Supplementary Materials

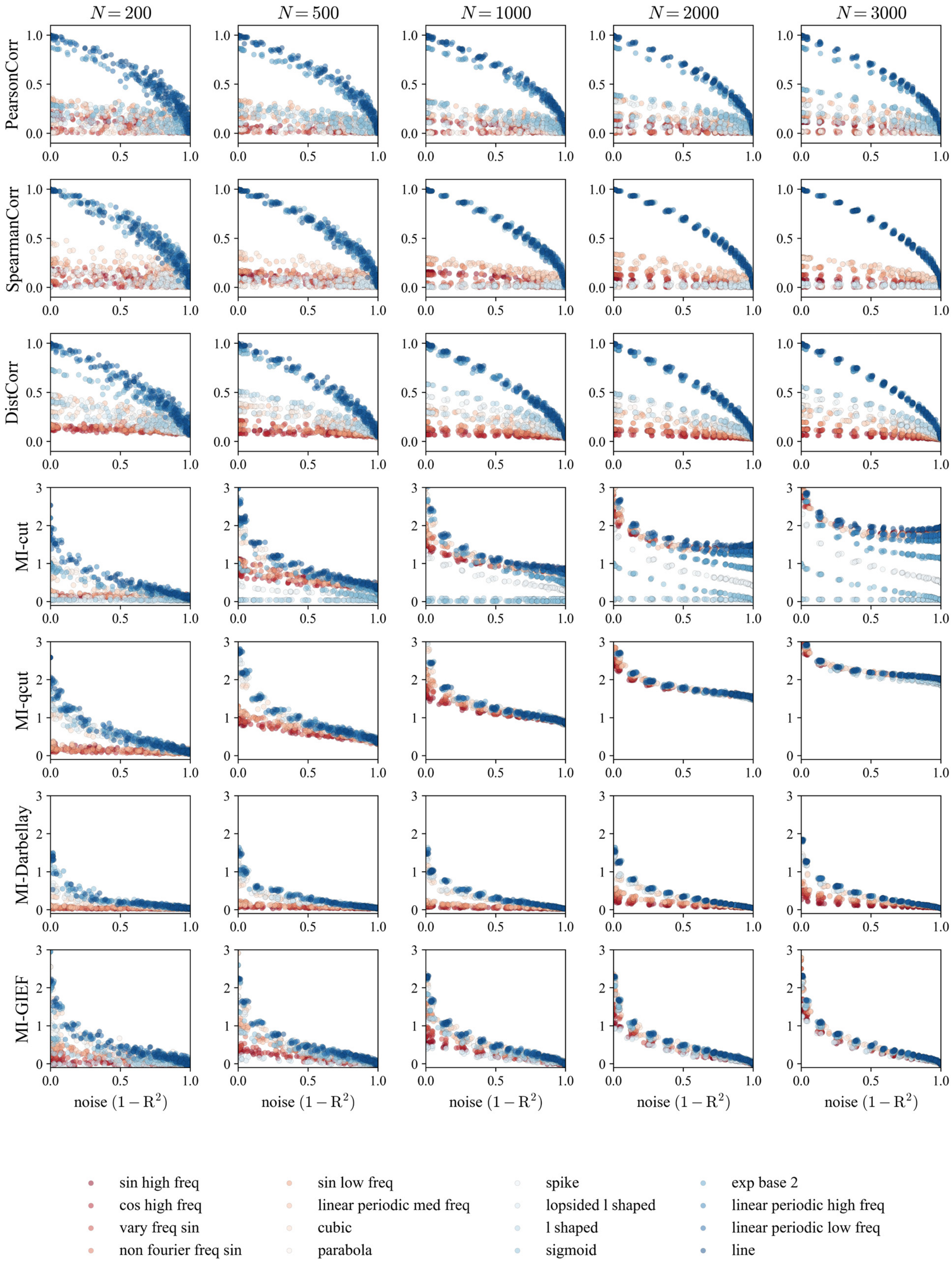

Figure S1. Association values obtained on 16 datasets with increasing noise levels and sample sizes.

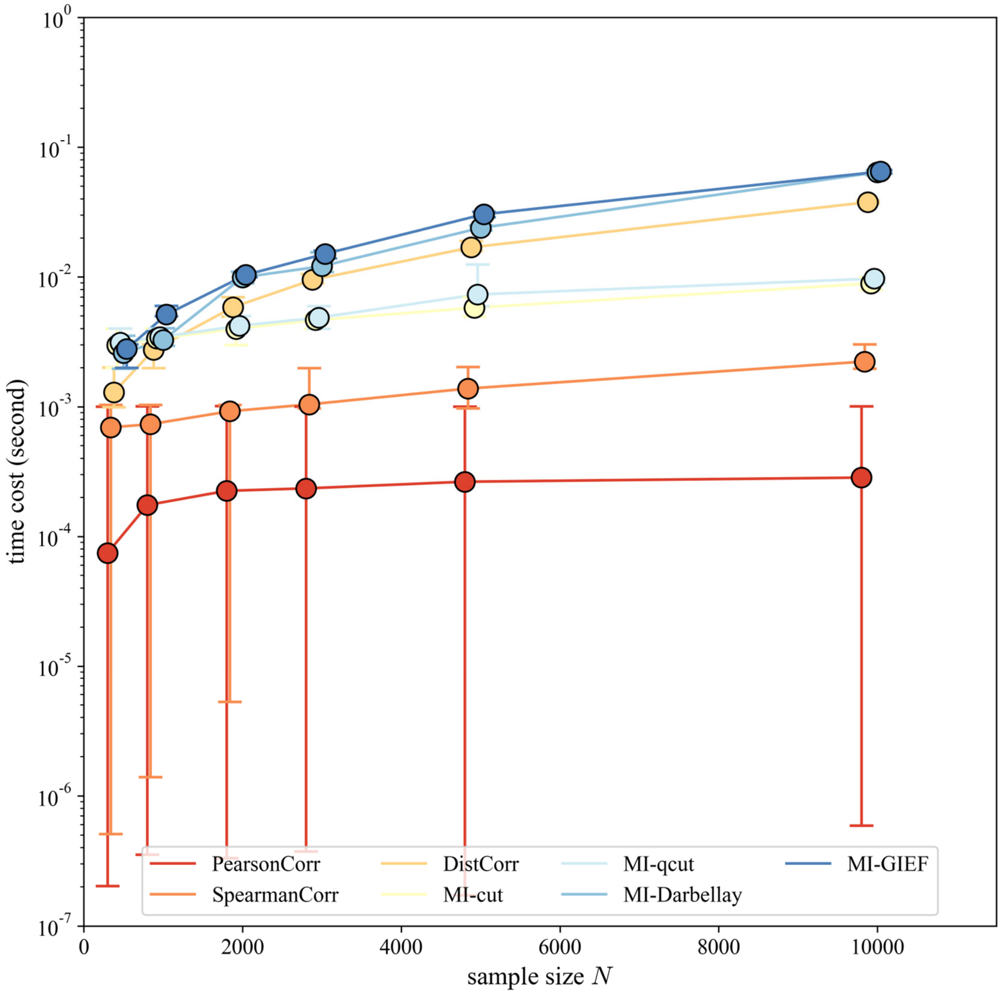

Figure S2. Comparison of time costs for different methods.

Table S1. Algorithm flowsheet of (conditional) independence test based on GIEF.

```
Algorithm: CHECKINDEP
Input:
    $x$: data of $X$
    $y$: data of $Y$
    $z$: data of $Z$, optional
    $\alpha$: significance level
    $rounds$: rounds for generating surrogate samples
Output:
    $is\_indep$: whether $X$ is independent of $Y$ (given $Z$)

    # Compute MI or CMI.
    $v \leftarrow \hat{I}_{\mathrm{g}}(x;y)$ or $\hat{I}_{\mathrm{g}}(x,y|z)$
    # Compute surrogate MI or CMI values.
    $\mathcal{MI}_{\mathrm{surrogate}} \leftarrow \emptyset$
    for $r$ in $rounds$, do
        permute samples of $X$ and get surrogate samples: $x^{\mathrm{s}}$
        compute MI or CMI value $v^{s} \leftarrow \hat{I}_{\mathrm{g}}(x^{s};y)$ or $\hat{I}_{\mathrm{g}}(x^{s},y|z)$ and add it to $\mathcal{MI}_{\mathrm{surrogate}}$:
            $\mathcal{MI}_{\mathrm{surrogate}} \leftarrow \mathcal{MI}_{\mathrm{surrogate}} \cup \{v^{s}\}$
    end for
    # Check independence.
    compute $P$ value:$P \leftarrow \mathrm{card}\left(\left\{m \middle| m \in \mathcal{MI}_{\mathrm{surrogate}}, m > v\right\}\right)/rounds$
    if $P < \alpha$ then
        $is\_indep \leftarrow$ True
    else
        $is\_indep \leftarrow$ False
    # Return the result.
    return $is\_indep$
```

Table S2. Algorithm flowsheet of resolving Markov blanket with CMIM-GIEF.

Algorithm: ResolveMarkovBlanket

```
Input:
    $x_{N\times D_x}$: $D_x$-dimensional data of the variables with sample size = $N$;
    $y_{N\times 1}$: one-dimensional data of the target with sample size = $N$;
    $K$: maximum number of iterations in IVS
    $\mathcal{S}_\mathrm{p}$: preselected variables set
    $\varepsilon$: threshold for terminating the iteration process
Output:
    $\mathcal{S}$: Markov blanket of $Y$

    # Initialize the arguments.
    $\mathcal{F} \leftarrow \mathcal{X} \backslash \mathcal{S}_\mathrm{p}$
    $\mathcal{S} \leftarrow \emptyset$
    $t \leftarrow 1$
    # Excute the IVS procedure for identifying the MB variables.
    for each step $t \le K$, do
        if $t = 1$, do
            for each variable $f \in \mathcal{F}$, do
                estimate MI-GIEF or CMI-G: $\hat{I}_\mathrm{g}(f;Y)$ or $\hat{I}_\mathrm{g}(f;Y|\mathcal{S}_\mathrm{p})$
            end for
            select the first MB variable:
                $S^{(1)} \leftarrow \underset{f\in\mathcal{F}}{\mathrm{argmax}}\, \hat{I}_\mathrm{g}(f;Y)$ or $\underset{f\in\mathcal{F}}{\mathrm{argmax}}\, \hat{I}_\mathrm{g}(f;Y|\mathcal{S}_\mathrm{p})$, $\mathcal{S} \leftarrow \{S^{(1)}\}$
        else if $t > 1$, do
            $\mathcal{R}^{(t)} \leftarrow \mathcal{S}_\mathrm{p} \cup \mathcal{S}^{(t-1)}$
            for each variable $f \in \mathcal{F} \backslash \mathcal{R}^{(t)}$, do
                for each variable $r \in \mathcal{R}^{(t)}$, do
                    extract data of $f$ and $r$ from $x$ and compute CMI-G: $\hat{I}_\mathrm{g}(f;Y|r)$
                end for
            end for
            if $m^{(t)} < \varepsilon$ then
                break
            else
                select the MB variable at step $t$:
                    $S^{(t)} \leftarrow \underset{f\in\mathcal{F}-\mathcal{S}^{(t-1)}}{\arg\max} \left\{ \min_{r\in\mathcal{R}^{(t)}} \hat{I}_\mathrm{g}(f;Y|r) \right\}$, $\mathcal{S} \leftarrow \mathcal{S} \cup \{S^{(t)}\}$
    end for
    # Return the results.
    return $\mathcal{S} \cup \mathcal{S}_\mathrm{p}$
```

Table S3. Algorithm flowsheet of differentiating variables.

```
Algorithm: DifferentiateVariables
Input:
    $x_{N\times D_x}$: $D_x$-dimensional data of variables with sample size = $N$;
    $y_{N\times 1}$: one-dimensional data of target with sample size =$N$
    $\mathcal{M}_Y$: Markov blanket of $Y$
Output:
    $\mathcal{X}^{\mathrm{sa}}$: set of variables strongly associated with $Y$;
    $\mathcal{X}^{\mathrm{ia}}$: set of variables interactively associated with $Y$;
    $\mathcal{X}^{\mathrm{ra}}$: set of variables redundantly associated with $Y$;
    $\mathcal{X}^{\mathrm{ir}}$: set of variables irrelevant with $Y$

    # Initialize sets of variables.
    $\mathcal{X}^{\mathrm{sa}} \leftarrow \emptyset$
    $\mathcal{X}^{\mathrm{ia}} \leftarrow \emptyset$
    $\mathcal{X}^{\mathrm{ra}} \leftarrow \emptyset$
    $\mathcal{X}^{\mathrm{ir}} \leftarrow \emptyset$
    # Execute variable differentiation.
    for each variable $X_i \in \mathcal{X}$, do
        if $X_i \in \mathcal{M}_Y$, then
            # Discriminate strong or interactional associated variables.
            if CheckIndep$(X_i; Y)$ = True then
                $\mathcal{X}^{\mathrm{ia}} \leftarrow \mathcal{X}^{\mathrm{ia}} \cup \{X_i\}$
            else
                $\mathcal{X}^{\mathrm{sa}} \leftarrow \mathcal{X}^{\mathrm{sa}} \cup \{X_i\}$
        else
            # Discriminate redundant or irrelevant variables.
            if CheckIndep$(X_i; Y)$ = True then
                $\mathcal{X}^{\mathrm{ir}} \leftarrow \mathcal{X}^{\mathrm{ir}} \cup \{X_i\}$
            else
                $\mathcal{X}^{\mathrm{ra}} \leftarrow \mathcal{X}^{\mathrm{ra}} \cup \{X_i\}$
    end for
    # Return the results.
    return $\mathcal{X}^{\mathrm{sa}}, \mathcal{X}^{\mathrm{ia}}, \mathcal{X}^{\mathrm{ra}}, \mathcal{X}^{\mathrm{ir}}$
```

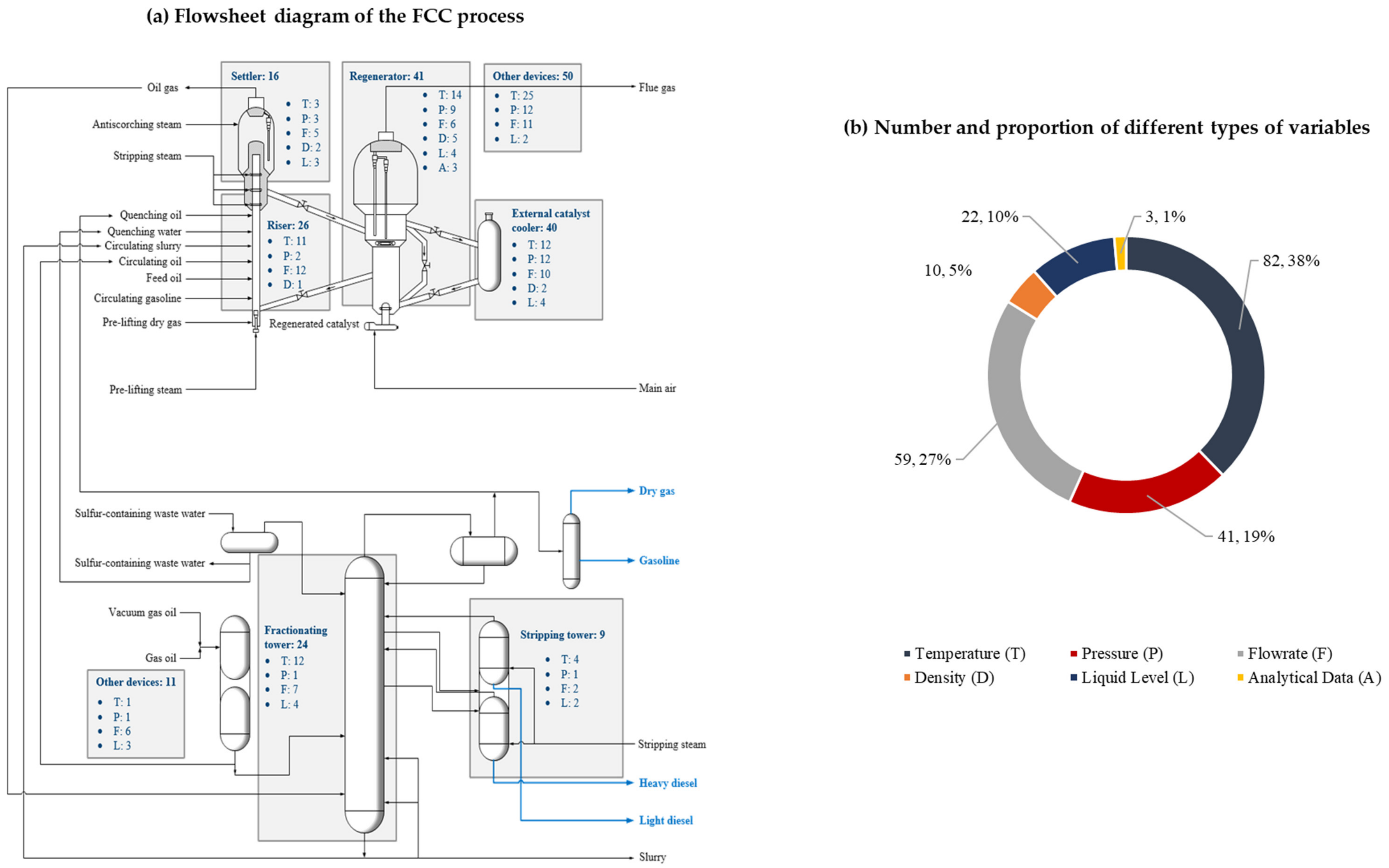

Figure S3. (a) Flowsheet diagram of the FCC process; **(b)** Numbers and percentages of different types of variables and targets in the process.

Table S4. Prediction effects of different models with full features on the four targets (the best results are bolded and underlined).

target	model	MAE	MSE	R^2	MAPE
Y_{LD}	RF	2.01E-02	**1.01E-03**	**0.985**	2.12E-01
	KNN	1.90E-02	1.10E-03	0.984	2.14E-01
	LightGBM	1.93E-02	1.11E-03	0.984	2.01E-01
	LR	2.38E-02	1.31E-03	0.981	**2.54E-01**
	XGBoost	**1.85E-02**	1.35E-03	0.980	2.08E-01
	Ridge	2.79 E-02	1.52E-03	0.977	2.94E-01
	MLP	4.69 E-02	4.01E-03	0.939	4.65E-01
	Lasso	2.46 E-01	6.61E-02	-4.01E-04	2.23E+00
Y_{HD}	RF	1.78E-02	**9.00E-04**	**0.991**	6.88E-02
	KNN	1.83E-02	1.00E-03	0.989	7.79E-02
	LightGBM	1.85E-02	1.10E-03	0.988	7.62E-02
	XGBoost	**1.77E-02**	1.20E-03	0.987	**6.38E-02**
	Ridge	3.05E-02	1.80E-03	0.981	1.34E-01
	MLP	4.71E-02	4.10E-03	0.956	1.85E-01
	LR	2.79E-02	5.80E-03	0.937	1.14E-01
	Lasso	2.95E-01	9.23E-02	-0.005	1.34E+00
Y_{GAS}	XGBoost	**1.17E-02**	**5.00E-04**	**0.948**	**4.20E-02**
	RF	1.26E-02	6.00E-04	0.945	7.25E-02
	LightGBM	1.28E-02	6.00E-04	0.943	5.91E-02
	LR	2.12E-02	9.00E-04	0.918	9.46E-02
	KNN	1.32E-02	1.00E-03	0.909	6.76E-02
	Ridge	2.17E-02	1.10E-03	0.895	1.24E-01
	MLP	4.62E-02	4.30E-03	0.573	2.44E-01
	Lasso	7.76E-02	1.04E-02	-0.001	3.42E-01
Y_{DG}	RF	**2.24E-02**	**1.50E-03**	**0.910**	**1.59E-01**
	LightGBM	2.33E-02	1.70E-03	0.901	1.65E-01
	XGBoost	2.33E-02	1.70E-03	0.899	1.61E-01
	KNN	2.47E-02	1.80E-03	0.898	1.65E-01
	LR	3.22E-02	2.30E-03	0.865	1.82E-01
	Ridge	3.19E-02	2.30E-03	0.863	1.76E-01
	MLP	4.86E-02	4.40E-03	0.745	1.96E-01
	Lasso	9.00E-02	1.72E-02	-0.003	3.15E-01

Table S5. Average prediction metrics obtained by different data-dimensionality reduction methods with RF on the four targets (the best results are bolded and underlined).

method	$\boldsymbol{Y_{\mathrm{LD}}}$				$\boldsymbol{Y_{\mathrm{HD}}}$				$\boldsymbol{Y_{\mathrm{GSL}}}$				$\boldsymbol{Y_{\mathrm{DG}}}$			
	R²	**MAE**	**MSE**	**MAPE**	**R²**	**MAE**	**MSE**	**MAPE**	**R²**	**MAE**	**MSE**	**MAPE**	**R²**	**MAE**	**MSE**	**MAPE**
total	0.984	2.06E-02	1.04E-03	20.2%	0.989	1.98E-02	1.05E-03	8.8%	0.933	1.30E-02	6.91E-04	12.5%	0.913	2.27E-02	1.43E-03	11.8%
CMIM-GIEF	**0.985**	**1.98E-02**	**9.68E-04**	18.8%	**0.990**	1.91E-02	9.64E-04	8.4%	**0.942**	**1.25E-02**	**5.92E-04**	**9.4%**	0.911	2.25E-02	1.46E-03	11.8%
CMIM-Q	0.984	2.15E-02	1.08E-03	21.3%	**0.990**	1.91E-02	**9.39E-04**	8.8%	0.934	1.27E-02	6.82E-04	12.0%	0.907	2.35E-02	1.53E-03	12.1%
Pearson	0.981	2.26E-02	1.24E-03	21.5%	0.987	2.07E-02	1.16E-03	9.1%	0.920	1.42E-02	8.25E-04	13.4%	0.904	2.40E-02	1.58E-03	12.2%
Spearman	0.983	2.09E-02	1.09E-03	20.5%	0.987	2.10E-02	1.21E-03	9.0%	0.920	1.45E-02	8.17E-04	13.2%	0.904	2.34E-02	1.58E-03	12.4%
DistCorr	0.983	2.10E-02	1.11E-03	20.1%	0.988	2.07E-02	1.15E-03	8.9%	0.922	1.43E-02	8.03E-04	13.0%	0.909	2.29E-02	1.49E-03	12.0%
MI	**0.985**	2.03E-02	1.02E-03	19.9%	0.989	1.98E-02	1.06E-03	8.5%	0.934	1.27E-02	6.78E-04	12.8%	0.917	**2.21E-02**	1.37E-03	11.7%
MDI	**0.985**	1.99E-02	9.78E-04	20.1%	**0.990**	**1.89E-02**	9.63E-04	**8.3%**	0.936	1.28E-02	6.53E-04	11.6%	0.916	2.24E-02	1.38E-03	**11.4%**
MDA	0.984	2.03E-02	1.04E-03	20.4%	0.989	1.95E-02	1.04E-03	8.4%	0.935	1.27E-02	6.67E-04	11.9%	**0.918**	2.22E-02	**1.35E-03**	11.6%
GA	0.979	2.28E-02	1.38E-03	21.2%	0.987	2.12E-02	1.16E-03	9.4%	0.929	1.35E-02	7.28E-04	12.8%	0.905	2.37E-02	1.55E-03	11.8%
Lasso	0.983	2.19E-02	1.12E-03	21.5%	0.989	2.00E-02	1.01E-03	8.9%	0.837	2.15E-02	1.66E-03	19.7%	0.883	2.72E-02	1.92E-03	13.0%
Ridge	0.972	2.80E-02	1.86E-03	25.4%	0.989	1.95E-02	1.00E-03	8.7%	0.921	1.49E-02	8.04E-04	13.7%	0.916	2.23E-02	1.38E-03	11.7%
PCA	0.957	3.19E-02	2.81E-03	30.8%	0.969	3.21E-02	2.83E-03	14.3%	0.874	1.86E-02	1.29E-03	16.0%	0.844	3.07E-02	2.57E-03	14.6%
KPCA	0.958	3.14E-02	2.78E-03	28.5%	0.970	3.22E-02	2.77E-03	14.0%	0.879	1.84E-02	1.24E-03	14.5%	0.847	3.02E-02	2.52E-03	14.5%
LLE	0.978	2.24E-02	1.43E-03	**18.2%**	0.986	2.03E-02	1.32E-03	8.8%	0.893	1.46E-02	1.10E-03	18.7%	0.872	2.63E-02	2.10E-03	12.7%
PLS	0.969	2.92E-02	2.02E-03	27.6%	0.980	2.83E-02	1.87E-03	13.2%	0.920	1.70E-02	8.15E-04	11.5%	0.889	2.64E-02	1.82E-03	13.0%