iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82772

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

A Streamlit Based Dashboard for Green Chemistry Reaction Analysis

Manoj H. N.[1], Hemanth Kumar[2]
[1]*Student, Department of MCA, JNN College of Engineering, Shivamogga, Karnataka, India*
[2]*Associate Professor, Department of MCA, JNN College of Engineering, Shivamogga, Karnataka, India*

Abstract: ***Green chemistry aim is to reduce pollution, energy usage and environmental damage caused by chemical reactions through selection of sustainable processes. This work focuses on Ni (Nickel) catalyst based Suzuki-Miyaura cross-coupling reactions taken from Open Reaction Database (ORD). The system processes reaction data and calculates green chemistry metrics such as Atom Economy (AE), Reaction Mass Efficiency (RME), Environmental Factor (E-Factor), Energy Consumption and CO_2 Emission. These metrics are analyzed through statistical methods such as Confidence Interval (CI), pairwise catalyst comparison (p-test), Analysis of Variance (ANOVA), Cohen's d effect size, eta-squared and Pearson correlation to evaluate the statistical significance of catalyst selection on reaction yield. A rule based Recommendation System operates on observed reaction patterns. A Custom Reaction Evaluator is also include to compute metrics from user defined reaction conditions. The system is developed using Streamlit based interactive dashboard where users can filter reactions and compare different conditions using interactive graphs along with downloadable report in PDF, DOCX and CSV formats.***

Keywords: Green Chemistry Metrics, Ni-Catalyzed Suzuki-Miyaura Reaction, Open Reaction Database, Streamlit Dashboard and Statistical Catalyst Analysis.

I. INTRODUCTION

In chemistry, reactions are mostly judged by how much product is formed. But this completely ignores what happens to the environment, how much waste is generated, how much energy is consumed and what materials are used up. Many reactions that look successful in lab may actually be producing large amounts of waste which makes them poor choice for sustainable chemistry. Green chemistry came as a solution to this by focusing on safer reaction design, less resource usage and lower environmental damage. But the real problem is that sustainability metrics are almost never reported alongside reaction data in literature and calculating them manually is both time consuming and error prone.

Open Reaction Database (ORD) changed this by making large scale structured reaction data publicly available with details on conditions, catalysts, solvents and outcomes. Ni-catalyzed Suzuki-Miyaura reactions are selected here because nickel (Ni) is cheaper and generally considered more eco-friendly than palladium which is traditionally used. But just because a catalyst is cheaper does not automatically make the reaction sustainable, full metric evaluation is still needed. By converting ORD data into green chemistry metrics and applying statistical analysis, this work allows proper multi-condition catalyst comparison without manual effort.

This work presents an interactive Streamlit dashboard for sustainability analysis of chemical reactions on one single platform. Users can filter reactions, view interactive charts and understand relationships between yield, waste, energy and environmental impact. Report export in PDF, DOCX and CSV is also supported for documentation purpose.

II. LITERATURE SURVEY

The Open Reaction Database (ORD) in [1] presents an open database to store chemical reaction data in a standard format. The authors mainly focus on making reaction data reusable, complete and easy for computer processing. This is one of the backbone of entire work to build any chemistry background application. The GitHub repository in [2] provides the technical implementation of ORD, which has data schema, protocol buffer formats and validation rules for reaction records. It focuses on reaction data that is stored in a structured and machine readable way, which supports large amount of data processing and reuse of experimental reactions for computational analysis. It does not include interpretation of reaction quality or sustainability metrics, so external analysis systems are required to evaluate reaction performance. Authors review in [3] presents a detailed discussion of green chemistry metrics such as AE, RME, E-factor, Process Mass Intensity (PMI) and eco-scale. They explain the difference between theoretical metrics like AE and practical metrics like RME, which includes yield and material usage.

©IJRASET: All Rights are Reserved

The paper concludes that single metric is insufficient and recommends combining multiple indicators to assess greenness. However, the study is limited to analytical chemistry and does not include statistical analysis which are addressed in this work. The review paper by Sheldon in [4] explains E-factor concept and its importance in the development of green chemistry. The author highlights that E-factor shifts focus from yield-based evaluation to waste reduction and resource efficiency in chemical processes. They discusses how metrics such as E-factor, AE and PMI help to identify reactions that generate large amounts of waste particularly in fine chemical and pharmaceutical industries. However, the paper does not provide a computational tool for applying these metrics in software tools. The work in [5] reviews the E-factor and its role in promoting waste prevention in chemical processes. It explains that the E-factor helped to shift attention from only yield to the amount of waste produced in reactions and discusses how waste relates to major environmental issues like climate change and pollution. This also highlights the importance of using catalysts, alternative solvents, renewable resources and energy-efficient methods to reduce waste. However, it does not provide a specific software tool for calculating these metrics.

Work in [6] presents a different reaction metrics used in green chemistry analysis. The author explains how primary metrics like reaction yield, AE can be combined into a general RME and linked to the E-factor. The paper shows how these relationships help compare the efficiency of different synthetic routes and identify the best or worst reactions in a process. However, it focuses on theory and spreadsheet implementation and does not provide a full software dashboard to visualize metrics from large experimental datasets. The AI4Green system in [7] is an open source Electronic Lab Notebook (ELN) designed for sustainable chemistry. It allows users to manually enter reaction details and automatically calculates green chemistry metrics. It also shows chemical hazards using colour codes. The system supports report export in PDF and CSV formats that are mainly used for academic and industrial laboratory. However, AI4Green works only for single reactions which entered manually and does not support bulk reaction data processing, analysis or comparison. Hae-Young Kim in [8] describe how ANOVA and post-hoc multiple comparison tests should be used when more than two groups are compared. The author explains that ANOVA can only confirm whether a difference exists or not, but it cannot show which specific groups are different. Because of this, post-hoc methods like Tukey, Tukey-Kramer, Bonferroni and Scheffé are needed for pairwise comparison when group sizes are not equal. This methodology helps to use the statistical approach for this work, where ANOVA is also applied to check overall differences in catalyst yield and p-value tests are used to identify which catalyst groups differ significantly. The authors in [9] discuss Ni-catalyzed Suzuki-Miyaura cross-coupling reactions using greener solvents as they contribute a major portion of waste in chemical synthesis and summarize the importance of solvent choice for sustainable reactions. In this work, solvent effect is not studied separately, but it is automatically included in waste-based metrics such as E-factor and PMI for molecular weight. This shows why mass-based sustainability metrics are useful even without directly selecting solvents. The work in [10] presents a data-driven approach to analyse reaction sustainability by systematically calculating green chemistry metrics for a large set of reactions. The authors show how using structured data makes it possible to compare reactions based on AE, E-factor and other waste metrics and shows patterns in reaction efficiency across reaction types. It highlights the importance of automated metric calculation for large datasets and suggests that data standardization improves the quality of sustainability analysis.

Authors in [11] guides to a method to rank common and bio-derived solvents based on safety, health and environmental criteria. The authors combine data from existing industrial solvent guides and classify solvents into recommended, problematic, hazardous and highly hazardous categories. This study highlights that solvents contribute a large portion of material usage in drug synthesis, so choosing safer solvents can reduce environmental impact. However, the work mainly provides solvent ranking guidelines and does not calculate reaction level metrics such as E-factor or PMI, which limits direct comparison between complete chemical reactions. Work in [12] give reviews about how digital chemical data and software tools can help to find more sustainable reaction pathways from large reaction databases. The authors explains that current sustainability evaluation is slow, manual and they suggest that linking data, sustainability metrics and decision-making tools could help automate the search for greener reactions. This work supports our idea that structured reaction data needs to be combined with automated metric calculation and analysis. Authors in [13] discusses barriers to implementing green chemistry in pharmaceutical industry and proposes Green Aspiration Level (GAL) metric to quantify the environmental impact of chemical production processes. The absence of clear and standardized sustainability metrics makes it difficult to compare the greenness of different synthesis routes. The GAL concept helps measure how close a chemical process is to best achievable environmentally friendly process used in industry. However, the work focuses on industrial process level evaluation and does not analyse reaction datasets, which highlights the need for data-driven systems that compute green metrics from reaction databases. Sheldon in [14] reviews the development of green chemistry and sustainability metrics used to evaluate chemical processes. Metrics such as AE, E-factor, carbon efficiency help to measure environmental impact of chemical reactions and industrial processes.

The work also discusses how these metrics evolved over time and how they support pollution prevention and resource efficiency in chemical manufacturing. This study supports the idea of using quantitative sustainability metrics to evaluate reaction environmental impact using data-driven analysis. The authors in [15] proposes a simple quantitative method to compare different chemical synthesis routes based on their resource usage and environmental impact. They presented the Environmental Assessment tool for Organic Syntheses (EATOS) tool, which helps to calculate these values using available reaction data. This work supports the idea that quantitative metrics are useful for evaluating reaction sustainability, but it focuses on comparison of individual reaction rather than analysing large reaction datasets.

An Author team discuss the use of PMI in [16] as an important metric to evaluate environmental performance of chemical reactions, which helps to estimate waste generation and resource usage. They also mentioned that PMI can be misunderstood if calculation boundaries are not clearly defined and gave guidelines to apply this metric correctly during reaction development. This study supports the use of mass-based sustainability indicators like PMI and E-factor, which are also used in this work to evaluate environmental impact. Jonathan H. Chen and Pierre Baldi in [17] presents a rule-based expert system designed to support decision making on chemicals using predefined reaction rules. The system uses reusable reaction rules and reagent models that represent different reaction conditions such as catalyst type and solvent. These rules are ranked based on observed chemical patterns to determine the most likely reaction results. This study is similar to the rule-based recommendation approach used in this work to suggest reaction conditions based on observed data patterns. Work in [18] explains the concept of confidence interval (CI) and how it is used to estimate the reliability of statistical results. CI gives a range of values within which the true population value is expected to lie based on sample data. This helps to understand precision of estimates and support interpretation of statistical results. This concept supports the use of CI analysis to evaluate the reliability of catalyst yield values obtained from reaction datasets. Authors in [19] explains different types of correlation coefficients used to measure the relationship between two variables, common methods such as Pearson's r, Spearman's rho, Kendall's tau and explains when each method should be used depending on the data type. Among these, Pearson method is suitable for measuring linear relationship between continuous variables. Pearson correlation gives a value between −1 and +1, where positive values show variables increase together and negative values show an opposite relation. It also explains how correlation helps to understand patterns and relationships in data. This concept supports to observe relationships between reaction sustainability metrics and reaction performance. Francisco Pena-Pereira et. al., in [20] introduces Analytical GREEnness (AGREE), a metric system used to evaluate the greenness of analytical procedures based on principles of green analytical chemistry. This tool converts multiple sustainability criteria into a single score and shows the result using a visual pictogram for easy interpretation. This study supports the idea of evaluating chemical methods using quantitative sustainability metrics, which is similar to the metric-based evaluation approach used in this work.

III. PROPOSED SYSTEM

A web based analytical dashboard developed using Streamlit framework with Python as the primary programming language. The system is designed as a modular pipeline where each module handles one specific responsibility and passes its output to the next module. The overall architecture consists of five interconnected modules which are as shown in Fig. 1.

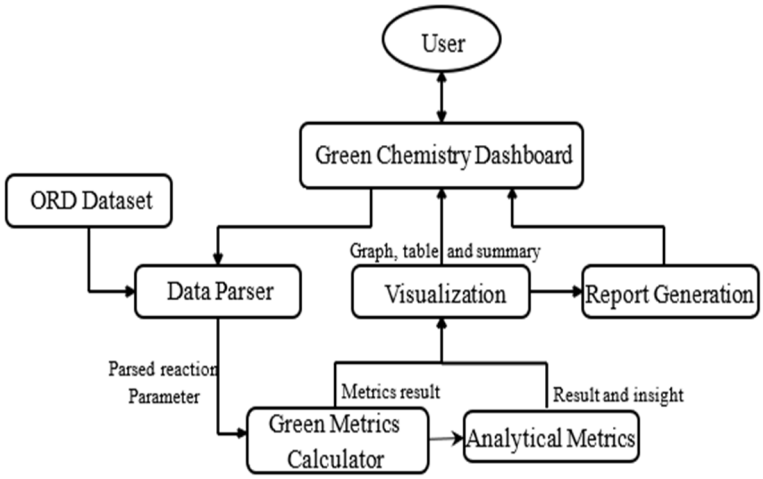


Fig. 1: Flowchart of Proposed System

Data Parser module reads raw ORD records stored in protocol buffer (.pb) binary format using protobuf library. ORD is an open access community driven repository which stores standardized experimental chemical reaction records with details of conditions, catalysts, solvents and measured outcomes. Since experimental parameters like temperature and time can be recorded in different units across different records, unit normalization to Celsius and hours is applied during extraction itself. All valid records are stored as structured in pandas DataFrame and passed to Green metrics module, which takes this cleaned data and applies green metrics uniformly across all reactions. All these calculations are done automatically using same logic for every reaction so there is no chance of human error and comparison is consistent.

Analytical module receives the enriched dataset and performs all statistical operations using SciPy library. Same module also handles recommendation generation using eq. 13 and custom reaction evaluation by passing user entered conditions which undergoes eq. 1 to 6 present in section IV. This ensures custom results are directly comparable with ORD reactions. Visualization module takes outputs from analytical and metric module and presents them as interactive charts using Plotly library. A sidebar panel allows user to filter active dataset by yield range, catalyst type and E-Factor limit. Report generation module uses ReportLab library for PDF export and python-docx library for DOCX export. CSV export is also available using pandas for users who want raw tabular data for further analysis. The complete module is presented through a multi-tab web interface and user can interact with any tab independently and apply filters at any point and all modules respond accordingly giving complete sustainability assessment experience within single platform.

IV. METHODOLOGY

The methodology followed in this work processes raw chemical reaction data through a structured computational pipeline to produce green chemistry sustainability insights. Each sub-section below describes the working logic of one stage in this pipeline.

A. Data Collection and Parsing

Data is collected from Open Reaction Database (ORD), where it stores data in protocol buffer (.pb) binary format which is not directly readable so we need to parse it. Data parser module reads each record and extract important reaction parameters like catalyst name, solvent type, temperature (°C), reaction time (hours) and product yield (%). If any record is having missing yield value or invalid data then it is removed during parsing phase. After parsing all valid records are stored in pandas DataFrame so that all reactions are in consistent format and ready for further analysis.

B. Green Chemistry Metrics Computation

The system calculates green metrics for each reaction. Since these data does not include molecular weights (MW), standardized assumptions are applied to all based on common green chemistry methods as in [3] and [6]. Fixed values like aryl halide MW = 150 g/mol, boronic acid MW = 140 g/mol, product MW = 200 g/mol, catalyst loading = 5 mol% (MW = 300 g/mol) under Ni-catalyzed Suzuki conditions in [9] and solvent volume = 2.0 mL are used. These assumptions ensure all reactions are evaluated in same way for fair comparison. So, the calculated metrics show relative sustainability differences rather than exact molecular accuracy. Below are the metrics calculated by the system:

1) Atom Economy (AE): Measures how efficiently reactant atoms are converted into desired product, where higher AE means lower atomic waste. It is considered a fundamental sustainability metric in [3] and [14].

$$AE = \frac{Molecular\ Weight\ of\ Desired\ Product}{\sum\ Molecular\ Weights\ of\ Reactants} \times 100 \qquad (1)$$

2) Reaction Mass Efficiency (RME): Measures how much reactant mass actually forms product by considering both reactant mass and experimental yield, making it more practical than AE alone is shown in [6].

$$RME = \frac{Mass\ of\ Product}{Total\ Mass\ of\ Reactants} \times 100 \qquad (2)$$

3) E-Factor: Measures waste generated relative to product mass, where lower value indicates more sustainable reaction and lower environmental impact, as discussed in [4] and [5].

$$E\text{-}Factor = \frac{Total\ Waste\ Mass}{Mass\ of\ Product} \qquad (3)$$

4) Process Mass Intensity (PMI): Measures total mass input that is needed to produce one unit mass of product, which is highlighted in [16].

$$PMI = \frac{Total\ Mass\ Input}{Mass\ of\ Product} \quad (4)$$

5) *CO_2 Emissions:* These are calculated from energy consumption using fixed emission factor, is presented in [13].

$$CO_2\ Emissions\ (kg) = Energy\ (kWh) \times Emission\ Factor\ (kg\ CO_2/kWh) \quad (5)$$

Where,

$$Energy\ (kWh) = k \times Reaction\ Time\ (h) \times Temperature\ Factor \quad (6)$$

Here, *k* represents a fixed proportional constant used to approximate energy demand under standardized laboratory assumptions. This estimation allows consistent comparison of energy usage across reactions even when exact equipment specifications are unavailable.

6) *Overall Green Score:* This is a composite weighted score that combines all above metrics into single value, thus reactions can be directly compared with each other which is discussed in [3] and [20]. Apart from this, each solvent present in the reaction record is also given a greenness score on scale of 1 to 10 based on solvent selection criteria as discussed in [11] where water and bio-based solvent have highest score and halogenated solvents are lowest. This solvent score is also included in overall green score calculation to consider solvent sustainability in reaction ranking.

C. Analytical Metrics

1) Statistical Analysis

Calculating green score for each reaction alone is not enough to make a proper conclusion. For Larger datasets, where same catalyst is tested many times in different conditions, average values can be misleading. Sometimes a catalyst tested fewer times may look better just by chance, not because it is actually more sustainable. So, statistical methods are used to check if the differences are real or just random, making the conclusion more accurate.

a. *Analysis of Variance (ANOVA):* One way ANOVA is used to compare reaction yields between catalyst groups by analyzing variation within and between groups. Before applying ANOVA, normality is checked using Shapiro-Wilk test. If data is normal, parametric ANOVA is used, otherwise Kruskal-Wallis test is applied. F-value and p-value determine whether catalyst choice significantly affects yield or differences are random. If $p < 0.05$, catalyst effect is considered significant.

$$F = \frac{MS_{between}}{MS_{within}} \quad (7)$$

Where,

$$MS_{between} = \frac{SS_{between}}{k-1} \text{ and } MS_{within} = \frac{SS_{within}}{N-k}$$

Here, *k* is number of catalyst groups, *N* is total reactions, SS is sum of squares and MS is mean square.

b. *Post-hoc Pairwise Comparison (p-test) and Effect Size:* After ANOVA, pairwise tests are used to identify which catalyst pairs differ significantly under same experimental conditions. Practical significance is measured using Cohen's d in [8].

$$d = \frac{|\bar{x}_1 - \bar{x}_2|}{s_{pooled}} \quad (8)$$

Where,

$$s_{pooled} = \sqrt{\frac{(n_1-1)var_1 + (n_2-1)var_2}{n_1 + n_2 - 2}} \quad (9)$$

Here, $\bar{x}$ is mean yield of a catalyst group, n is number of reactions available for that catalyst in the dataset and var is variance of yield values within that catalyst group. where, values of $d < 0.2$ is negligible, 0.2-0.5 is small, 0.5-0.8 is medium and > 0.8 is large difference.

c. *Eta-Squared (η^2) Analysis (Catalyst Importance):* It measures how much total yield variation is explained only by catalyst selection in [8]. Higher η^2 value means catalyst choice has stronger influence on yield.

$$\eta^2 = \frac{SS_{between}}{SS_{total}} \quad (10)$$

Where, $\eta^2 > 0.14$ indicates large catalyst influence on yield outcome.

d. *Confidence Interval (CI) Estimation:* CI estimates the range where true average catalyst yield is likely to fall, also helps to evaluate correctness when sample size is small or data variation is high. It is calculated using t-distribution as discussed in [18].

$$CI = \bar{x} \pm t_{\alpha/2,\, n-1} \times \frac{s}{\sqrt{n}} \qquad (11)$$

Here, $t(\alpha/2,\ n\text{-}1)$ is critical t-value at 95% confidence level with n-1 degrees of freedom and s is standard deviation of yield values for that catalyst group. Narrow CI indicates consistent results, while wide CI indicates high variation or insufficient data.

e. Correlation Based Sensitivity Analysis: Pearson correlation (r) measures linear relationship between sustainability metrics like yield, E-factor, RME, PMI, energy and CO_2 as in [19] where values range between -1 to +1. These correlations are used to generate heatmap analysis in dashboard.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}} \qquad (12)$$

Here, $\bar{x}$ is mean of first sustainability metric across all reactions, $\bar{y}$ is mean of second sustainability metric across all reactions, x_i is individual reaction value of first metric and y_i is individual reaction value of second metric. Where, values near ±1 show strong relationship and near 0 show weak relationship.

2) Expert Recommendation System:

This rule-based expert system approached in [17], where the recommendation module predicts suitable reaction conditions using reaction patterns without requiring machine learning models. Each reaction is evaluated using a composite score that balances yield and environmental impact equally by normalizing E-factor to a 0–100 scale.

$$Composite\ Score = (0.5 \times Yield) + \left(0.5 \times \frac{100}{1 + E\text{-}Factor}\right) \qquad (13)$$

This module has three options where best balanced priority option selects reaction with highest composite score, while yield priority selects maximum yield reaction and sustainability priority selects minimum E-factor reaction. All recommendations update dynamically based on filters which is applied by user.

3) Custom Reaction Evaluator:

Custom Reaction Evaluator is another useful feature where user can enter their own hypothetical reaction conditions like solvent, temperature (°C), reaction time (hours), scale (mmol) and expected yield (%). The system calculates all green metrics for user-entered conditions using the same computation applied to the dataset. This is helpful for researchers who want to check sustainability of a reaction before actually doing experiment in lab.

D. Interactive Visualization

The calculated metrics are shown in dashboard which is having multiple tabs. User can apply filters like yield range, catalyst type and E-factor limit and all charts will update in real time. These visualizations are also help to view more detail (such as numerical values) within graph and some graph have annotations which gives conclusion of entire graph which is dynamic and works through filter.

E. Report Generation

After completion of analysis, user can download the results. System supports three formats for export. CSV format is for tabular data. PDF format gives formatted analytical report with charts and tables. DOCX format is editable word document. Report generation is fully automated in backend and it ensures that whatever is shown on dashboard is same as what is in downloaded report.

V. RESULT AND DISCUSSION

The system was evaluated using Ni-catalyzed Suzuki-Miyaura reaction records extracted from ORD. Results across all analytical modules are discussed below.

A. Green Chemistry Metrics Output

After computing sustainability metrics for all valid reactions, yield showed wide variation from 1% to 88% with an average of 29.7%, indicating strong influence of catalyst and solvent conditions. AE remained constant at 68.97% due to fixed molecular assumptions, while RME varied with yield, reflecting actual mass efficiency differences. E-Factor showed high variation, where solvent-intensive reactions produced significantly higher waste, confirming solvent as major contributor to environmental impact.

Average CO_2 emission was approximately 0.32 kg per reaction, influenced by temperature and reaction time. Fig. 2 compares AE (Fig. 2.1), RME (Fig. 2.2) and E-Factor (Fig. 2.3) for top yield reactions. Showing that high AE does not always correspond to low E-Factor, proving that multiple metrics must be evaluated together for accurate sustainability assessment.

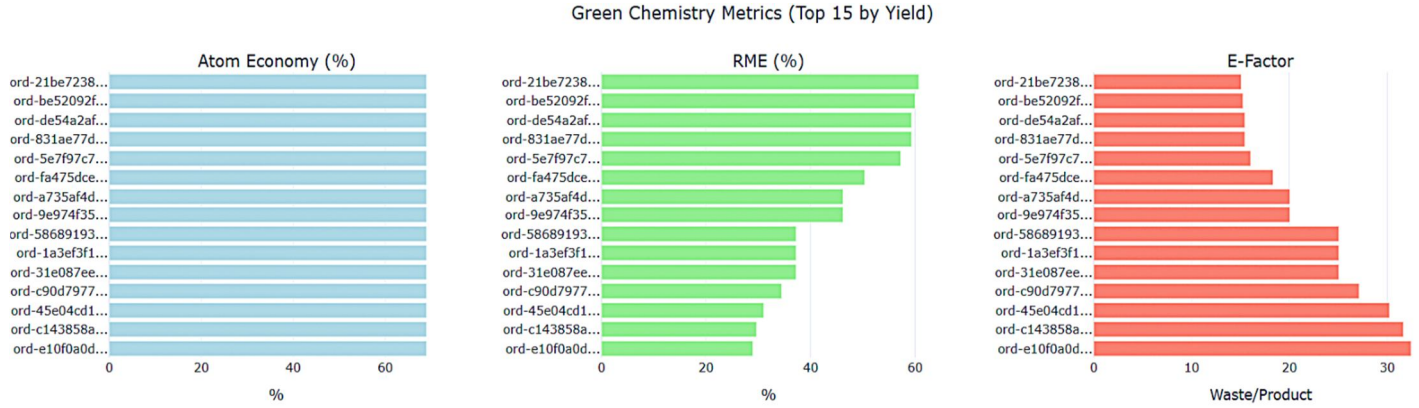


Fig. 2.1: AE Fig. 2.2: RME Fig. 2.3: E-Factor

Fig. 2: Comparative Analysis of Green Chemistry Metrics for Top Yield Reactions

Fig. 3 shows yield vs E-Factor scatter plot where each point represents a reaction coloured by catalyst type, with green region indicating ideal high yield and low waste zone. Most reactions lie outside this zone, showing only few catalyst combinations achieve both high yield and low E-Factor. The trend line indicates a moderate negative correlation, meaning higher yield reactions generally produce less waste per unit product.

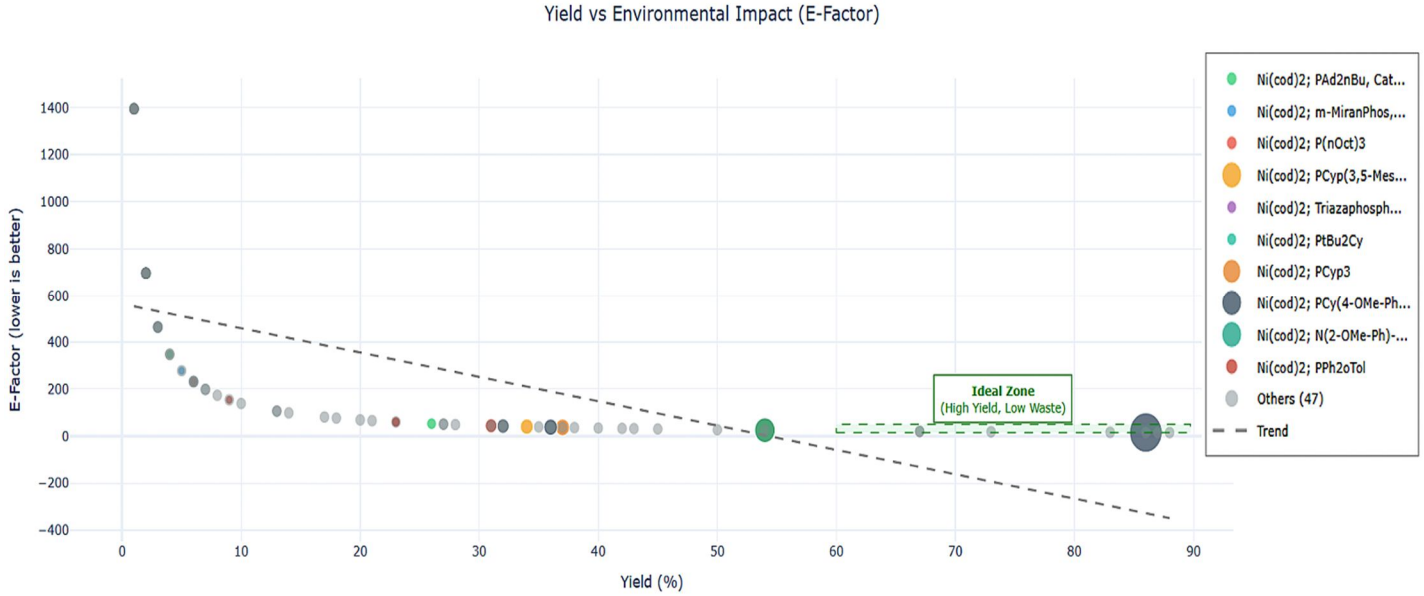


Fig. 3: Yield vs E-Factor Scatter Plot with Ideal Zone and Trend Line

B. Statistical Analysis Output

One-way ANOVA on yield across catalyst groups showed statistically significant effect ($p < 0.05$), confirming catalyst choice strongly influences sustainability outcomes. Cohen's d indicated medium to large practical differences, and η^2 showed catalyst selection explains a substantial portion of yield variation. Fig. 4 presents 95% CI error bars for mean yield, where narrow intervals indicate consistent catalyst performance while wide intervals suggest high variation or limited data. This confirms statistical analysis is essential beyond simple averages to identify reliable catalysts.

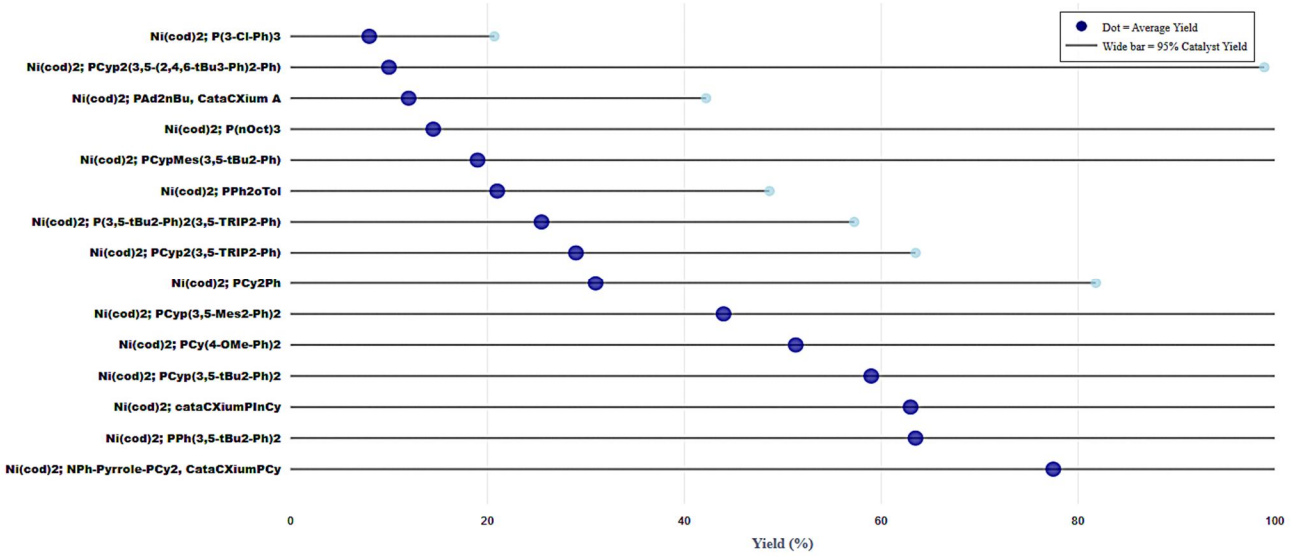


Fig. 4: Confidence Interval Plot for Catalyst Yield Comparison

Pearson correlation analysis showed strong positive correlation between E-Factor and PMI, while yield had moderate negative correlation with E-Factor consistent with Fig 3. Energy consumption and CO_2 emissions also showed strong positive correlation, as CO_2 is directly derived from energy in this system.

C. Custom Reaction Evaluator Output

Fig. 5 evaluates using multiple hypothetical inputs with varying scale, solvent, temperature and reaction time. For each case, the system computes full green metrics and compares them with dataset for yield, E-Factor and Green Score. The obtained results follow the same trends observed in ORD dataset analysis, showing that the computation pipeline gives consistent and comparable outputs for both real and user-defined reactions.

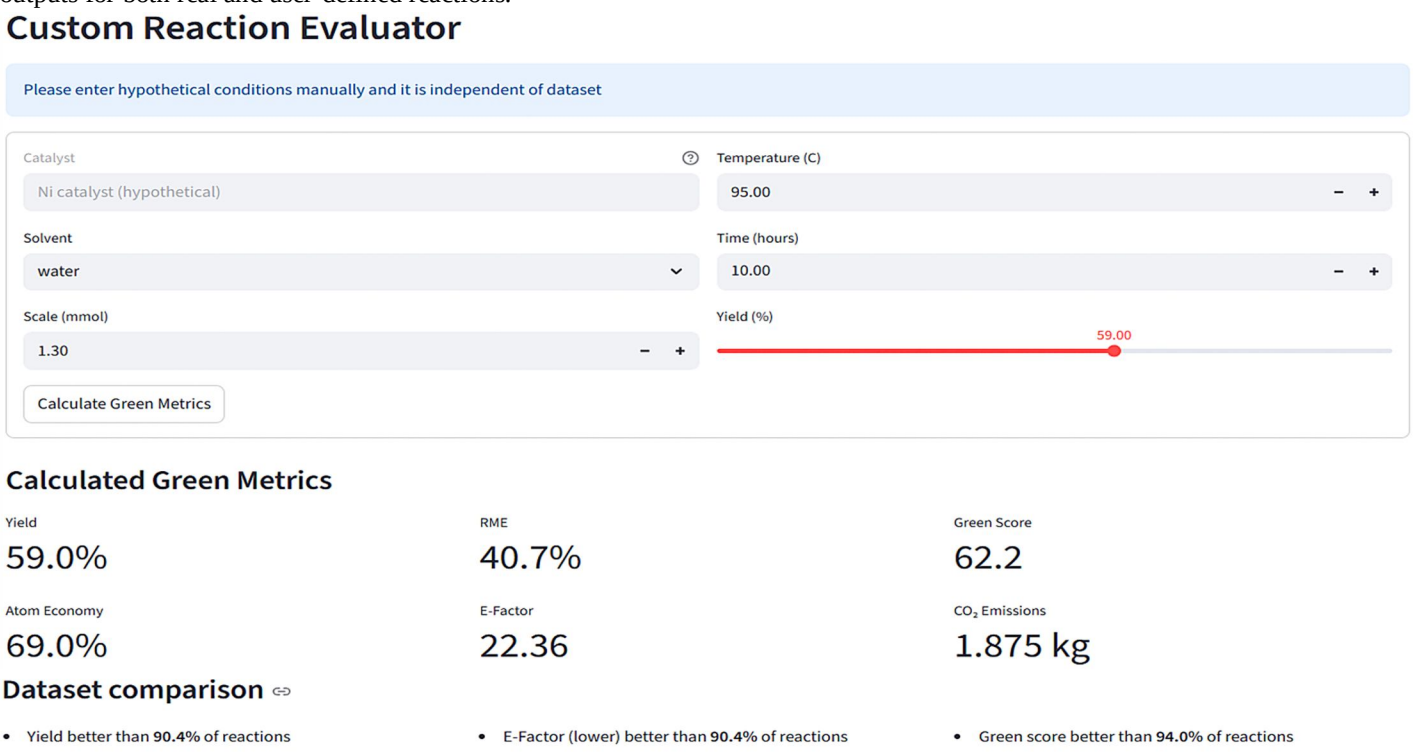


Fig. 5: Custom Reaction Evaluator output for a hypothetical Ni-catalyzed reaction

VI. CONCLUSION

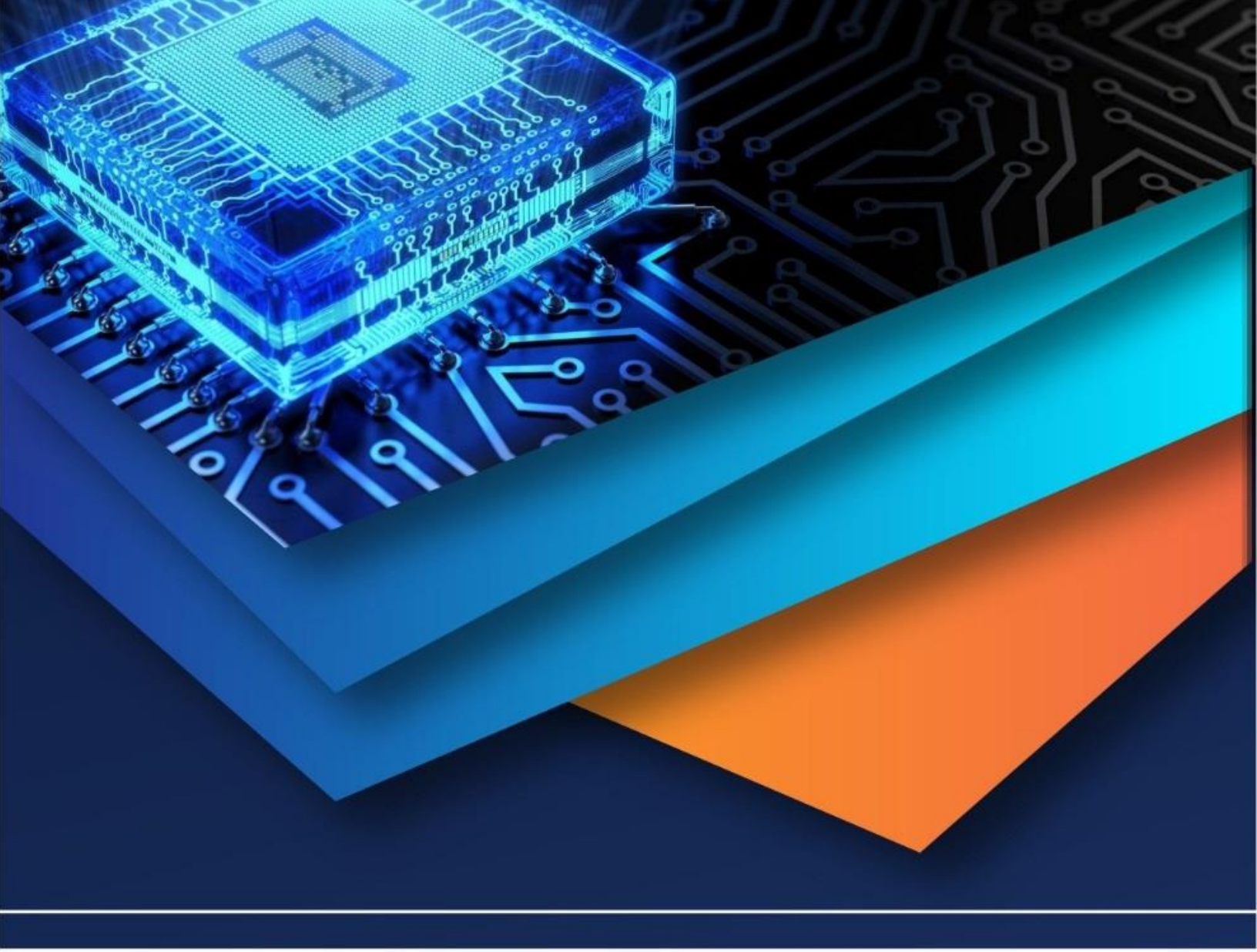
Green Chemistry Dashboard was built to make sustainability evaluation of chemical reactions simpler and more reliable than manual methods. The system takes real reaction data from ORD, computes green chemistry metrics, runs statistical analysis and presents everything through interactive visualizations in one platform. By combining all metrics into a single dashboard, users can compare catalyst performance and reaction conditions. Statistical methods further strengthen the analysis by confirming whether observed differences between catalysts are statistically real or just random variation. This work shows that data driven sustainability analysis gives much better insight into catalyst performance than just looking at yield alone. Features like Expert Recommendation System, Custom Reaction Evaluator and multi-format report generation make it useful for both documentation purposes.

In short, the dashboard provides a simple and effective platform for sustainability analysis of chemical reactions. The system can be further improved by implement Machine Learning and larger datasets in future, but the current dashboard processed and visualized reaction data consistently without any major issues and all modules worked as expected.

REFERENCES

[1] Steven M. Kearnes**,** Michael R. Maser, Michael Wleklinski, Anton Kast, Abigail G. Doyle, Spencer D. Dreher, Joel M. Hawkins, Klavs F. Jensen and Connor W. Coley, "The Open Reaction Database," Journal of the American Chemical Society, Vol. 143, No. 45, pp. 18820-18826, 2021.
[2] Open Reaction Database (ORD) Contributors, "Open Reaction Database ORD-Schema" GitHub Repository, Available: https://github.com/Open-Reaction-Database.
[3] M. Tobiszewski, M. Marć, A. Gałuszka and J. Namieśnik, "Green Chemistry Metrics with Special Reference to Green Analytical Chemistry," Molecules, Vol. 20, No. 6, pp. 10928-10946, 2015.
[4] R. A. Sheldon, "The E factor 25 years on: the rise of green chemistry and sustainability," Green Chemistry, Vol. 19, pp. 18-43, 2017.
[5] R. A. Sheldon, "The E factor at 30: a passion for pollution prevention," Green Chemistry, Vol. 25, pp. 1704-1728, 2023.
[6] John Andaros, "Unification of reaction metrics for green chemistry: Applications to reaction analysis," Organic Process Research & Development, Vol. 9, No. 2, pp. 149-163, 2005.
[7] Samuel Boobier, Joseph C. Davies, Ivan N. Derbenev, Christopher M. Handley and Jonathan D. Hirst "AI4Green: An Open-Source ELN for Green and Sustainable Chemistry," Journal of Chemical Information and Modeling, Vol. 63, No. 10, pp. 2895-2901, 2023.
[8] Hae-Young Kim "Statistical notes for clinical researchers: Post-hoc multiple comparisons," Restorative Dentistry & Endodontics, Vol. 40, No. 2, pp. 172-176, 2015.
[9] Stephen D. Ramgren, Liana Hie, Yuxuan Ye and Neil K. Garg, "Nickel-Catalyzed Suzuki-Miyaura Couplings in Green Solvents," Organic Letters, Vol. 15, No. 15, pp. 3950-3953, 2013.
[10] Fotouh R. Mansour, Mahmoud A. El Hassab, Taghreed A. Majrashi, Wagdy M. Eldehna, "A Modified Green Star Area (MoGSA) and software to assess greenness of reactions in the chemistry laboratories," PLOS One, Vol. 19, No. 12, Article e0314421, 2024.
[11] Denis Prat, Andy Wells, John Hayler, Helen Sneddon, C. Robert McElroy, Sarah Abou-Shehada and Peter J. Dunn, "CHEM21 selection guide of classical- and less classical-solvents," Green Chemistry, Vol. 18, pp. 288-296, 2016.
[12] Jana M. Weber, Zhen Guo, Chonghuan Zhang, Artur M. Schweidtmann and Alexei A. Lapkin, "Chemical data intelligence for sustainable chemistry," Chemical Society Reviews, Vol. 50, pp. 12013-12036, 2021.
[13] F. Roschangar, R. A. Sheldon and C. H. Senanayake, "Overcoming barriers to green chemistry in the pharmaceutical industry - The Green Aspiration Level concept," Green Chemistry, Vol. 17, pp. 752-768, 2015.
[14] Roger A. Sheldon, "Metrics of Green Chemistry and Sustainability: Past, Present, and Future," ACS Sustainable Chemistry & Engineering, Vol. 6, No. 1, pp. 32-48, 2018.
[15] Markus Eissen and Juergen O. Metzger, "Environmental Performance Metrics for Daily Use in Synthetic Chemistry," Chemistry - A European Journal, Vol. 8, pp. 3580-3585, 2002.
[16] Edward R. Monteith, Pieter Mampuys, Louise Summerton, James H. Clark, Bert U. W. Maes, and C. Robert McElroy, "Why we might be misusing Process Mass Intensity (PMI) and a methodology to apply it effectively as a discovery level metric," Green Chemistry, Vol. 22, pp. 123-135, 2020.
[17] Jonathan H. Chen and Pierre Baldi, "No Electron Left Behind: A Rule-Based Expert System to Predict Chemical Reactions and Reaction Mechanisms," Journal of Chemical Information and Modeling, Vol. 49, No. 9, pp. 2034-2043, 2009.
[18] Ana-Maria Šimundić, "Confidence interval," Biochemia Medica, Vol. 18, No. 2, pp. 154-161, 2008.
[19] Haldun Akoglu, "User's Guide to Correlation Coefficients," Turkish Journal of Emergency Medicine, Vol. 18, No. 3, pp. 91-93, 2018.
[20] Francisco Pena-Pereira, Wojciech Wojnowski and Miguel Tobiszewski, "AGREE-Analytical GREEnness Metric Approach and Software," Analytical Chemistry, Vol. 92, No. 14, pp. 10076-10082, 2020.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)