Supporting Information

for *J Water Environ Nanotechnol* 2026, 11(SI)

DOI: 10.22090/jwent.2026.2083999.2061

Sustainable Bioremediation of Hydrocarbon-Contaminated Soils Using ANN and RSM with Enhanced Feature Engineering and Data Augmentation

Hyginus Obinna Ozioko* , Emmanuel Ebube Eze, Ogumere, Henry Ihechi

Department of Civil Engineering, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

In the format provided by the authors and unedited

APPENDIX

Table 9: Preprocessed and Augmented Geotechnical Dataset (PAGD)

Oil Content (%)	Time (Weeks)	MDD (kg/m^3)	OMC (%)	CBR Value (%)	Permeability (k, cm/s)	Initial TPH (Mg/kg)	Final TPH (Mg/kg)	TPH Removal (%)	Microbial Count (CFU/g)	TPH Reduction Rate(mg/kg/week)	Microbial Efficiency	Normalized CBR
0	0	1850	14.5	12.5	8.50E-05	0	0	0	0	0	0	0.006757
3	0	1785	16.5	9.7	4.90E-05	32540	32540	0	0	0	0	0.005434
3	4	1802	16.1	10.2	5.70E-05	32540	11850	63.6	4500000	5172.499	1.41E-05	0.00566
3	8	1818	15.6	10.9	7.00E-05	32540	6120	81.2	88000000	3302.5	9.23E-07	0.005996
3	12	1835	15.2	11.7	7.80E-05	32540	2410	92.6	1.2E+08	2510.833	7.72E-07	0.006376
5	0	1715	18.2	7.3	3.00E-05	56120	56120	0	0	0	0	0.004257
5	4	1735	17.6	8.1	4.30E-05	56120	26590	52.6	4000000	7382.498	1.31E-05	0.004669
5	8	1758	17	9	5.60E-05	56120	13540	75.9	81000000	5322.499	9.37E-07	0.005119
5	12	1785	16.3	9.9	6.70E-05	56120	5180	90.8	1.1E+08	4245	8.25E-07	0.005546
7	0	1640	19.8	5	1.70E-05	81870	81870	0	0	0	0	0.003049
7	4	1665	19.1	5.9	2.60E-05	81870	44250	45.9	3300000	9404.998	1.39E-05	0.003544
7	8	1690	18.5	7.2	3.80E-05	81870	23810	70.9	68000000	7257.499	1.04E-06	0.00426
7	12	1720	17.7	8.6	5.30E-05	81870	11540	85.9	98000000	5860.833	8.77E-07	0.005
10	0	1575	21.3	2.9	8.00E-06	110230	110230	0	0	0	0	0.001841
10	4	1595	20.8	4.1	1.40E-05	110230	63550	42.4	2800000	11670	1.51E-05	0.002571
10	8	1620	20.1	5.6	2.30E-05	110230	38720	64.9	55000000	8938.749	1.18E-06	0.003457
10	12	1655	19.3	6.9	3.60E-05	110230	18550	83.2	91000000	7639.999	9.14E-07	0.004169
0	0	1866.298	13.84511	11.89831	8.71E-05	0	0	0	0	0	0	0.006429
3	0	1821.164	17.08281	9.300699	4.90E-05	31153.28	32613.16	0	0	0	0	0.005387
3	4	1808.308	15.46342	10.2238	5.94E-05	31765.14	11438.46	61.60399	4385124	4944.62	1.45E-05	0.005686
3	8	1975.58	15.58259	10.82503	7.04E-05	34343.96	6104.182	76.22806	91843685	3290.582	9.49E-07	0.006093
3	12	1797.395	15.53009	11.77405	7.97E-05	32702.74	2360.037	89.4163	1.17E+08	2525.663	7.72E-07	0.006376
5	0	1664.406	18.69223	7.306923	3.02E-05	54319.98	56093.18	0	0	0	0	0.004047
5	4	1771.103	18.69021	8.208356	4.38E-05	54969.87	25917.4	53.77317	3948281	7952.53	1.38E-05	0.004694
5	8	1777.85	16.38732	9.416826	5.57E-05	55123.87	13305.27	79.74293	83671861	5099.339	9.42E-07	0.005171
5	12	1781.486	16.98975	10.47431	7.07E-05	55101.47	5081.109	89.66564	1.07E+08	4048.164	8.47E-07	0.005264
7	0	1720.899	20.25014	4.986894	1.65E-05	82618.47	81453.87	0	0	0	0	0.002959
7	4	1545.094	18.98201	5.81052	2.67E-05	81987.72	42271.34	42.96632	3141518	8866.71	1.43E-05	0.003636
7	8	1641.468	18.29865	7.183074	3.70E-05	82099.34	23475.12	66.93542	68611929	7283.103	1.03E-06	0.004326
7	12	1653.013	18.04384	8.468065	5.39E-05	77827.31	11842.34	89.1396	1.01E+08	5992.82	9.04E-07	0.004629
10	0	1550.885	21.00274	2.795536	8.03E-06	110484.4	112614.3	0	0	0	0	0.001871
10	4	1665.556	19.53646	4.294203	1.44E-05	111492.8	64487.58	42.74398	2709051	11320.01	1.53E-05	0.002649
10	8	1602.111	18.21974	5.722642	2.29E-05	101439.5	38210.98	66.59137	54686586	8791.515	1.23E-06	0.00337
10	12	1603.313	18.56964	6.858488	3.57E-05	107313.4	19538.72	84.39157	92321212	7679.15	9.21E-07	0.00427
0	0	1781.427	14.14348	11.99045	8.88E-05	0	0	0	0	0	0	0.006741
3	0	1793.435	16.3099	9.843965	5.00E-05	32041.63	33335.74	0	0	0	0	0.005487
3	4	1804.269	16.23002	10.5516	5.58E-05	33831.66	11698.24	63.66458	4374480	5348.992	1.37E-05	0.005792
3	8	1860.171	15.87012	10.5563	7.35E-05	33115.76	6142.478	80.95443	88480133	3235.66	8.77E-07	0.005646
3	12	1876.58	14.3716	11.85144	7.73E-05	33148.84	2389.286	93.56169	1.19E+08	2526.831	7.61E-07	0.00643

5	0	1723.575	18.20951	7.161585	3.03E-05	53798.71	55851.49	0	0	0	0	0.004302
5	4	1852.403	18.07614	8.286222	4.24E-05	59526.49	25820.85	53.27902	4066490	7634.263	1.30E-05	0.004599
5	8	1728.592	17.9363	9.446252	5.31E-05	55980.75	13073.22	76.7866	78352145	5330.25	9.76E-07	0.005166
5	12	1807.764	15.93868	9.876158	6.49E-05	54473.71	5350.912	83.25335	1.09E+08	4123.673	8.11E-07	0.005505
7	0	1585.181	20.6173	5.0528	1.68E-05	78752.32	84678.86	0	0	0	0	0.002894
7	4	1689.287	18.73803	5.693223	2.79E-05	82594.13	41733.47	46.16817	3442028	9472.416	1.40E-05	0.003534
7	8	1760.099	20.06037	7.175342	3.77E-05	81231.49	23790.86	69.399	69522774	6797.006	1.05E-06	0.004223
7	12	1720.762	17.90233	7.810266	4.99E-05	82846.13	11169.8	87.37633	96720915	6010.166	8.80E-07	0.004809
10	0	1632.999	21.48456	2.920004	7.70E-06	104707	107418.4	0	0	0	0	0.001798
10	4	1646.63	20.48472	4.151487	1.42E-05	107535.4	63726.62	41.70979	2838768	11272.23	1.54E-05	0.002491
10	8	1598.511	20.62844	5.873164	2.19E-05	117183.4	39012.37	61.7201	54712405	8794.14	1.21E-06	0.003357
10	12	1702.741	18.81778	6.802529	3.57E-05	109200.4	18530.36	84.85201	88630907	7529.664	9.66E-07	0.004056
0	0	1742.942	15.1283	12.0996	8.62E-05	0	0	0	0	0	0	0.006306
3	0	1727.302	17.11898	9.331498	4.75E-05	33316.13	33580.73	0	0	0	0	0.005578
3	4	1742.175	15.93544	9.854454	5.52E-05	31534.9	12172.82	62.39378	4472408	5392.635	1.41E-05	0.005536
3	8	1812.085	15.66329	11.21716	7.11E-05	32455.9	6113.93	82.46559	87797469	3257.532	9.39E-07	0.006216
3	12	1749.31	14.43355	10.99237	7.88E-05	33938.22	2512.101	94.3071	1.23E+08	2382.681	7.44E-07	0.006386
5	0	1731.78	19.14827	7.536999	2.87E-05	55509.72	53919.33	0	0	0	0	0.004176
5	4	1727.494	17.33492	8.359716	4.47E-05	57810	26841.59	52.33876	4240737	7400.392	1.31E-05	0.004562
5	8	1713.149	17.86954	9.263335	5.73E-05	57121.38	13540.06	69.82782	84690327	5233.329	9.25E-07	0.005214
5	12	1685.523	15.91529	9.474269	6.76E-05	54766.26	5175.718	90.49242	1.08E+08	4298.872	8.05E-07	0.005741
7	0	1652.021	19.0884	5.129956	1.83E-05	81187.99	84334.72	0	0	0	0	0.003102
7	4	1675.689	18.99008	6.069094	2.54E-05	86967.82	42548.12	43.49678	3242425	9341.431	1.45E-05	0.003557
7	8	1608.228	17.89354	7.128528	3.92E-05	78879.36	25261.83	70.07531	68775448	7300.547	1.05E-06	0.004354
7	12	1736.342	17.46528	8.493031	5.03E-05	81667.58	11035.38	87.94312	97448183	5875.642	8.81E-07	0.005102
10	0	1578.96	21.36526	2.957293	8.05E-06	115051.3	110807.2	0	0	0	0	0.001836
10	4	1556.116	20.75758	4.100845	1.30E-05	106932.2	62289.84	43.83581	2844930	11335.08	1.50E-05	0.00256
10	8	1682.786	20.39232	5.3683	2.30E-05	110550.8	38871.84	64.73433	56860876	8827.241	1.17E-06	0.003374
10	12	1640.858	19.89668	6.824725	3.50E-05	106509.9	18776.49	83.38034	90728077	7513.786	9.49E-07	0.00401
0	0	1781.502	13.83253	12.40613	8.66E-05	0	0	0	0	0	0	0.006417
3	0	1754.506	16.74253	9.645301	4.88E-05	32466.1	32641.26	0	0	0	0	0.00537
3	4	1789.092	15.99304	9.784236	5.66E-05	33078.25	12239.88	61.12057	4538632	5179.477	1.44E-05	0.005623
3	8	1786.792	15.0462	10.50438	7.09E-05	32540.98	5900.01	83.80919	83185915	3221.938	9.32E-07	0.006304
3	12	1844.656	15.53387	11.51653	7.93E-05	33182.81	2397.995	86.90896	1.22E+08	2500.066	7.72E-07	0.006554
5	0	1799.815	18.47374	7.424324	3.02E-05	57624.45	55528.78	0	0	0	0	0.004237
5	4	1806.201	17.62306	7.67892	4.33E-05	55163.85	25670.49	51.45901	3858963	7547.75	1.33E-05	0.004813
5	8	1698.751	17.09368	8.891284	5.68E-05	58990.78	14369.28	80.04801	79178076	5202.688	9.39E-07	0.00543
5	12	1884.972	17.15173	9.808331	6.52E-05	54539.67	5195.39	86.61	1.16E+08	4301.543	8.39E-07	0.005473
7	0	1688.313	18.9822	4.875134	1.65E-05	83504.55	85067.07	0	0	0	0	0.00296
7	4	1672.678	18.85392	5.969637	2.52E-05	82695.76	43469.23	45.37199	3272589	9885.422	1.40E-05	0.003557
7	8	1674.272	19.00003	7.616104	3.66E-05	84322.72	22929.15	71.95116	64879838	7381.626	1.01E-06	0.004308
7	12	1767.075	17.44953	7.942999	5.23E-05	84492.25	11488.64	84.36667	96717510	5903.506	8.44E-07	0.004955
10	0	1497.386	21.23196	2.957159	7.81E-06	110730	110899.9	0	0	0	0	0.001792
10	4	1625.676	21.96304	4.226937	1.39E-05	115589.1	61315.88	42.93295	2758366	11136.27	1.49E-05	0.002454
10	8	1618.565	19.88701	5.761793	2.32E-05	107959.5	41433.6	62.83944	55319613	9111.459	1.20E-06	0.003418

10	12	1691.304	19.47161	6.764581	3.55E-05	107528.8	18372.55	89.03517	88913471	7318.847	9.10E-07	0.004334
0	0	1772.319	14.2366	12.58029	8.26E-05	0	0	0	0	0	0	0.006484
3	0	1727.092	16.88212	9.60318	4.73E-05	32392.63	33625.61	0	0	0	0	0.005466
3	4	1828.011	16.95476	10.07287	5.98E-05	34389.53	11855.71	61.1231	4478160	5245.665	1.37E-05	0.00594
3	8	1792.545	15.68776	11.5068	7.09E-05	33504.67	5787.63	79.67157	88589739	3255.445	9.37E-07	0.005829
3	12	1724.921	14.52291	11.88508	7.93E-05	32253.13	2457.056	96.44753	1.13E+08	2609.423	7.67E-07	0.006371
5	0	1756.52	17.32578	7.303489	3.04E-05	57239.67	59539.36	0	0	0	0	0.004162
5	4	1735.314	17.34895	8.583951	4.19E-05	56162.59	25171.02	51.94589	4026224	7252.462	1.32E-05	0.004709
5	8	1744.371	16.01312	9.299839	5.69E-05	53464.85	13601.8	77.35062	86019593	5335.389	9.47E-07	0.005142
5	12	1796.068	16.52731	9.822448	6.83E-05	56891.1	5037.946	95.52178	1.11E+08	4068.149	8.73E-07	0.005607
7	0	1555.39	19.22304	5.158046	1.64E-05	83973.71	88817.94	0	0	0	0	0.003085
7	4	1663.768	19.28968	5.567124	2.65E-05	86708.96	42866.96	47.16617	3166123	9721.961	1.42E-05	0.003475
7	8	1743.265	18.39548	7.390472	3.94E-05	80720.24	24730.8	69.17232	65648684	7020.593	1.01E-06	0.00409
7	12	1747.123	17.66109	8.842776	5.31E-05	80771.05	11482.73	86.40607	96001419	5620.913	8.65E-07	0.005041
10	0	1543.608	22.20466	2.733124	8.43E-06	102717.7	103365.8	0	0	0	0	0.001849
10	4	1521.267	20.13769	4.197655	1.41E-05	113852.5	63247.44	43.85377	2807646	11873.97	1.55E-05	0.002579
10	8	1580.232	19.5121	5.45608	2.42E-05	112276.9	39237.85	64.94704	57278391	9107.603	1.18E-06	0.003503
10	12	1650.563	19.51882	6.975708	3.65E-05	105646.9	18773.1	78.93253	93507227	7996.854	9.84E-07	0.004282

3. Non augmented data modelling Matlab code

```
%% PHYSICALLY ROBUST TPH MODELING WITH SENSITIVITY ANALYSIS
% Comparison of Response Surface Methodology vs. Artificial Neural Networks
clear; clc; close all;

%% 1. Data Initialization
data = [ 0, 0, 1850, 14.5, 12.5, 8.50E-05, 0, 0, 0, 0, 0, 0, 0.006757;
        3, 0, 1785, 16.5, 9.7, 4.90E-05, 32540, 32540, 0, 0, 0, 0, 0.005434;
        3, 4, 1802, 16.1, 10.2, 5.70E-05, 32540, 11850, 63.6, 4500000, 5172.499, 1.41E-05, 0.00566;
        3, 8, 1818, 15.6, 10.9, 7.00E-05, 32540, 6120, 81.2, 88000000, 3302.5, 9.23E-07, 0.005996;
        3, 12,1835, 15.2, 11.7, 7.80E-05, 32540, 2410, 92.6, 1.2E+08, 2510.833, 7.72E-07, 0.006376;
        5, 0, 1715, 18.2, 7.3, 3.00E-05, 56120, 56120, 0, 0, 0, 0, 0.004257;
        5, 4, 1735, 17.6, 8.1, 4.30E-05, 56120, 26590, 52.6, 4000000, 7382.498, 1.31E-05, 0.004669;
        5, 8, 1758, 17, 9, 5.60E-05, 56120, 13540, 75.9, 81000000, 5322.499, 9.37E-07, 0.005119;
        5, 12,1785, 16.3, 9.9, 6.70E-05, 56120, 5180, 90.8, 1.1E+08, 4245, 8.25E-07, 0.005546;
        7, 0, 1640, 19.8, 5, 1.70E-05, 81870, 81870, 0, 0, 0, 0, 0.003049;
        7, 4, 1665, 19.1, 5.9, 2.60E-05, 81870, 44250, 45.9, 3300000, 9404.998, 1.39E-05, 0.003544;
        7, 8, 1690, 18.5, 7.2, 3.80E-05, 81870, 23810, 70.9, 68000000, 7257.499, 1.04E-06, 0.00426;
        7, 12,1720, 17.7, 8.6, 5.30E-05, 81870, 11540, 85.9, 98000000, 5860.833, 8.77E-07, 0.005;
        10, 0, 1575, 21.3, 2.9, 8.00E-06, 110230, 110230, 0, 0, 0, 0, 0.001841;
        10, 4, 1595, 20.8, 4.1, 1.40E-05, 110230, 63550, 42.4, 2800000, 11670, 1.51E-05, 0.002571;
        10, 8, 1620, 20.1, 5.6, 2.30E-05, 110230, 38720, 64.9, 55000000, 8938.749, 1.18E-06, 0.003457;
        10, 12,1655, 19.3, 6.9, 3.60E-05, 110230, 18550, 83.2, 91000000, 7639.999, 9.14E-07, 0.004169 ];
X = data(:, 1:2); Y = data(:, 9); n = length(Y);

%% 2. Simplified ANN (3 Neurons) with LOOCV
yPredANN_CV = zeros(n,1);
best_net = []; % To save a representative net for sensitivity

for i = 1:n
```

```
    % Ensure reproducibility by resetting the random seed for each fold
    rng(i);

    X_tr = X; X_tr(i,:) = []; Y_tr = Y; Y_tr(i) = [];
    mX = mean(X_tr); sX = std(X_tr); mY = mean(Y_tr); sY = std(Y_tr);

    net = fitnet(3, 'trainbr');
    net.trainParam.showWindow = false;

    net = train(net, ((X_tr-mX)./sX)', ((Y_tr-mY)./sY)');
    yPredANN_CV(i) = (net(((X(i,:)-mX)./sX)') * sY) + mY;

    if i == 1; best_net = net; end % Capture weights from one fold
end

%% 3. Garson's Algorithm (Sensitivity Analysis)
w_input_hidden = best_net.IW{1};   % Weights from Input to Hidden
w_hidden_output = best_net.LW{2,1}; % Weights from Hidden to Output
Q = abs(w_input_hidden) .* abs(w_hidden_output');
S_sum = sum(Q, 1);
Relative_Importance = (S_sum / sum(S_sum)) * 100;

%% 4. Simplified RSM (Parsimony)
tbl = table(X(:,1), X(:,2), Y, 'VariableNames', {'Oil','Time','TPH'});
mdlRSM = fitlm(tbl, 'TPH ~ Oil*Time + Oil^2 + Time^2');
mdlRSM = step(mdlRSM); % Removes insignificant terms
yPredRSM_CV = zeros(n,1);
for i = 1:n
    tbl_cv = tbl; tbl_cv(i,:) = [];
    mdl_temp = fitlm(tbl_cv, mdlRSM.Formula);
    yPredRSM_CV(i) = predict(mdl_temp, tbl(i,1:2));
end

%% 5. Results & Metrics
R2_RSM = corr(Y, yPredRSM_CV)^2;
R2_ANN = corr(Y, yPredANN_CV)^2;
rssRSM = sum((Y - yPredRSM_CV).^2);
rssANN = sum((Y - yPredANN_CV).^2);
aicRSM = n*log(rssRSM/n) + 2*mdlRSM.NumCoefficients;
aicANN = n*log(rssANN/n) + 2*13;

fprintf('\n--- VALIDATED PERFORMANCE ---\n');
fprintf('RSM R2: %.4f | AIC: %.2f\n', R2_RSM, aicRSM);
fprintf('ANN R2: %.4f | AIC: %.2f\n', R2_ANN, aicANN);

%% 6. Visualizations
figure('Color', 'w', 'Position', [100, 100, 950, 450]);

% Subplot 1: Sensitivity
subplot(1,2,1);
hBar = bar(Relative_Importance, 'FaceColor', [0.2 0.6 0.8]);
set(gca, 'XTickLabel', {'Oil Content', 'Time'});
ylabel('Relative Importance (%)');
title('ANN Sensitivity Analysis (Garson)');
grid on;
% Label percentages on bars
```

```
xtips = 1:numel(Relative_Importance);
ytips = Relative_Importance;
labels = string(round(Relative_Importance,1)) + '%';
text(xtips, ytips, labels, 'HorizontalAlignment', 'center', 'VerticalAlignment', 'bottom');

% Subplot 2: Accuracy
subplot(1,2,2);
scatter(Y, yPredRSM_CV, 70, 'r', 'filled', 'MarkerFaceAlpha', 0.6); hold on;
scatter(Y, yPredANN_CV, 70, 'b', 's', 'filled', 'MarkerFaceAlpha', 0.6);
plot([0 100], [0 100], 'k--', 'LineWidth', 1.5);
xlabel('Experimental TPH Removal (%)'); ylabel('Predicted TPH Removal (%)');
legend(sprintf('RSM (R^2 = %.4f)', R2_RSM), ...
      sprintf('ANN (R^2 = %.4f)', R2_ANN), ...
      'Perfect Fit', 'Location', 'northwest');
title('Cross-Validated Model Accuracy');
grid on;
```

4. TAYLOR DIAGRAM COMPARATIVE CODE

```
%% Taylor Diagram: Summarized Model Comparison
% Optimized for MATLAB 2024
clear; clc; close all;

%% 1. Data Preparation
% Reference Standard Deviation (Adjust this to your actual data's Std Dev)
std_ref = 10;

% Augmented Data (Averaged R2 from Training, Validation, and Testing)
ann_aug_r2 = mean([0.9935, 0.9940, 0.9856]); % ~0.9910
rsm_aug_r2 = mean([0.9960, 0.9932, 0.9857]); % ~0.9916

% Non-Augmented Data (Validated Performance)
ann_non_r2 = 0.8727;
rsm_non_r2 = 0.9707;

% Combine for plotting
models = {'ANN Augmented', 'RSM Augmented', 'ANN Non-Augmented', 'RSM Non-Augmented'};
r2_vals = [ann_aug_r2, rsm_aug_r2, ann_non_r2, rsm_non_r2];
r_vals  = sqrt(r2_vals);

% Calculated/Estimated Standard Deviations for the plot
% Note: In a true Taylor diagram, these are derived from your model outputs.
std_devs = [9.9, 10.1, 8.0, 9.2];

%% 2. Figure Initialization
figure('Color', 'w', 'Position', [100, 100, 900, 700]);
hold on;
axis equal;

% Define plot limits
max_rho = 12;
theta = linspace(0, pi/2, 100);

%% 3. Background Grid (Handling Visibility for Legend)
```

```
% Radial lines for Correlation (R)
corrs = [0.2 0.4 0.6 0.7 0.8 0.9 0.95 0.99];
for c = corrs
    % Setting 'HandleVisibility' to 'off' removes "data1, data2..." from legend
    line([0 max_rho*c], [0 max_rho*sin(acos(c))], ...
        'Color', [0.8 0.8 0.8], 'LineStyle', ':', 'HandleVisibility', 'off');
    text(max_rho*c*1.02, max_rho*sin(acos(c))*1.02, num2str(c), 'FontSize', 9);
end

% Reference Arc (Dashed Black Line)
plot(std_ref*cos(theta), std_ref*sin(theta), 'k--', ...
    'LineWidth', 1.5, 'HandleVisibility', 'off');
text(std_ref, -0.5, 'Reference', 'FontSize', 10, 'FontWeight', 'bold');

%% 4. Plotting the Models
% Define colors and markers
colors = [
    0.85, 0.33, 0.10; % Red/Orange (ANN Aug)
    0.93, 0.69, 0.13; % Gold/Yellow (RSM Aug)
    0.00, 0.45, 0.74; % Blue (ANN Non-Aug)
    0.47, 0.67, 0.19  % Green (RSM Non-Aug)
];
markers = {'o', 's', 'o', 's'};

for i = 1:length(models)
    % Convert Polar (StdDev, R) to Cartesian (x, y)
    x = std_devs(i) * r_vals(i);
    y = std_devs(i) * sin(acos(r_vals(i)));

    plot(x, y, markers{i}, 'MarkerSize', 12, ...
        'MarkerFaceColor', colors(i,:), 'MarkerEdgeColor', 'k', ...
        'DisplayName', sprintf('%s (R^2=%.4f)', models{i}, r2_vals(i)));
end

%% 5. Formatting & Aesthetics
grid on;
set(gca, 'TickDir', 'out', 'Box', 'off');
xlabel('Standard Deviation', 'FontSize', 12);
ylabel('Standard Deviation', 'FontSize', 12);
title('Comparison of Summarized Model Performance', 'FontSize', 14);

% Final Legend (Only shows the 4 models)
legend('Location', 'northeastoutside', 'FontSize', 10);

% Annotate Correlation label
text(max_rho*0.7, max_rho*0.8, 'Correlation Coefficient (R)', ...
    'Rotation', -45, 'FontWeight', 'bold', 'Color', [0.4 0.4 0.4]);

hold off;
```