# Statistical Analysis Report: Tractor Fuel Consumption (Synthetic Data)

This report contains a synthetic dataset generated to study the impact of tillage depth, implement width, soil moisture, and speed on tractor fuel consumption (L/h) for a spring tine cultivator. The dataset contains 300 simulated observations.

## 1. Data preview

First 8 rows of the dataset:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| tillage\_depth\_cm | implement\_width\_m | soil\_moisture\_pct | speed\_kmh | fuel\_L\_per\_hr |
| 12.49 | 1.63 | 14.22 | 3.25 | 15.893 |
| 24.01 | 2.83 | 16.96 | 2.16 | 20.548 |
| 19.64 | 2.85 | 14.43 | 3.09 | 16.535 |
| 16.97 | 3.09 | 12.22 | 5.5 | 16.851 |
| 8.12 | 3.32 | 13.02 | 4.53 | 16.644 |
| 8.12 | 3.94 | 21.52 | 7.36 | 24.455 |
| 6.16 | 2.79 | 15.16 | 6.9 | 17.046 |
| 22.32 | 2.31 | 19.11 | 4.05 | 20.708 |

## 2. Descriptive statistics

tillage\_depth\_cm: mean=14.904, std=5.887, min=5.100, 25%=9.780, 50%=15.235, 75%=20.137, max=24.800

implement\_width\_m: mean=2.777, std=0.757, min=1.530, 25%=2.127, 50%=2.805, 75%=3.433, max=4.000

soil\_moisture\_pct: mean=21.777, std=7.086, min=10.120, 25%=15.245, 50%=21.320, 75%=27.925, max=34.920

speed\_kmh: mean=5.117, std=1.781, min=2.040, 25%=3.558, 50%=5.280, 75%=6.725, max=7.990

fuel\_L\_per\_hr: mean=19.099, std=2.732, min=12.425, 25%=17.185, 50%=19.189, 75%=20.872, max=27.452

## 3. Correlation matrix

Pearson correlation between variables:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | tillage\_depth\_cm | implement\_width\_m | soil\_moisture\_pct | speed\_kmh | fuel\_L\_per\_hr |
| tillage\_depth\_cm | 1.000 | -0.101 | -0.047 | 0.006 | 0.386 |
| implement\_width\_m | -0.101 | 1.000 | 0.063 | 0.027 | 0.291 |
| soil\_moisture\_pct | -0.047 | 0.063 | 1.000 | -0.047 | 0.306 |
| speed\_kmh | 0.006 | 0.027 | -0.047 | 1.000 | 0.624 |
| fuel\_L\_per\_hr | 0.386 | 0.291 | 0.306 | 0.624 | 1.000 |

## 4. Regression model

Model formula: fuel\_L\_per\_hr ~ tillage\_depth\_cm + implement\_width\_m + soil\_moisture\_pct + speed\_kmh + tillage\_depth\_cm:implement\_width\_m + soil\_moisture\_pct:speed\_kmh

Model summary (coefficients):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| term | coef | std err | t | P>|t| | [0.025 0.975] |
| Intercept | 7.3936 | 1.1403 | 6.484 | 3.783e-10 | [5.149, 9.638] |
| tillage\_depth\_cm | 0.2278 | 0.0495 | 4.606 | 6.126e-06 | [0.130, 0.325] |
| implement\_width\_m | 1.2484 | 0.2779 | 4.492 | 1.015e-05 | [0.701, 1.795] |
| soil\_moisture\_pct | 0.0140 | 0.0372 | 0.378 | 0.7057 | [-0.059, 0.087] |
| speed\_kmh | 0.4910 | 0.1527 | 3.216 | 0.001445 | [0.191, 0.791] |
| tillage\_depth\_cm:implement\_width\_m | -0.0106 | 0.0176 | -0.605 | 0.5454 | [-0.045, 0.024] |
| soil\_moisture\_pct:speed\_kmh | 0.0222 | 0.0068 | 3.262 | 0.001238 | [0.009, 0.036] |

R-squared: 0.756; Adj. R-squared: 0.751; F-statistic: 151.52 (p=9.61e-87)

## 5. ANOVA (Type II)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | sum\_sq | df | F | PR(>F) |
| tillage\_depth\_cm | 405.2492 | 1.0000 | 218.3115 | 0.0000 |
| implement\_width\_m | 200.3968 | 1.0000 | 107.9556 | 0.0000 |
| soil\_moisture\_pct | 250.0908 | 1.0000 | 134.7262 | 0.0000 |
| speed\_kmh | 881.6996 | 1.0000 | 474.9799 | 0.0000 |
| tillage\_depth\_cm:implement\_width\_m | 0.6802 | 1.0000 | 0.3664 | 0.5454 |
| soil\_moisture\_pct:speed\_kmh | 19.7507 | 1.0000 | 10.6399 | 0.0012 |
| Residual | 543.8925 | 293.0000 | nan | nan |

## 6. Multicollinearity (VIF)

const: VIF = 40.326

tillage\_depth\_cm: VIF = 1.012

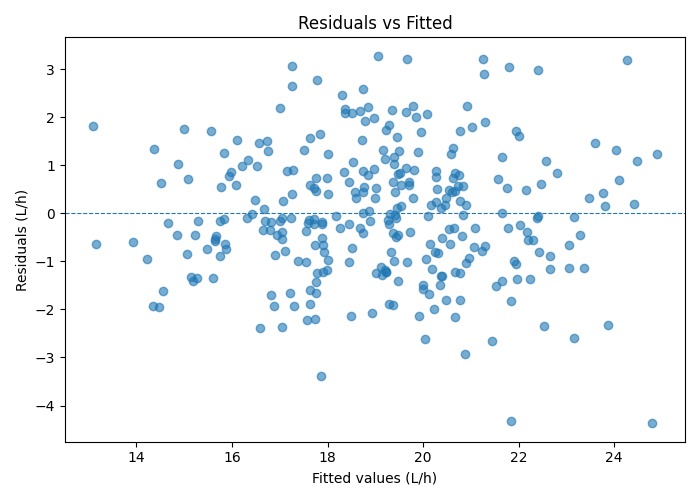
implement\_width\_m: VIF = 1.015

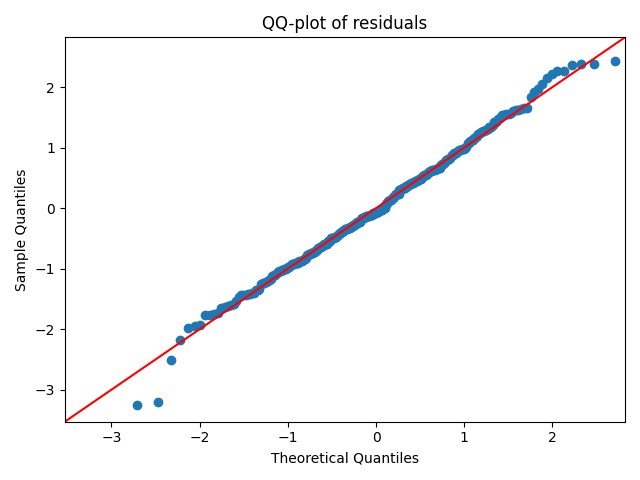
soil\_moisture\_pct: VIF = 1.008

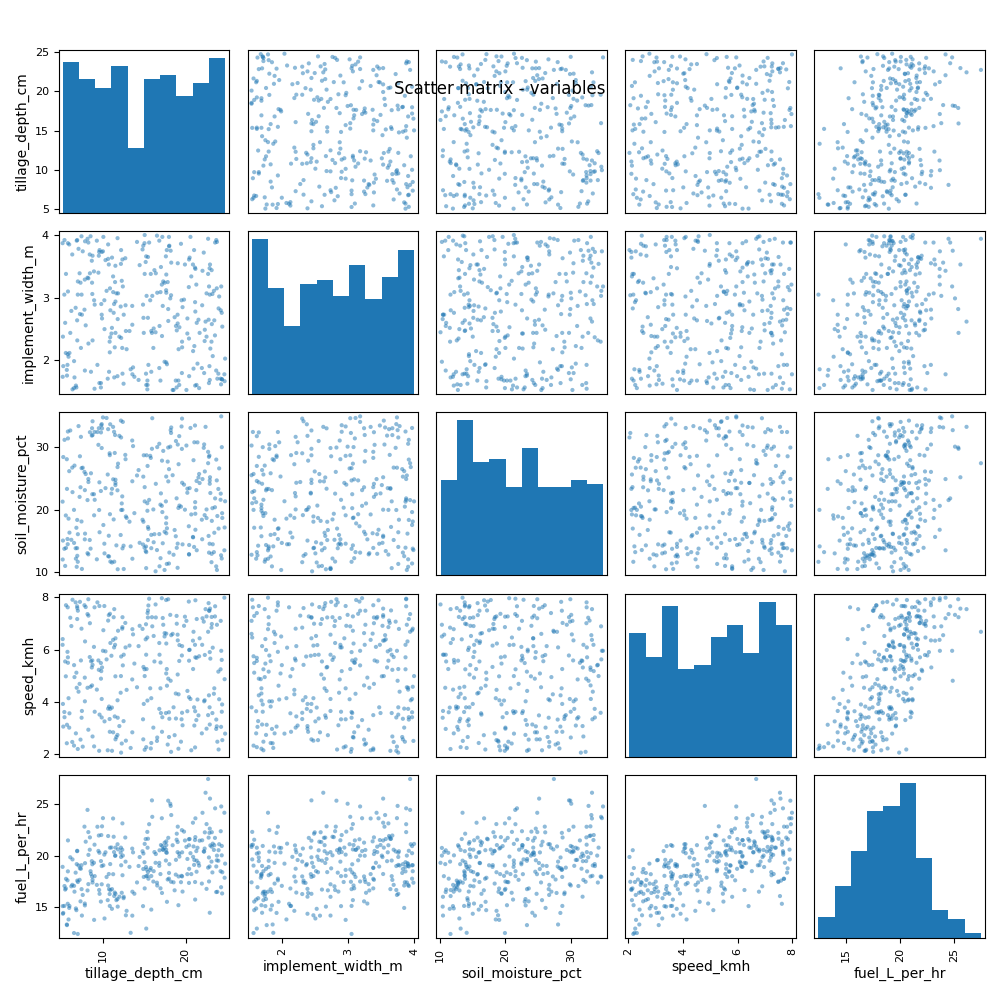
speed\_kmh: VIF = 1.003

## 7. Diagnostic plots

Included: residuals vs fitted, QQ-plot of residuals, scatter matrix.







## 8. Key findings and recommendations

• In this synthetic experiment, tillage depth, implement width, soil moisture, and speed show positive associations with fuel consumption (L/h).

• Interactions between depth & width and between soil moisture & speed capture combined effects and were included in the model.

• Diagnostics (QQ plot, residuals) should be checked in real data; heteroscedasticity can be present. Consider robust SEs if present in field data.

• For practical field reporting, consider fuel per area (L/ha), include tractor engine rpm, gear, soil texture, and implement setting in future experiments.