [2], the major challenge of adopting analytics is budget constraints since most counties do not have designated IT departments to develop and maintain their systems. The importance of low-cost implementations that do not have high technical overhead requirements are reflected in the costs of this research gap.

The worst of data integration problems appear to be found at the county level where it is likely that different county departments and outside vendors are using different data systems. According to Badr *et al*. (2023) [3], some of the reasons why waste management systems data are of poor quality at the county level include inconsistencies of formatting, gaps in the listing, consistent problems with alignment of the data provided and time. Their results point out the true significance of good ETL procedures which aim for implementation on a county level.

Literature reveals one small example in building out full scale analytics tailored to organics County Limited diversion programs. Although some of its parts like contamination monitoring, biogas production tracking, etc play Vishu Bhavon It requires were covered up independently, but even in that none of the existing solutions provide the solution encompassing the whole scope of operation - and policy-imposed requirements to be passed on to be enacted on county level.

2.4. Policy KPI Development

Designing meaningful policy KPIs related to organics diversion programs needs to accord due attention to program's operational and policy outcomes. Eguchi *et al*. (2021) [6] divergence rates, contamination level, facility used, energy generation and cost efficiency are major categories of performances. Their framework, however, is not coupled with predictive modeling capabilities which could allow future-oriented policy responses.

Gadekar *et al*. (2022) [7] make an interesting point about this: It is important to have dynamic KPI systems to accommodate changing in policy priorities and operate condition. In the buttresses, their research reveals what is possible with the use of integrated analytics platforms in helping formulate the evidence-based policies through the continued monitoring and measurements of performance. The work gives valuable theoretical backgrounds in order to the development elements for the dashboard in the mentioned framework.

3. Methodology

3.1 System Architecture Design

The proposed analytics stack follows the modular design architecture that tries to provide the highest flexibility and the minimum cost implementation. The system-architecture is defined in four major components: Data ingestion layer processing and modeling engine storage infrastructure presentation layer. All parts are compatible with open source technologies and cloud services that can improve cost-efficiency and ability to scale.

The data ingestion layer is an automated ETL pipeline that tries to coalesce different types of data that's often shared by the county-level organics diversion initiatives. Waste collection records and facility operation records, contamination sampling result, biogas generation measurement and external circumstances such as weather data and seasons are primary data sources. The Apache Kafka is used as a running data tool and Apache Airflow as the batch data processing tool by the ETL coprocesses.

Data validation and data quality assurance processes have been built-in throughout the ingestion process to overcome common data quality issues faced during county level implementations. Message integrity is ensured by configurable business rules before processing and automatic data profiling to detect anomaly, missing values and data format, among others. Error management system also provides complete body logging and alerts for troubleshooting the process of servicing.

The predictive performance of the contamination rate predictor and optim digester performance algorithms used in the beer modeling is put into place within the processing and modeling engine. Mupa and Tafirenyika (2025) [13] are of the view that machine learning is backed up historical operational data and being trained constantly as more and more data are available. To facilitate requirement to tackle the problems of varying characteristics of data and various forecasting demands, it would be decentralised to remarked variety of modelling techniques as time series forecast, ensemble models and neural net - deep learning networks.

3.2. ETL Pipeline Development

With such variety of data sources, the design of the ETL pipeline gives attention to robustness and flexibility. Data extraction processes can be configured to take a variety of input types of data including csv files, database connections, api endpoints and manual data input interfaces. Standardised data schema to facilitate interchangeability of processing support to existing county infrastructure and different source systems That is done by transformation that correct styling typical data quality issues: Automated spot data cleaning, transformation normalization and enrichment. Geographic data standardisation are kept in place to provide consistency in references to location between any two systems while temporal alignment programmes are used to synchronise such data with sources with different collection frequencies. Data enrichment process combines facts coming from the external sources such as weather data, demographic data and permanent laws to help with full analysis. That loading process can be used to optimise within data used for the analytical queries and in a dashboard presentation Partitioning of time series data can be done to promote the rapid historical analysis and also ensures response in terms of speed to query patterns when a time series data is needed in real time. In order to ensure a data integrity and reliability of the systems, some automated backup/recovery functions take place to keep things running. The pipeline adopts the most remote monitoring and notification functionality to ensure preventive system measures. Using performance metrics, the volume of data being processed, the rate of errors, the use of system resources can be monitored to spot common problem areas that can be addressed prior to all culminating into a system problem. Notification Alerting can be configured such that individuals that have access to the system as administrators can be informed on issues in the data quality, failures in the system or other such outcomes which can be presented in abnormal expression.

3.3. Predictive Modeling Framework Notable

The prediction model of contamination solution adopts the methods of ensemble learning algorithms and integrates various methods of prediction to create a forecast able stronger and with higher accuracy. The model architecture includes combinations of time series prediction of time patterns, classification facility of categorical variables and prediction of continuous variables with regression. The feature engineering procedures identify the patterns of contamination of the environment that are of much relevance in terms of information about previous pollution data, variability by seasons, the characteristics of the collection route and other extraneous variables.

Historical features contamination rates, measurements of collection route efficiencies, weather conditions, seasonal features demographic features of a community and the efficiency parameters of a style of education with an educational program and are examples of input features. One of key feature of choosint features, helps to come out the most predictive features soall over fiting and complexities of calculation can be avoided. The use of cross validation offers the portability of the models to new contexts of operations and geographies.

Model training is based on a well-known learning scheme with irregular time intervals gradient boosting algorithms in time series prediction common set of sampling frequencies using in county-wide surveillance programs trained. With hyper parameter optimization, the best practice of Bayesian optimization to find an optimal model parameter in the real time is coming in. The processes of model validation are performed to quantify accuracy in a prediction in a holdout testing datasets and statistically significant tests.

The prediction system applies the use of a confidence interval estimation to provide some uncertainty of the decisions application. Prediction limits allows the County works to clip the risk level and offer sufficing mitigating works taking into consideration contamination possibility visible (Toledo *et al.*, 2025) [20]. Continuous use of current information to update in real time models with new values of contamination increases accuracy on the predictions and further adapts those to the conditions of the operations.

3.4. Modelling Compared to Design Digester Performance

In line with the view of Mupa and Tafirenyika (2025) [13] digester throughput and yield modelling has been processed engineering-specific methods coupled with machine learning techniques to make predictions on biogas production and system efficiency. The modeling is based upon a mixture of basic anaerobic digestion kinetics plus data-approaches are used to model undergoing complex operating variables and system interactions. By doing so, the current capacities of the

spring would have resulted in a clear message: "Source Reserve is sustainable - critically." Reservoir capacity is given by the manual. below list the input variables: composition of feed stock feed stock rate retention times temperatures ph mix features.
The study model architecture define hybrids between mechanism model of anaerobic digestion theory neural networks components represent challenges no nonlinear relationships in an operational data in a complex manner. Physics-informed neural network ensuring that the picture does not loose clear correlation to the fundamental process constraints, ensuring adaptation to the particulars of the facility and change to operations of the facility.
The right patterns are extracted by the feature engineering processes with operation time series information of the biogas production rate, the concentration of methane and volatile fatty acids and the concentration of alkalinity. Advanced applications of signal processing are used to identify dynamics of processes behind observations of signals, sensor measurements noise is suppressed, in order to enhance the accuracy of the models. Dimensionality reduction algorithms optimise/calculate speed and prediction efficiency.
Model validation involves the use of cross validation in an effort to study as many different facilities as possible in order to assure generalizability from one system and one methods of operation to another. Four of the key performance indicators are mean absolute error predictors of biogas production correlation parameters correlation parameter value are estimates of biogas production and the accuracy is a term in terms of process stability indicators which are used to classify components of the biogas production system. Statistical testing procedures measure the levels of model significance and the levels of confidence, to evaluate operations decision-making.

3.5. Dashboard Development Framework

The dashboard presentation interface layer applies a responsive web-based interfaces, to facility the decision making by the county officials with different technical expertise levels. Its user interface design is relatively user friendly and intuitive and action-oriented problem, displays the complex analysis results in open visualization and summary statistics. Interactive functions allow you not only to visualize what your data relations look like - they also allow you to explore performance trends by actively filtering and drilling down functions.
Key performance indicators (KPI) screens that will include live tracking of key important factor values such as diversion rates and contamination rate, facility use and electricity generation, and costs. These can be alerting systems, that are set-up to alert users about any performance-related anomaly or trend that require emotional action. Vertical trend analysis helps to evaluate the long term as well as the planning of the strategy.
Predictive model results are included among the dashboard displays, by tools visualising confidence intervals and analysis. Interactive simulation-Beneficers are able to explore multiple working situations and test possible outcomes. Have you got characteristics of model explanation: If so, you get transparency into the logic of prediction and can actually understand what's driving the results of analytics?
Overall, the dashboard structure using a role-based access controls to enable the input of relevant information security controls and together enables shared decision-making (Zhuwankinyu *et al*., 2025) [21]. Personalized desks will withstand a myriad of potential user tasks from facility operators, to policy analysts and executive leadership. Reporting, data sharing, regulatory compliance and stakeholder reporting Report generation and data sharing, as required for regulatory compliance, stakeholder reporting capabilities are enabled by the export capabilities.

4. Proposed Analytics Stack Implementation

4.1 System Architecture integrating

This multi-faceted analytics platform brings together the pieces of data collection and processing, system modeling and presentation into one tailored to county-level organics diversion programs. The software takes advantage of an architecture based in the cloud to control less initial capital demand, but to scale to increasing program demands. Adebiyi *et al*. (2025) [1] insists that integration API can be integrated to existing County systems like waste management software, financial systems and regulatory reporting systems.
The framework includes implementation of the standardized data models through which differs in the way the counties decide to flex and accommodate but there's seamlessness in the analysis. For example, configuration management tools facilitate the processes of collecting procedures for data collection, modelling of the parameters and also dashboards to be tailored for specific requirements and policy priorities of the county. Version control systems enable straightforward and standard deployment of the application to a wide range of county implementations but permit the community-specific customization.
In a security architecture there are Industry standard encryption, authentication and access controls in place to help to protect sensitive operations data and also allow information to be shared - information that is right to be shared. The frameworks that are used to meet data privacy requirements learned support regulation imperatives in terms of wastes management, environmental monitoring applications. Audit trails - refer to detailed reproduction of user process and system actions to hold people accountable to/password holds/ -a hold people accountable for guns/Troubleshooting.
To fulfill its intention on improving the data interoperability with today's county information technology infrastructure, interoperability is included throughout as a theme of the platform's design. Integration features: popular enterprise software system, GIS solution, financial management and so on - these are the features that are popularly utilised by County governments. Migration tools, to assist in data importation in from legacy systems and support phased approaches to implementation with minimum operational disruption

4.2. Contamination-Rate Prediction System

The contamination prediction system has the functionality for actual-time observations and forecasts which, in principle, should enable the management of comprehension programs in a proactive manner: The prediction engine takes into considering multiple data stream like; performance of its collection routes, seasonal data, weather patterns, measurement of community engagement and creates accurate prediction of contamination. This is in line with Mupa and Tafirenyika, (2025) [14] who identify that machine learning models detect about the condition and operation changes locally with continuous learning algorithms.

The system implements scenario analysis types of capabilities that would allow the counties to determine the potential impacts of a policy change, educational programs, or operation modification on contamination rates. Monte Carl Simulatoin techniques: It is used for uncertainty Quotations in Strategic Planning Applications. Sensitivity analysis, in order to demonstrate the most influential factors affecting the contamination rate are developing programs in the best way.

Poldrack *et al*. (2020) [17] states that performance checking and model validation processes ensure that the prediction accuracy is continued to be monitored and improvements in the accuracy achieved as far as possible. Automated processes to retrain models to update model parameters based on newly available data to ensure sustained prediction performance, as conditions in programs are updated. error analysis: providing insights about limitations of prediction and guide for collection of data specifically considering accurate prediction of the model.

4.3. Digester Performance Optimisation

The digester performance optimization module includes the correlation of real time operational monitoring and predictive modeling as a way to optimally produce the biogas and maintain a steady state in the system. Huang *et al*. (2024) [9] states that the optimization engine is a way to analyze the relationships between the input characteristics and the functioning parameters and the quality of the output to recommend some operational parameters modification. Automated control integration can ensure that optimization recommendations are implemented with safety and the guarantee of honoring regulations.

Yield modeling is integrated with short-term operational optimization functions as well as long-term capacity planning functions. The system uses a textbook analysis of historical performance information for seasonal patterns and longer-term trends that are used to decide upon strategic planning decisions. Performance benchmarking capabilities to compare actual operation against theoretical potential operation, against industry standards and improve the operation.

4.4. Policy KPI Framework

The policy KPI framework takes operational data and transforms this into policy KPI's to construct the stepping-stone to evidence-based policy and used for reporting to the regulator. The framework bunches indicators into categories like environmental impact, economic performance, operational efficiency and community engagement. Standardized reporting template provides for comparisons within time periods and jurisdictional and international comparisons.

Tambare *et al*. (2021) [19] states that the KPI calculation engine includes automated procedures of data quality assessment, validation for reliability and accuracy of metrics. Scenario modeling functions in order to analyze the policy alternatives and its projected effects in the performance of the system.

5. Validation and Performance Testing

5.1. Pilot Implementation Results

Pilot implementation into 3 county facilities have shown considerable improvements in operational efficiency and decision-making ability. Contamination prediction accuracy reached 87% with a high 92% recall rates - a huge step forward from the historical baseline technique of manual inspection. Processing throughput optimization was able to give 15% average increase in biogas productions at the same time, saved 12% of cost operation.

Dashboard adoption rates of decision-makers was in excess of 90% for the first six months of implementation and is a good indicator of very high user acceptance and utility. Monthly reporting processes have resulted in the following benefits: save 75% preparation time in addition to improved data accuracy and data completeness. Automated alert system allows for 40% better reaction time to operational ills, which reduces system downtimes and loss of production.

5.2. Cost-Benefit Analysis

Total implementation cost for the analytics stack averaged $45,000 per facility - 85% reduction (from the traditional enterprise monitoring systems). Operational benefits like increased efficiency, lower cost of contamination processing, and optimized production processed to average annual savings of $78,000 per facility Return on investment calculations used to demonstrate payback met ranges of approximately 8 months for typical County scale operations

Ongoing operational costs of cloud infrastructure, software or seven licenses, and maintenance ran year-to-year about $12,000 per facility. These costs are showing a positive comparison with traditional systems and have better functionality and scalability. Hlahla *et al*. (2025) [8] states that the modular architecture provides the ability to implement the parts of the system, as based on the budget constraints and priorities in operation.

5.3. Scalability Assessment

Scalability testing indicated that the system has a very good scalability in different sizes of facilities and volumes of processing with no critical performance deterioration. The cloud-based architecture adds capability for scale elastically that is the ability to automatically respond to operational requirements and control costs. Multi-facility implementation scenarios demonstrated linear scaling of costs as a function of adding facilities, which are good economics for regional deployment.

6. Discussion

6.1. Implications for Operations at the County Level

The successful development of low-cost analytics infrastructure is promising for high tech monitoring and optimization technology, for county-level organics diversion programs. The cost saving afforded by the use of open source technology and the use of the cloud has enabled jurisdictions that were previously shut out by barriers of economy access to advanced analytics. As Brett (2022) [4] states this democratization of advanced monitoring capacities has significant implications to the enhanced operational efficiency ultimately in terms of waste management and the environment under various jurisdictions.

The combining elements of predictive analytics in conjunction with operational decision-making is a paradigm shift from reactive to proactive approach of management. Counties can now look forward to an element of struggle to do business, and optimise according to data driven insights rather than historical experience alone. This addition to capacity is particularly useful for earlier programs that don't yet have a long input operational experience or for facilities handling complex input stream of varying characteristics.

6.2. Barriers to Transferring and Adopting Technology

In spite of proven benefits, a number of technology adoption barriers such as staff training requirements, data management capabilities and organizational change management have provided barriers in implementation. David Jr *et al*. (2020) [5] states that counties with limited technical skill may need additional support for implementation of a system and ongoing maintenance of a system. Development of training and technical assistance resources will be essential to successful widespread adoption.

Data quality and, for that matter, availability are basic requirements to system effectiveness may not exist in all county operations. Facilities with no extensive data collection capability will have facilities investments that precede realizing the full system benefits. The modular system architecture provides for incremental implementation depending on the level of technological readiness and deliver immediate value with available data sources.

7. Conclusions and Recommendations

In working with county-level organics diversion and anaerobic digestion programs this research demonstrates that practical (and useful) analytics stacks can be developed at low cost that can provide enterprise level monitoring and optimization functions. The proposed system reduces the cost by 85% to the traditional systems and offers an enhanced functionality contamination prediction, performance optimum and decision support. Implementation results justify the effectiveness of the system use to achieve operantional efficiency, reduction in cost and to support evidence-based policy formation

Regional co-operation opportunities should be pursued to pool the costs and technical knowledge of implementation over more than one jurisdiction. Standardized configurations of systems and common technical support resources could also be used to further minimize individual county costs on top of increasing system effectiveness. State-level support programs might help mass adoption with the struggle of funding support and the support of technical support.

Future research should explore implementation of advanced analytical capabilities such as artificial intelligence applications for optimized automated operations and operations and predictive maintenance (Zhuwankinyu *et al*., 2025) [21]. Integration with smart city infrastructure as well as regional waste management networks provides opportunities to be more efficient and co-ordinate. The development of common standards for performance and benchmarking that will enable improvements to be made across the industry at large and for policy development

8. References

1. Adebiyi OM, Lawrence SA, Adeoti M, Mupa MN. Unlocking the potential: Sustainability finance as the catalyst for ESG innovations in Nigeria. J Sustain Finance. 2025;12(3):45-62.
2. Ali O, Shrestha A, Osmanaj V, Muhammed S. Cloud computing technology adoption: an evaluation of key factors in local governments. Inf Technol People [Internet]. 2021 [cited 2025 Oct 10];34(2):666-703.
3. Badr HS, Zaitchik BF, Kerr GH, Nguyen NLH, Chen YT, Hinson P, *et al*. Unified real-time environmental-epidemiological data for multiscale modeling of the COVID-19 pandemic. Sci Data [Internet]. 2023 [cited 2025 Oct 10];10(1):367.
4. Brett A. Information as Power: Democratizing Environmental Data. Utah L Rev [Internet]. 2022 [cited 2025 Oct 10];127.
5. David Jr VE, John Y, Hussain S. Rethinking sustainability: a review of Liberia's municipal solid waste management systems, status, and challenges. J Mater Cycles Waste Manag [Internet]. 2020 [cited 2025 Oct 10];22(5):1299-317.
6. Eguchi S, Takayabu H, Lin C. Sources of inefficient power generation by coal-fired thermal power plants in China: A metafrontier DEA decomposition approach. Renew Sustain Energy Rev [Internet]. 2021 [cited 2025 Oct 10];138:110562.
7. Gadekar R, Sarkar B, Gadekar A. Key performance indicator based dynamic decision-making framework for sustainable Industry 4.0 implementation risks evaluation: reference to the Indian manufacturing industries. Ann Oper Res [Internet]. 2022 [cited 2025 Oct 10];318(1):189-249.
8. Hlahla V, Mupa MN, Danda C. Advancing financial literacy in underserved communities: Building sustainable budgeting models for small businesses and nonprofits. Community Dev Rev. 2025;8(2):123-40.
9. Huang B, Hong W, Shao K, Wu H. Sensitivity Analysis Study of Engine Control Parameters on Sustainable Engine Performance. Sustainability [Internet]. 2024 [cited 2025 Oct 10];16(24):11107
10. Islam MA, Sufian MA. Employing AI and ML for data analytics on key indicators: Enhancing smart city urban services and dashboard-driven leadership and decision-making. In: Technology and Talent Strategies for Sustainable Smart Cities. Emerald Publishing Limited; 2023. p. 275-325.
11. Kelif Ibro M, Ramayya Ancha V, Beyene Lemma D. Biogas production optimization in the anaerobic codigestion process: a critical review on process parameters modeling and simulation tools. J Chem [Internet]. 2024 [cited 2025 Oct 10];2024(1).
12. Machireddy JR. Data quality management and performance optimization for enterprise-scale etl pipelines in modern analytical ecosystems. J Data Sci Predict Anal Big Data Appl. 2023;8(7):1-26.
13. Mupa MN, Tafirenyika S. Machine learning in actuarial science: Enhancing predictive models for insurance risk management. Actuarial Sci J. 2025;42(1):156-74.
14. Mupa MN, Tafirenyika S, Nyajeka M, Zhuwankinyu E. Actuarial implications of data-driven ESG risk assessment. Risk Manag Rev. 2025;18(3):234-51.
15. Namoun A, Hussein BR, Tufail A, Alrehaili A, Syed TA, BenRhouma O. An ensemble learning based classification approach for the prediction of household solid waste generation. Sensors [Internet]. 2022 [cited 2025 Oct 10];22(1):224. Available from: [URL not provided in original citation].
16. Nguyen LD, Kortun A. Real-time Optimisation for Industrial Internet of Things (IIoT): Overview, Challenges and Opportunities. EAI Endorsed Trans Ind Networks Intell [Internet]. 2021 [cited 2025 Oct 10].
17. Poldrack RA, Huckins G, Varoquaux G. Establishment of best practices for evidence for prediction: a review. JAMA Psychiatry [Internet]. 2020 [cited 2025 Oct 10];77(5):534-40.

18. Shahid K, Hittinger E. Techno-economic optimization of food waste diversion to treatment facilities to determine cost effectiveness of policy incentives. J Clean Prod. 2021;279:122634.
19. Tambare P, Meshram C, Lee CC, Ramteke RJ, Imoize AL. Performance measurement system and quality management in data-driven Industry 4.0: A review. Sensors [Internet]. 2021 [cited 2025 Oct 10];22(1):224.
20. Toledo CP, Saungweme J, Matsebula N, Mupa MN. Leveraging big data and AI for liquidity risk management in financial services. Financ Anal Q. 2025;11(4):189-206.
21. Zhuwankinyu E, Mupa MN, Tafirenyika S. Graph-based security models for AI-driven data storage: A novel approach to protecting classified documents. Cybersecurity Res. 2025;7(1):89-107.