# DVD
# EVALUATION 1

TEAM 14

# CONTENT

# PVT GENERATION

- Temperature range should be from -55 to 125 Celsius (uniform distribution).
- Nominal Voltage = 1V. $\pm 10\%$ variations must be considered (uniform distribution).
- For Process, consider the values in the PTM file as nominal standards and vary it with $\pm 3\sigma$ variations satisfying Monte Carlo distribution, i.e, mean = nominal, std = mean/30.
- For Delay consider cqload variations from 0.01f to 5f (uniform distribution).

# NETLISTS

GENERATED AS PER REQUIREMENTS

SEPARATE SCRIPTS FOR MODULE DEFINITIONS AND CALCULATIONS

DONE FOR THE 45NM HP TECH NODE

# SCRIPTS FOR DATAGEN

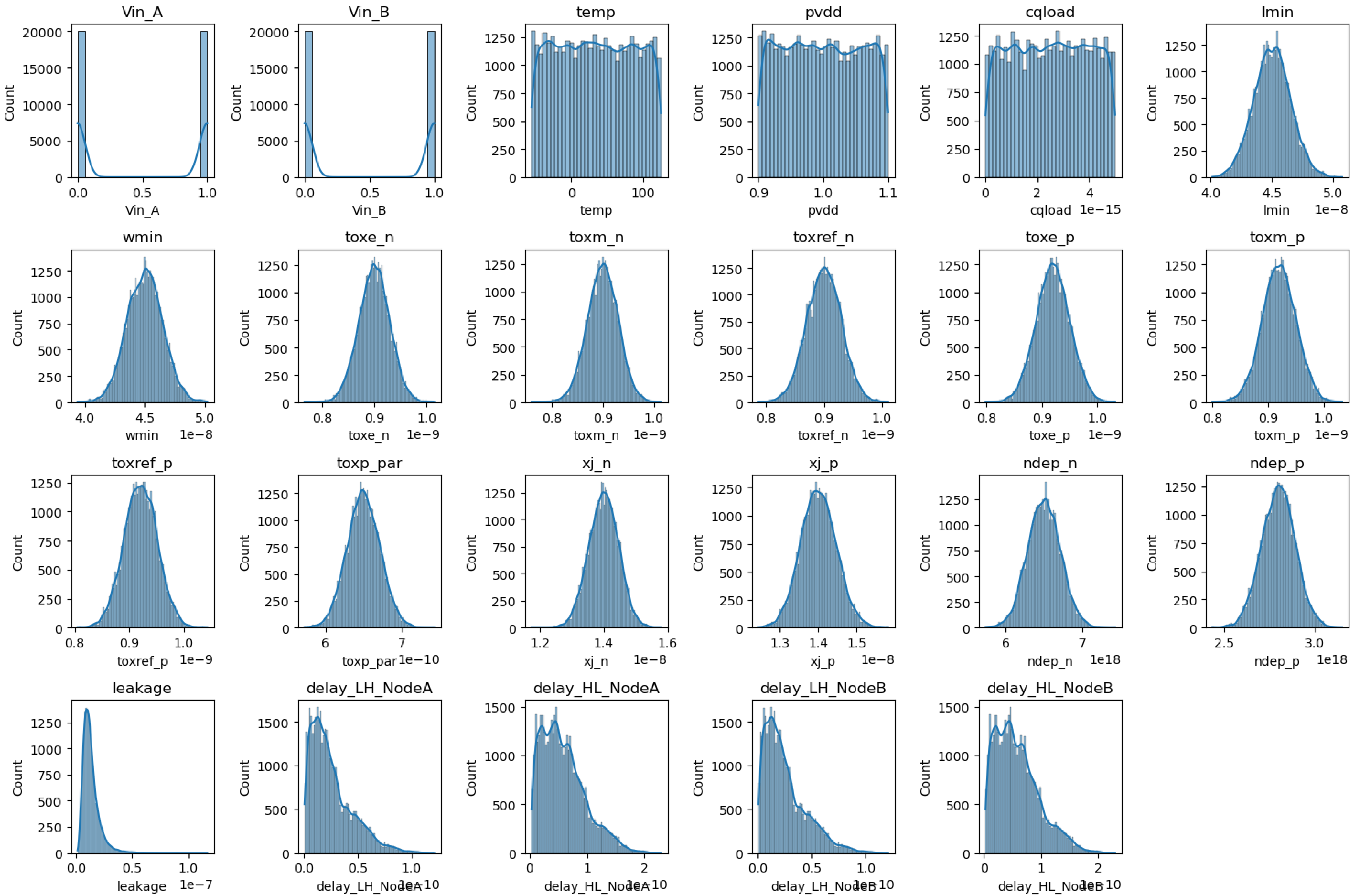
SEPARATE SCRIPTS FOR GATES

COMPUTED LEAKAGE AND DELAYS TOGETHER

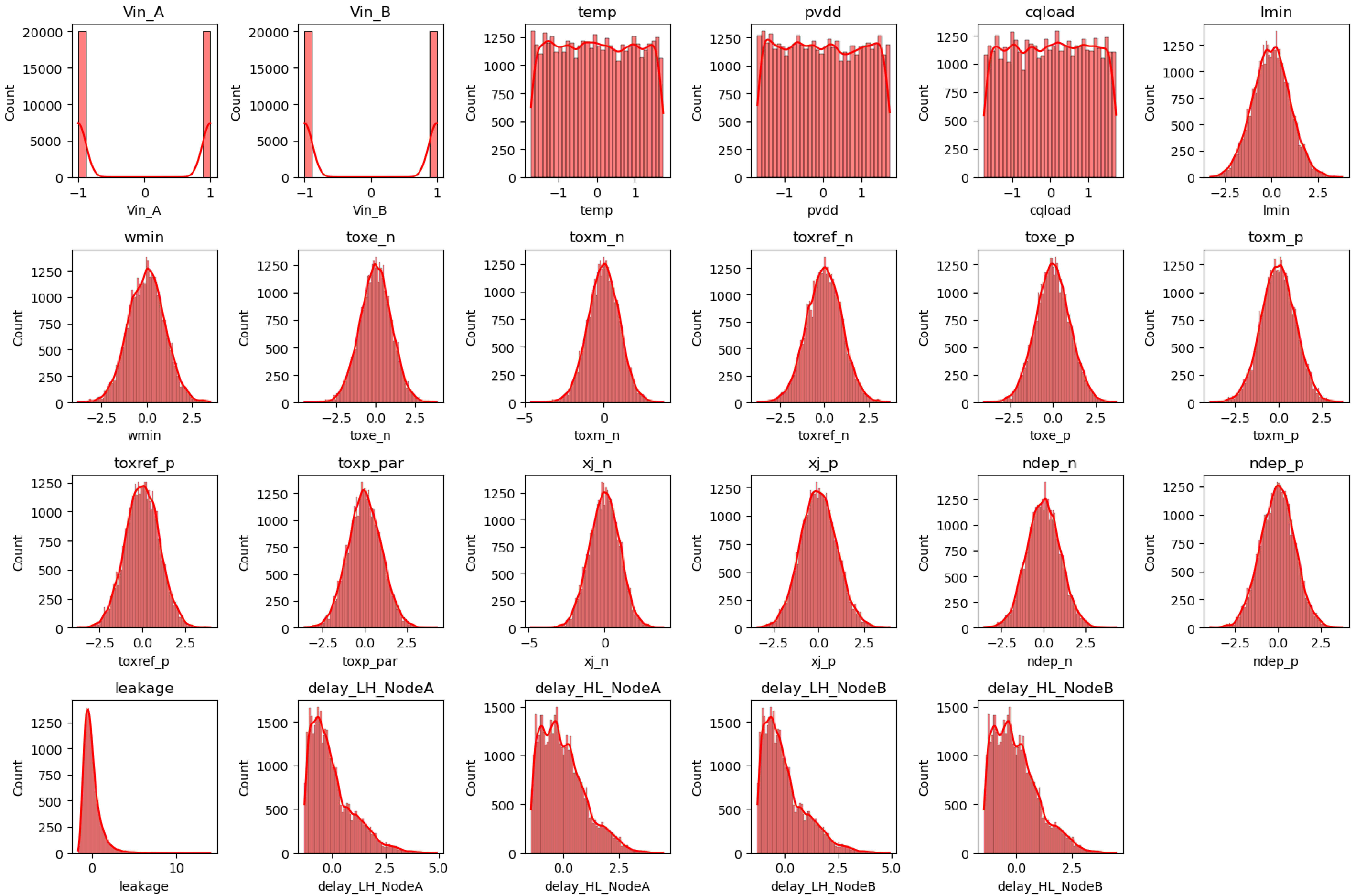
VALUES GENERATED FOR ALL 40K SAMPLES (20K FOR INVERTER)

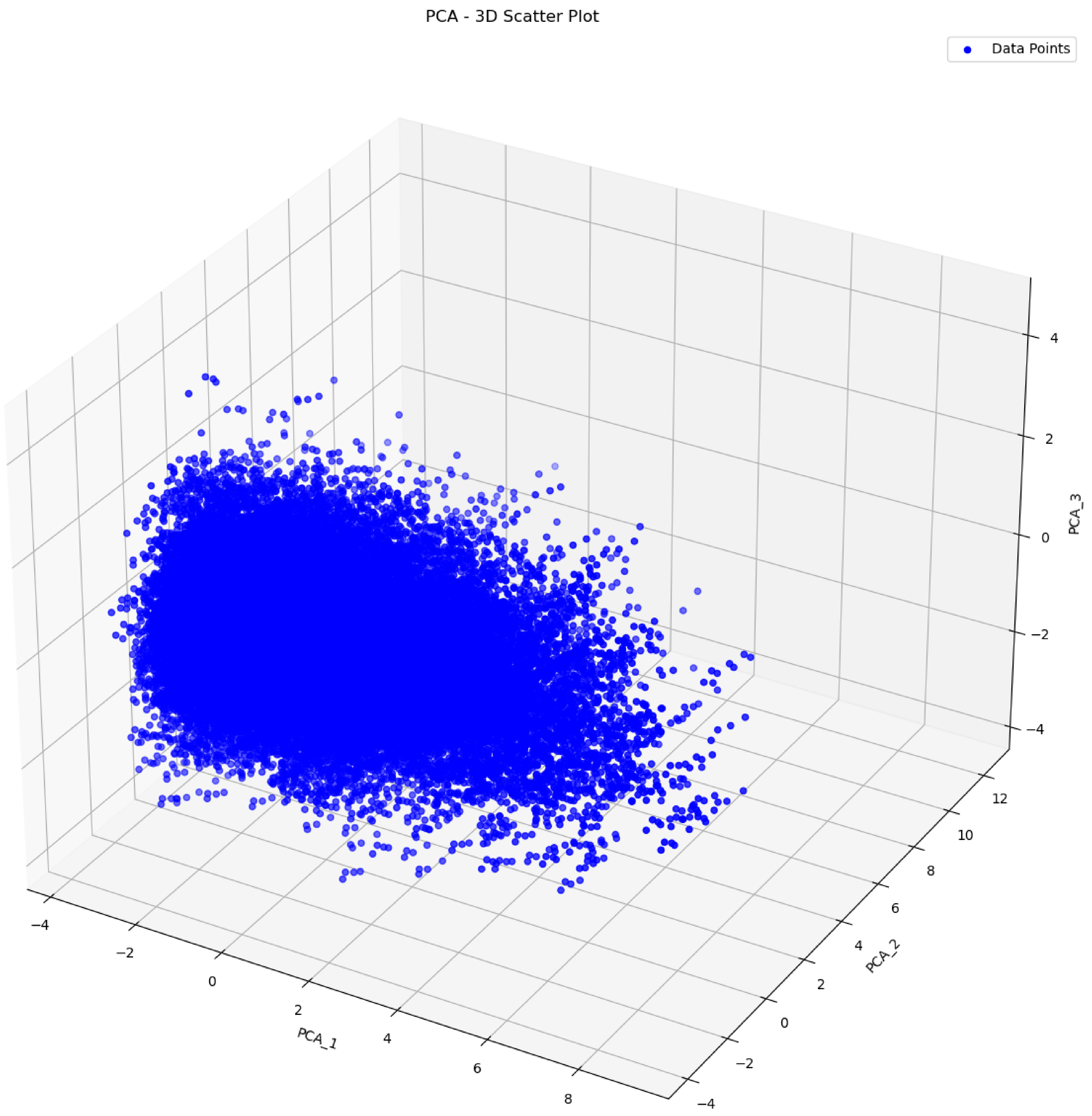
# EDA & VISUALIZATIONS

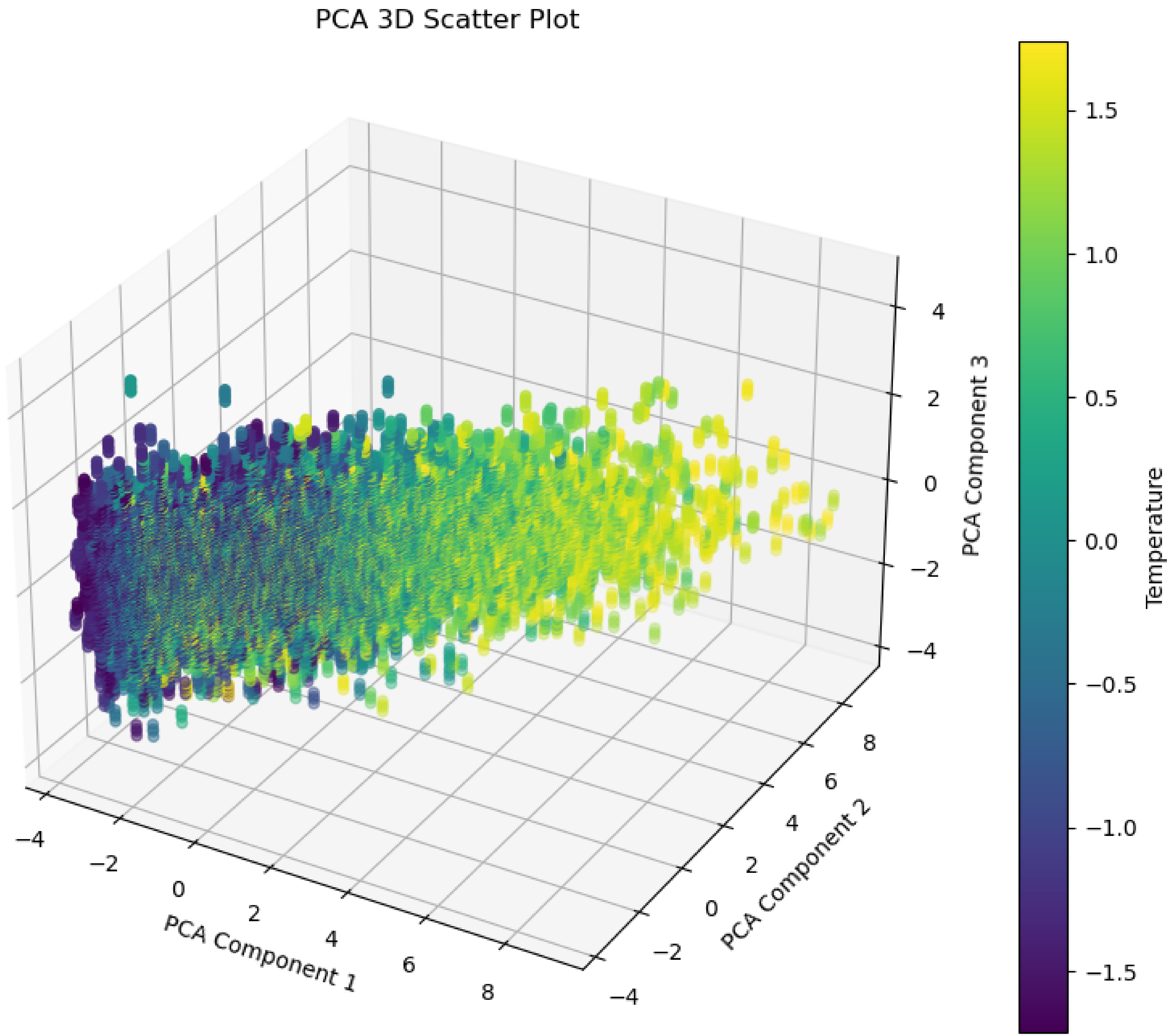
# NAND2

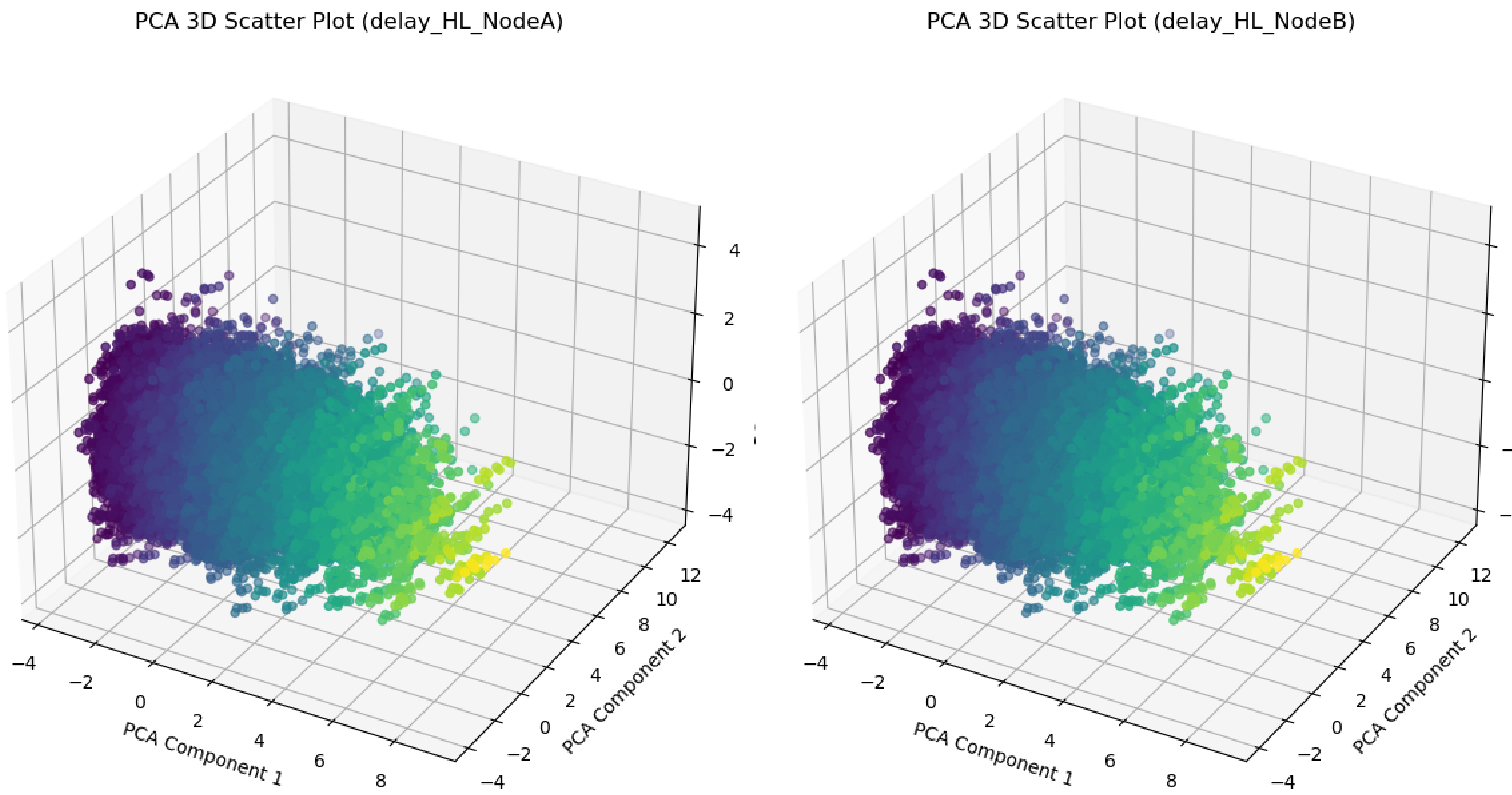


Vin_A
Vin_B
temp
pvdd
cqload
lmin
wmin
toxe_n
toxm_n
toxref_n
toxe_p
toxm_p
toxref_p
toxp_par
xj_n
xj_p
ndep_n
ndep_p
leakage
delay_LH_NodeA
delay_HL_NodeA
delay_LH_NodeB
delay_HL_NodeB
Count
1e−15
1e−8
1e−9
1e−10
1e18
1e−7


Vin_A
Vin_B
temp
pvdd
cqload
lmin
wmin
toxe_n
toxm_n
toxref_n
toxe_p
toxm_p
toxref_p
toxp_par
xj_n
xj_p
ndep_n
ndep_p
leakage
delay_LH_NodeA
delay_HL_NodeA
delay_LH_NodeB
delay_HL_NodeB
Count


PCA - 3D Scatter Plot
Data Points
PCA_1
PCA_2
PCA_3
−4
−2
0
2
4
6
8
10
12


PCA 3D Scatter Plot
PCA Component 1
PCA Component 2
PCA Component 3
Temperature
−4
−2
0
2
4
6
8
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5


PCA 3D Scatter Plot (delay_HL_NodeA)
PCA Component 1
PCA Component 2
PCA 3D Scatter Plot (delay_HL_NodeB)
PCA Component 1
PCA Component 2

# Correlation map of PCA factors and original columns

## PCA_1:

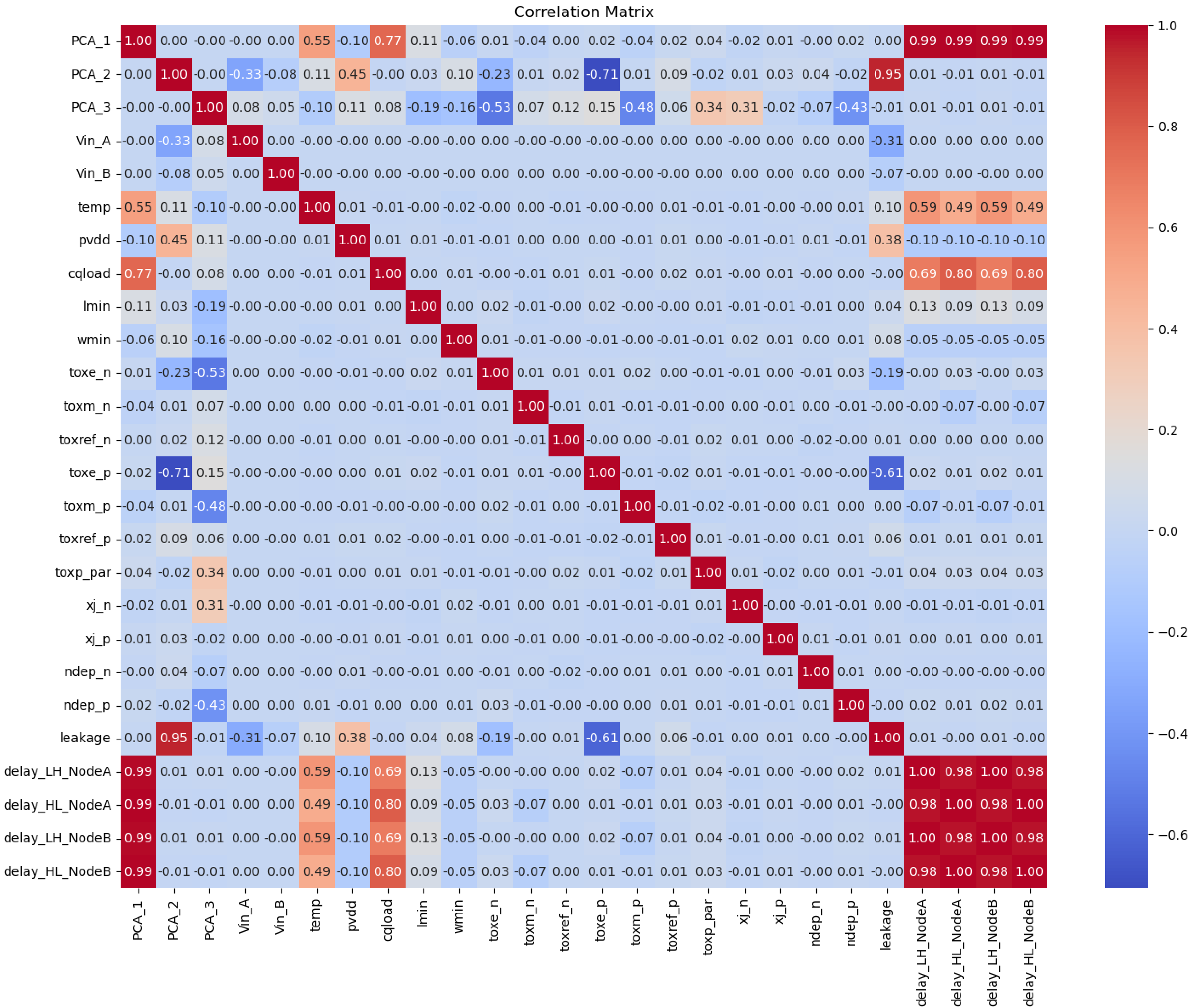
delay_HL_NodeA
delay_HL_NodeB
delay_LH_NodeB
delay_LH_NodeA

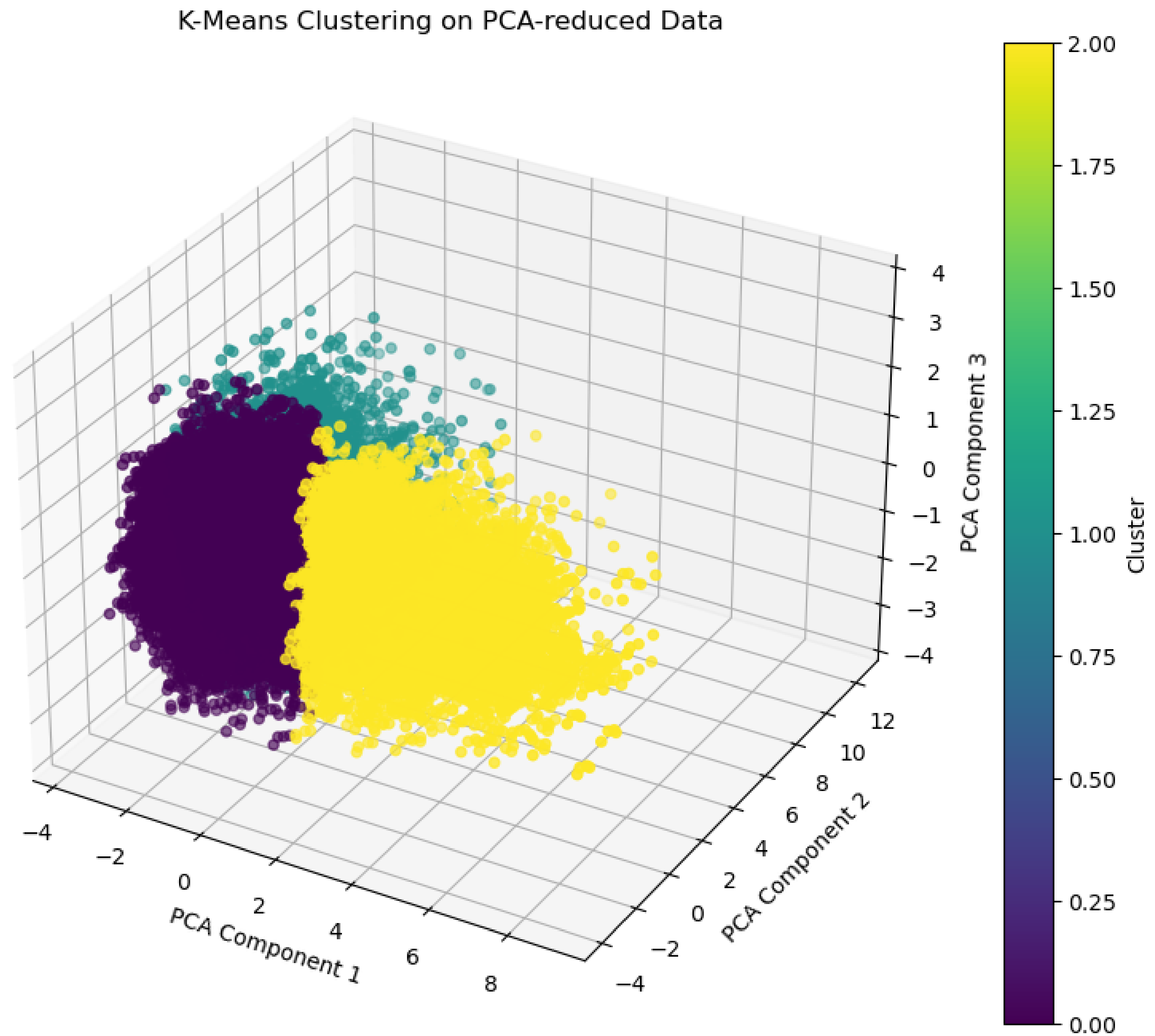
## PCA_2

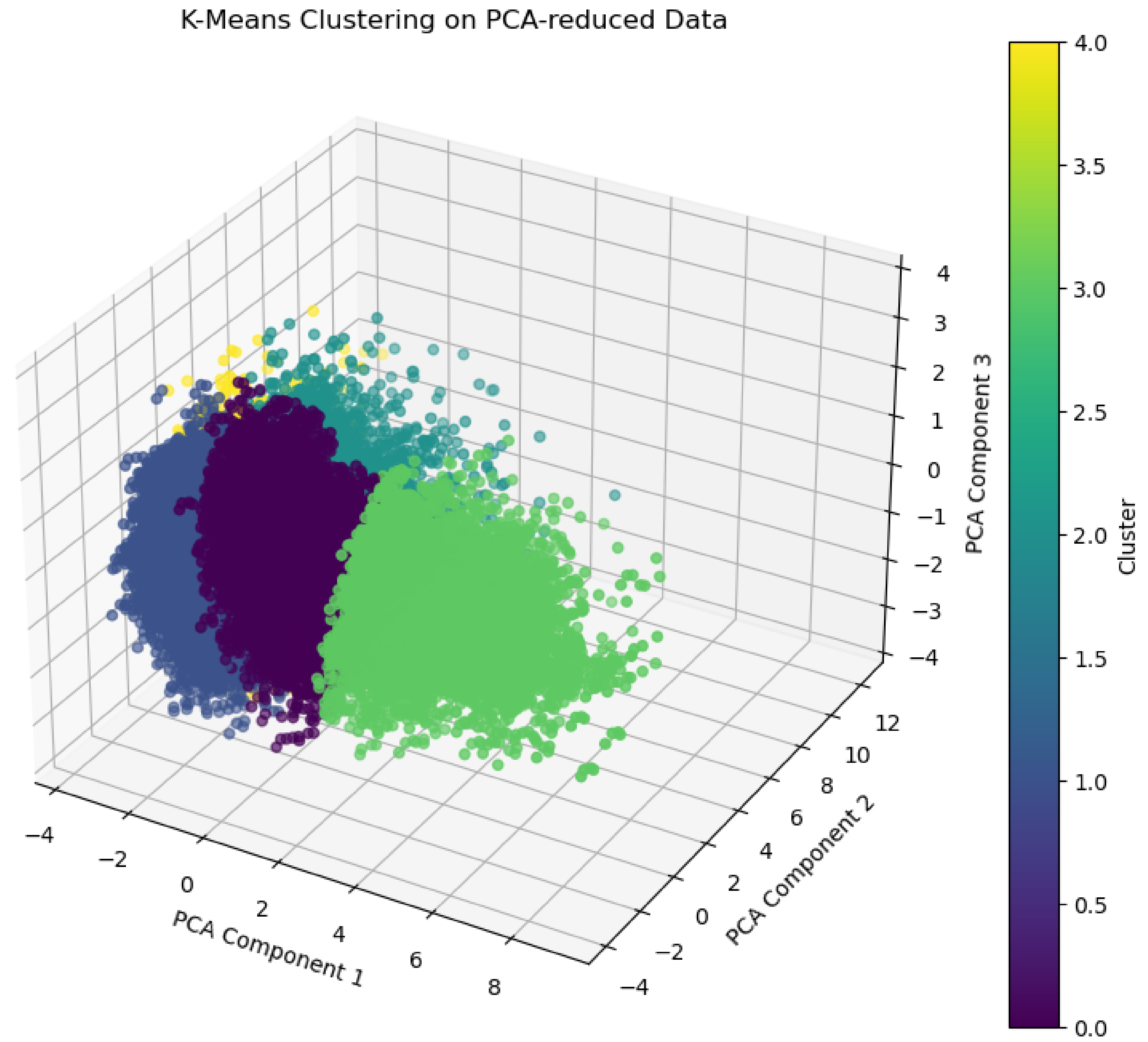
leakage
toxe_p
pvdd
Vin_A

## PCA_3

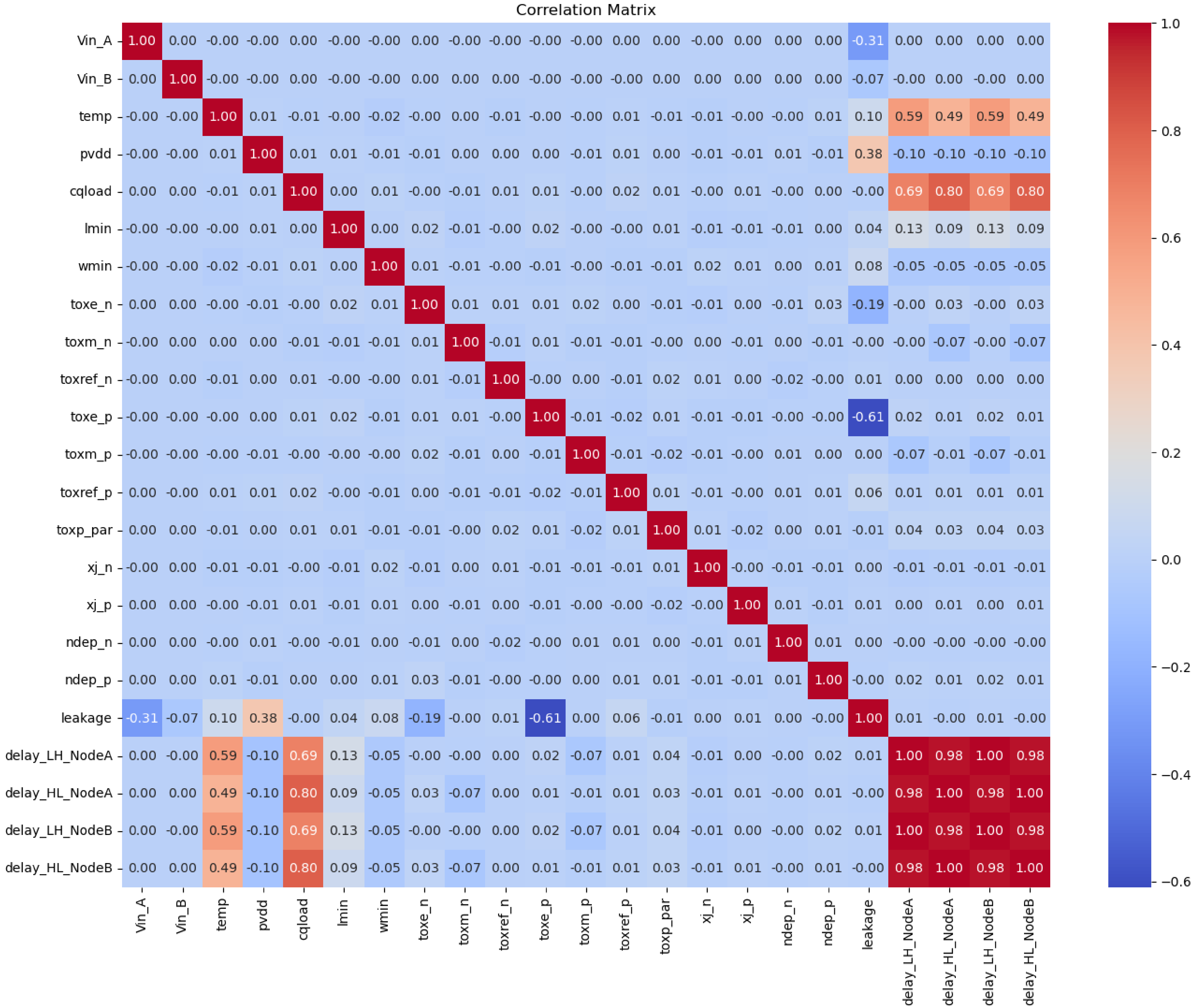
toxe_n
toxm_p
ndep_p
toxp_par



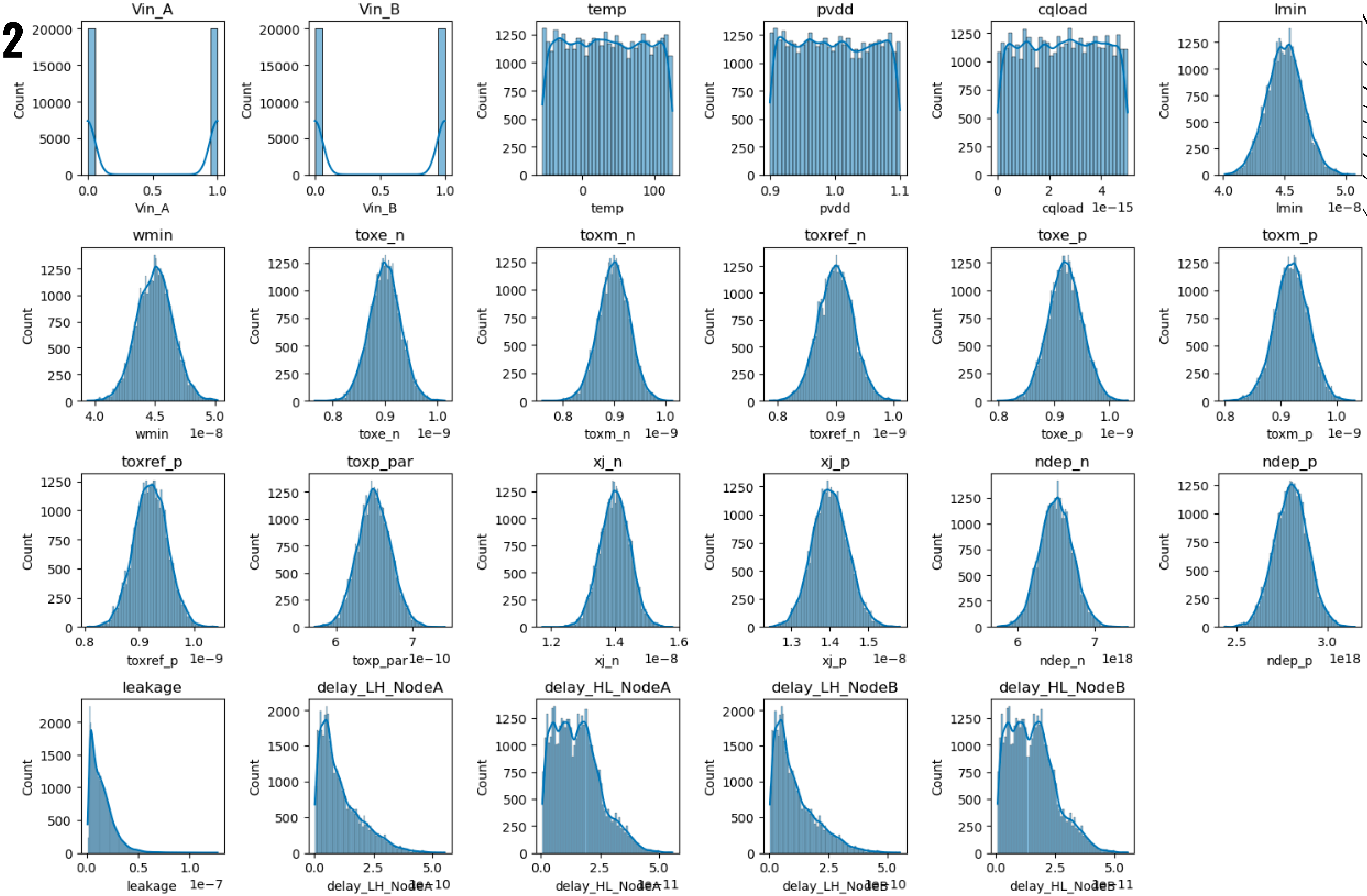

K-Means Clustering on PCA-reduced Data
PCA Component 1
PCA Component 2
PCA Component 3
Cluster
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

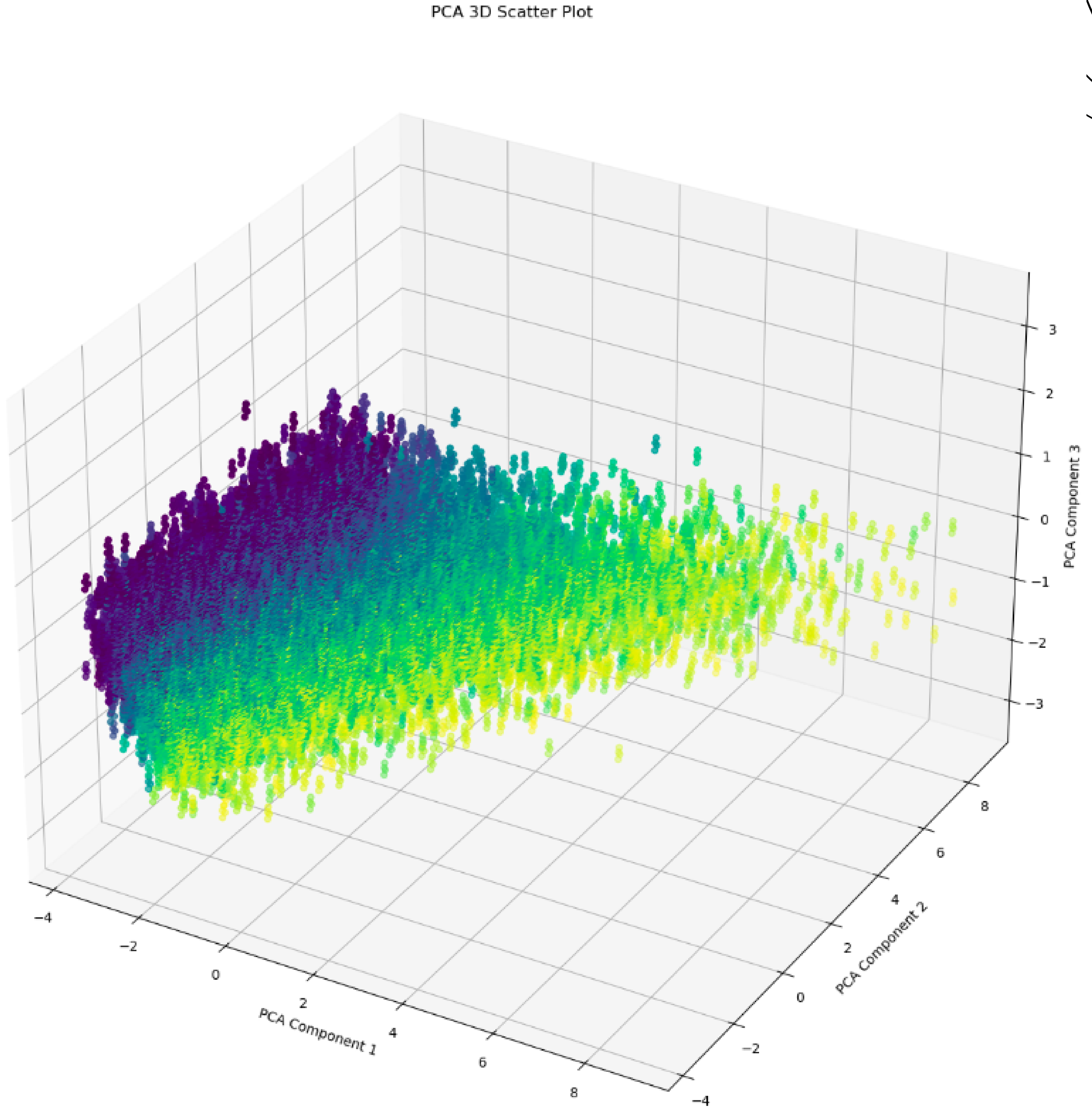

K-Means Clustering on PCA-reduced Data
PCA Component 1
PCA Component 2
PCA Component 3
Cluster
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0

# Correlation map of NAND DATA

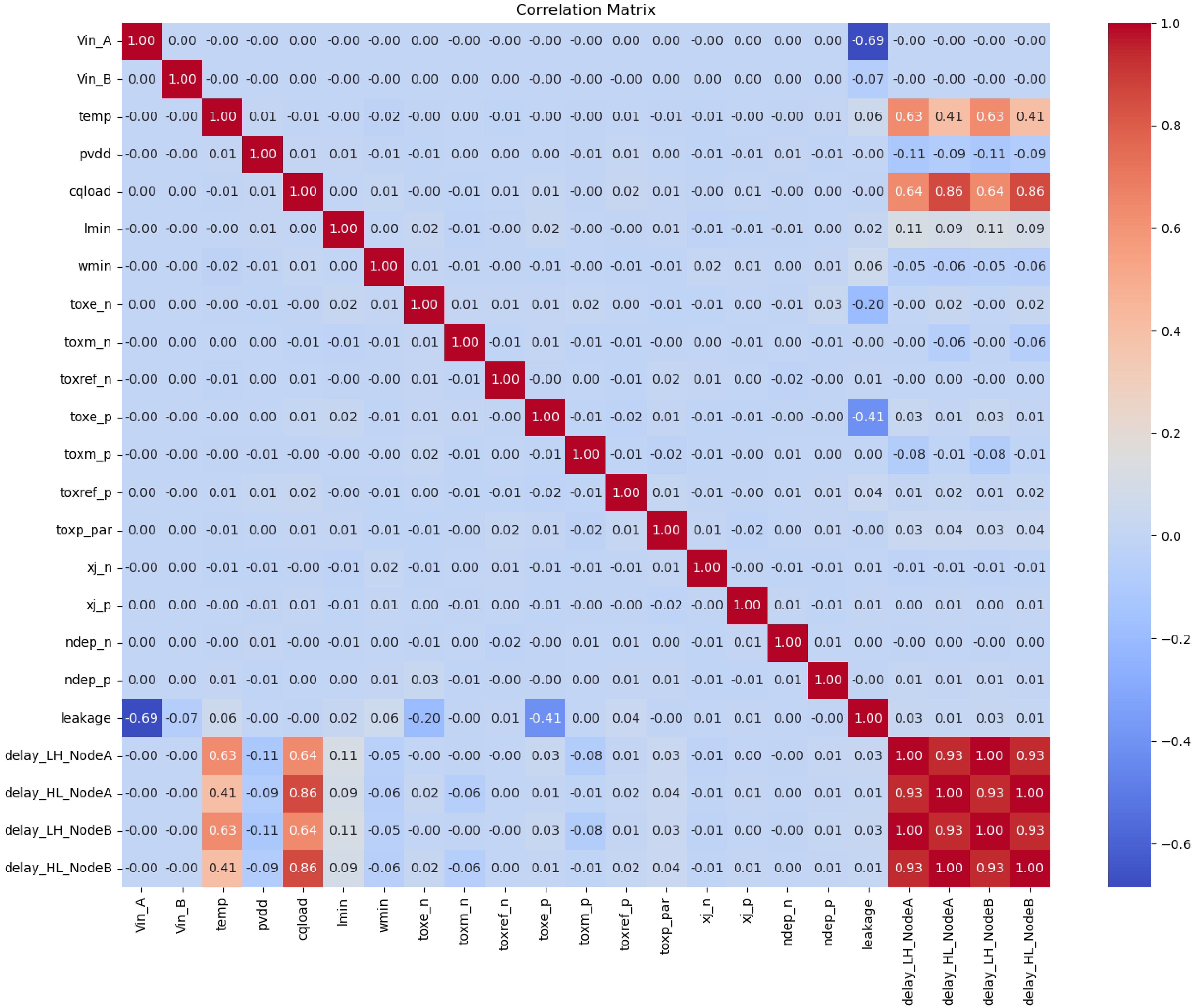


# NOR2

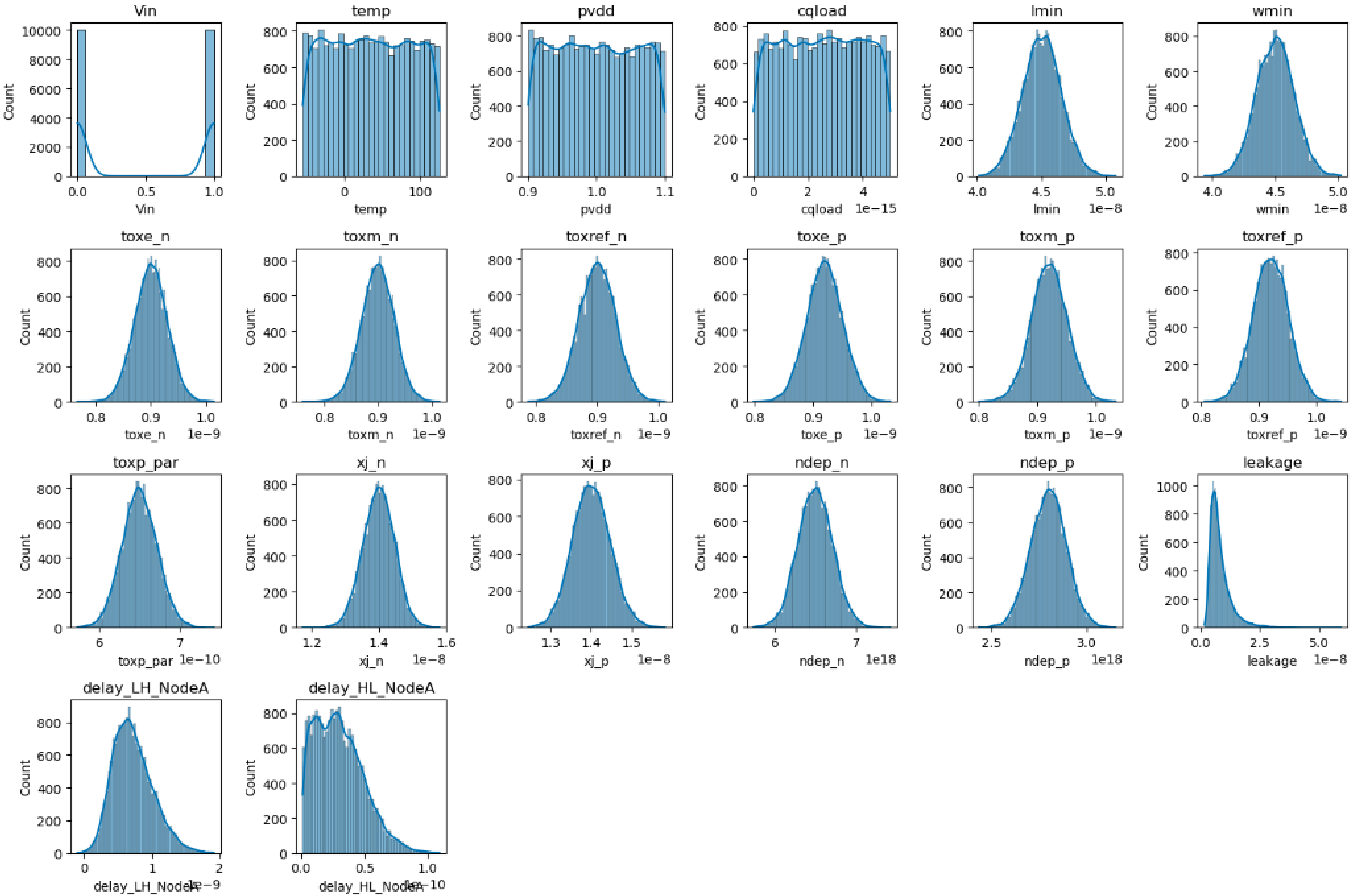

Vin_A
Vin_B
temp
pvdd
cqload
1e−15
lmin
1e−8
wmin
1e−8
toxe_n
1e−9
toxm_n
1e−9
toxref_n
1e−9
toxe_p
1e−9
toxm_p
1e−9
toxref_p
1e−9
toxp_par
1e−10
xj_n
1e−8
xj_p
1e−8
ndep_n
1e18
ndep_p
1e18
leakage
1e−7
delay_LH_NodeA
1e−10
delay_HL_NodeA
1e−11
delay_LH_NodeB
1e−10
delay_HL_NodeB
1e−11
Count


PCA 3D Scatter Plot
PCA Component 1
PCA Component 2
PCA Component 3
−4
−2
0
2
4
6
8
3
2
1
0
−1
−2
−3

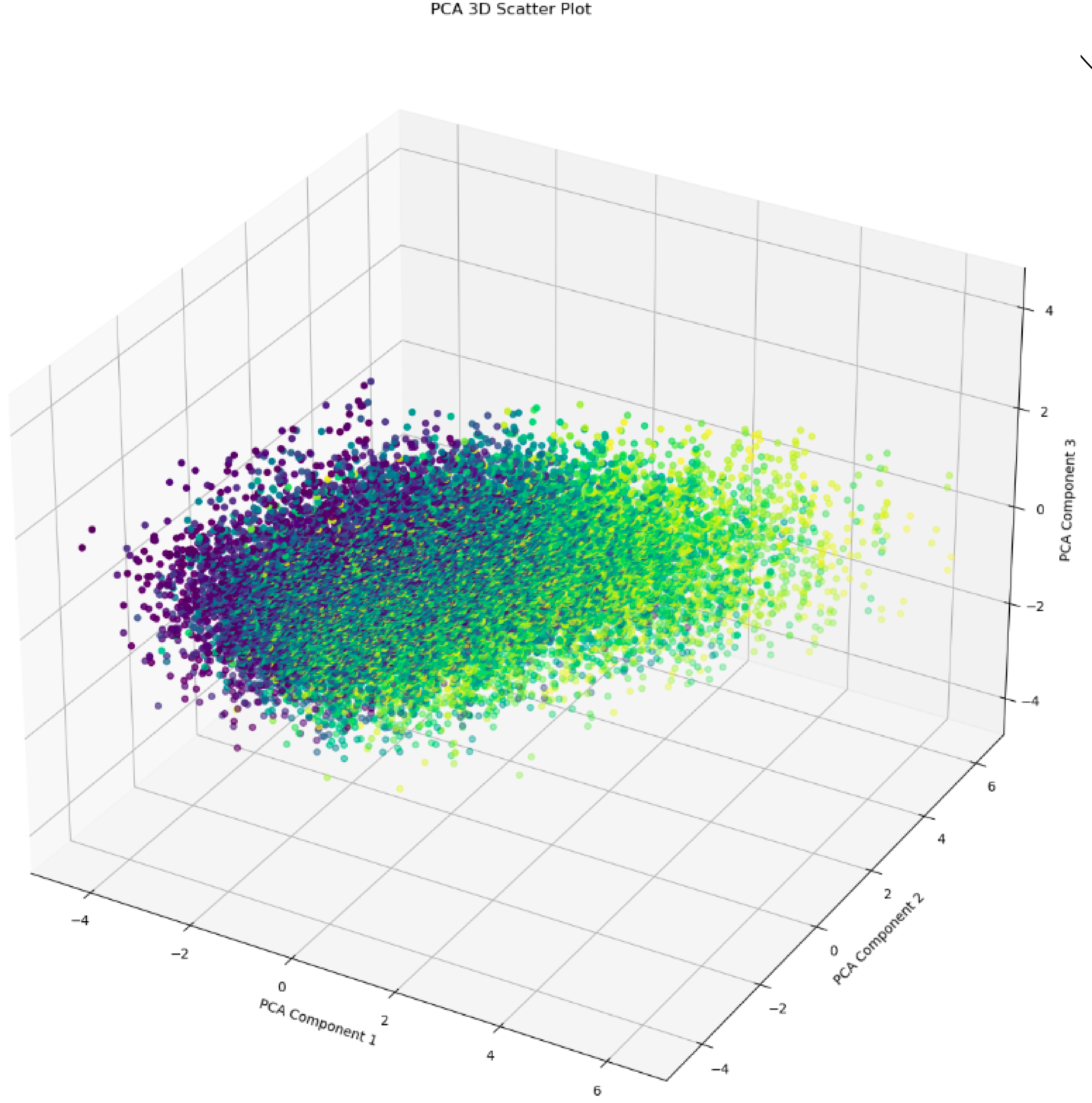
Correlation map of NOR DATA



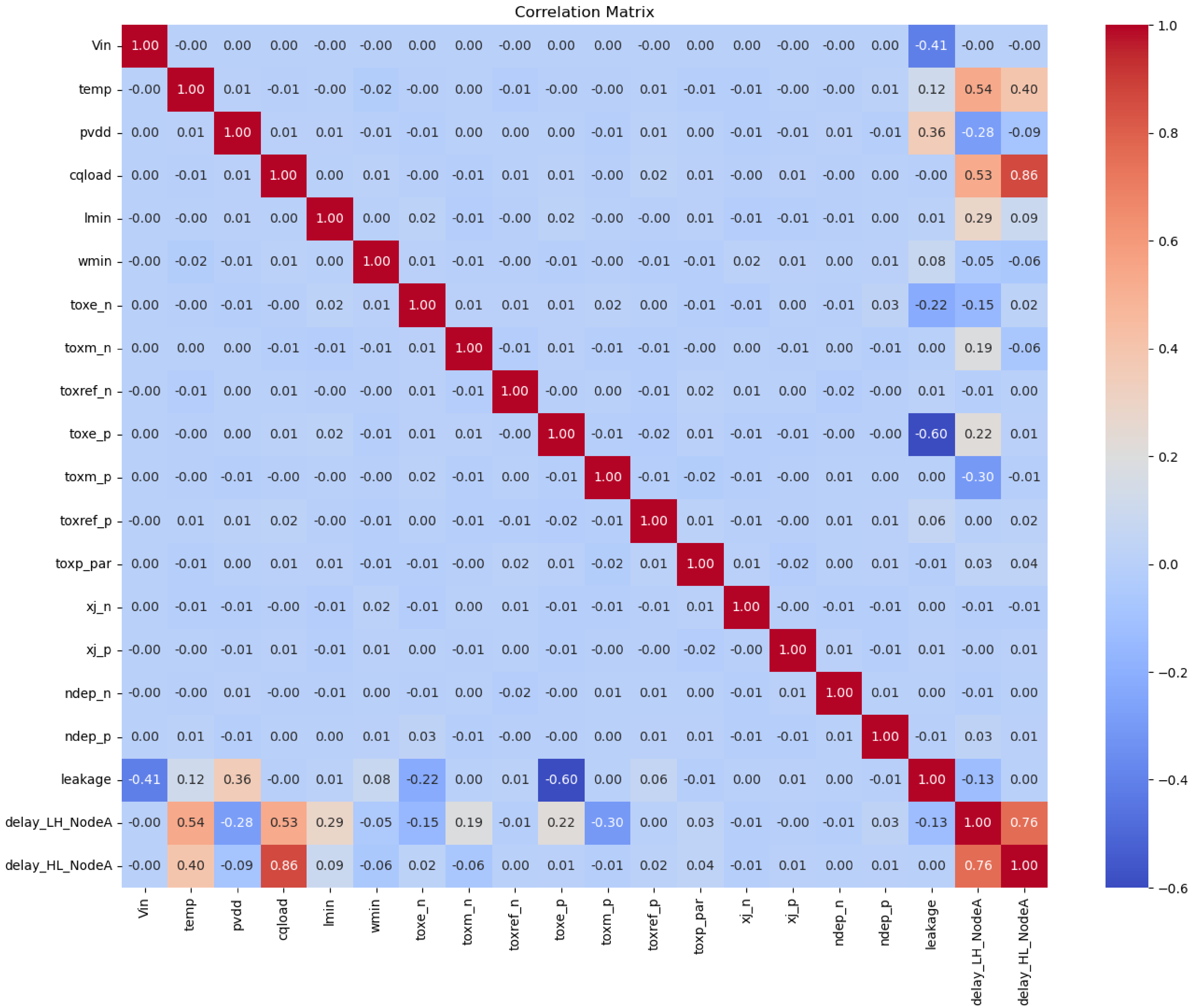
# NOT



Vin
temp
pvdd
cqload
lmin
wmin
toxe_n
toxm_n
toxref_n
toxe_p
toxm_p
toxref_p
toxp_par
xj_n
xj_p
ndep_n
ndep_p
leakage
delay_LH_NodeA
delay_HL_NodeA
Count
1e−15
1e−8
1e−9
1e−10
1e18


PCA 3D Scatter Plot
PCA Component 1
PCA Component 2
PCA Component 3
−4
−2
0
2
4
6

# Correlation map of NOT DATA



# ML PROPOSAL

# Ensemble Learning with Gradient Boosting

- **Effectiveness**: Gradient boosting algorithms like XGBoost, LightGBM, or CatBoost are chosen because they are effective in capturing complex relationships in the data and can handle both numerical and categorical features well.
- **Handling Non-linearity**: The dataset may contain non-linear relationships between input features and target variables (leakage and delay). Gradient boosting can capture these non-linearities effectively through its ensemble of decision trees.
- **Feature Importance**:  Provides feature importance scores, allowing us to identify which features are most influential in predicting leakage and delay.
- **Robustness:** It is robust to outliers and noise in the data, which might be present in real-world datasets like the one under consideration.

- Leakage and delay can be found through the following steps: Model selection, Feature Importance Analysis, Identification of Influential Features, Hyperparameter Tuning

# MODEL SELECTION

- Utilize gradient boosting algorithms such as XGBoost, LightGBM, or CatBoost.
- These algorithms are well-suited for regression tasks and are effective in handling complex relationships in the data.
- By training these models on the dataset, they can learn the underlying patterns and associations between the input features and the target variables (leakage and delay).

# FEATURE IMPORTANCE ANALYSIS

- After training the ensemble models, leverage feature importance scores provided by these models.
- These scores indicate the relative importance of each feature in predicting the target variables (leakage and delay).
- Features with higher importance scores are deemed to have a stronger influence on the prediction, while those with lower scores have less impact.

# IDENTIFICATION OF INFLUENTIAL FEATURES

- Analyze the feature importance scores to identify the most influential features contributing to the prediction of leakage and delay.
- This analysis helps in understanding which features are critical in determining the values of the target variables.

# HYPERPARAMETER TUNING

- Employ techniques like grid search or random search to fine-tune the hyperparameters of the gradient boosting models.
- Hyperparameters are parameters that control the learning process of the model, such as the learning rate, maximum depth of trees, and regularization parameters.
- By optimizing these hyperparameters, the model's performance and generalization capability can be improved, leading to more accurate predictions of leakage and delay.

# THANK YOU