

The application of robust regression to a production function comparison – the example of Swiss corn*

Robert Finger

Werner Hediger

* IED Working Paper No.2: www.ied.ethz.ch



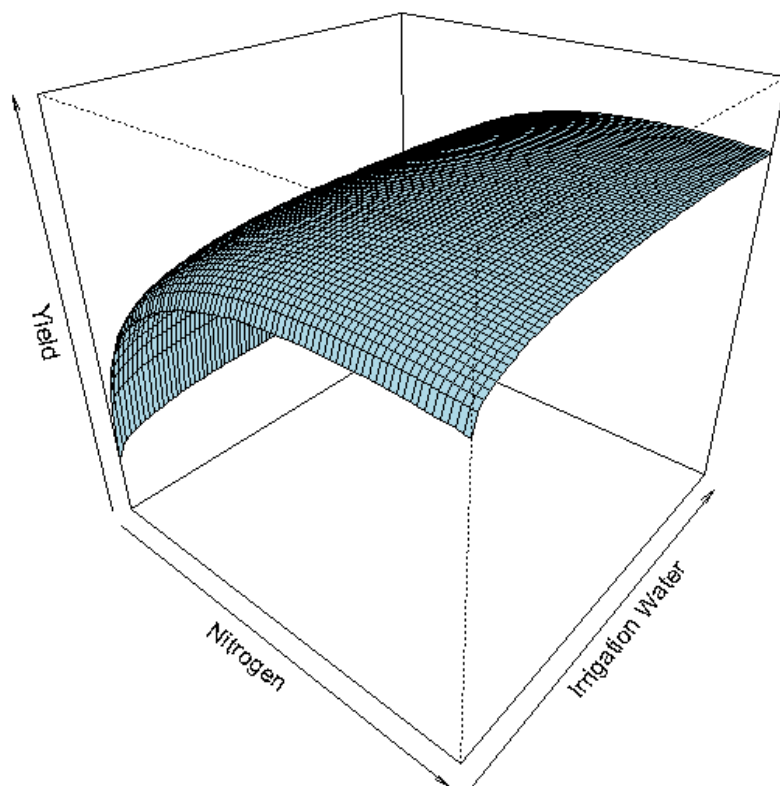


The Problem

- Goal of (Crop) Production Functions: adequate representation of the input-output relationship
- So far, crop production function evaluation was based on functional forms – not on estimation techniques
- Goal of the paper: (efficient) estimation and evaluation of crop production functions in the presence of extreme weather events

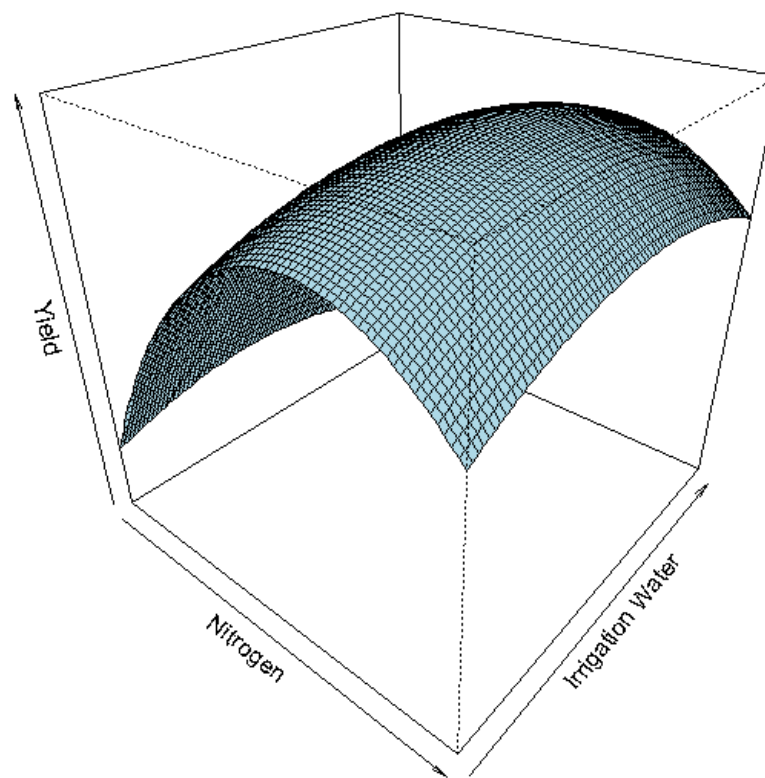
The Functional Forms: Yield ~ Nitrogen Fertilizer + Irrigation Water

Square Root Function



$$Y = \beta_0 + \beta_1 \sqrt{N} + \beta_2 \sqrt{W} + \beta_3 N + \beta_4 W + \beta_5 \sqrt{NW}$$

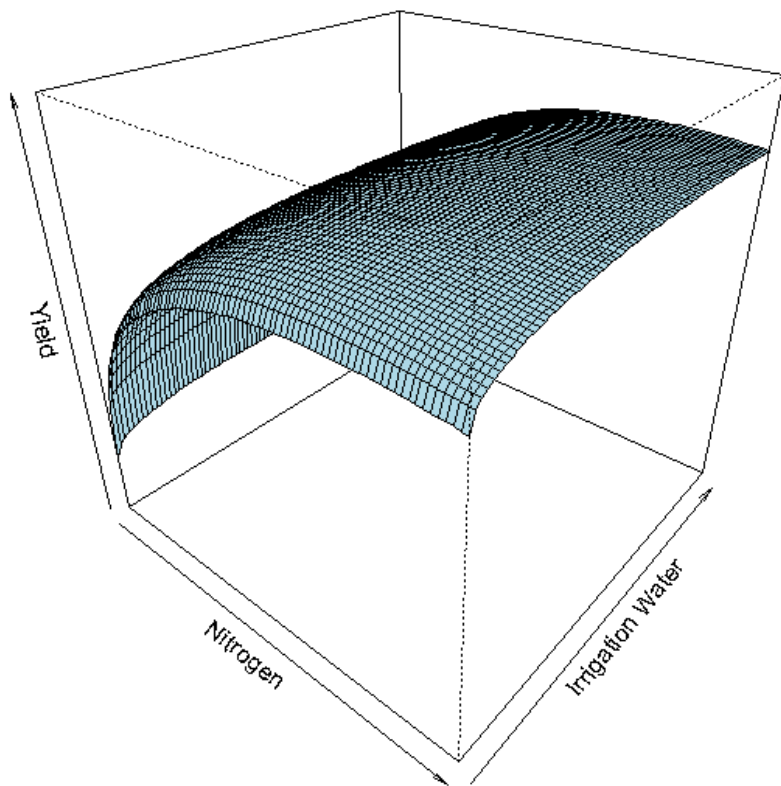
Quadratic Function



$$Y = \beta_0 + \beta_1 N + \beta_2 W + \beta_3 N^2 + \beta_4 W^2 + \beta_5 NW$$

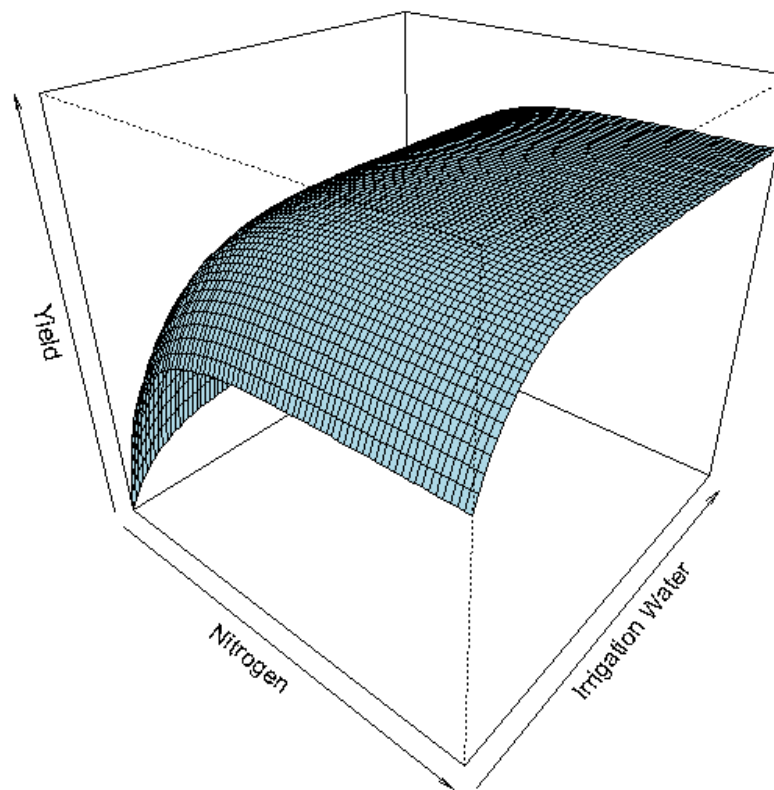
The Functional Forms: Yield ~ Nitrogen Fertilizer + Irrigation Water

Square Root Function



$$Y = \beta_0 + \beta_1 \sqrt{N} + \beta_2 \sqrt{W} + \beta_3 N + \beta_4 W + \beta_5 \sqrt{NW}$$

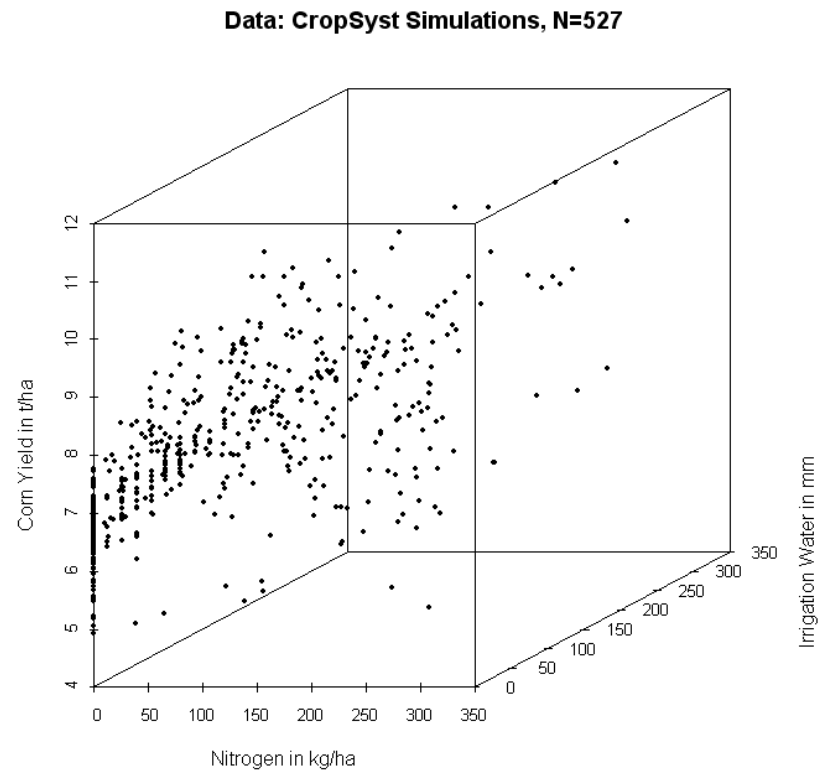
Mitscherlich-Baule Function



$$Y = \beta_0 \times (1 - \exp(-\beta_2 \times (\beta_3 + N))) \times (1 - \exp(-\beta_4 \times (\beta_5 + W)))$$

The Data

- Simulated Field Experiment: *Corn Yields~Nitrogen+Irrigation*, 1983-2003
weather data on the Swiss Plateau

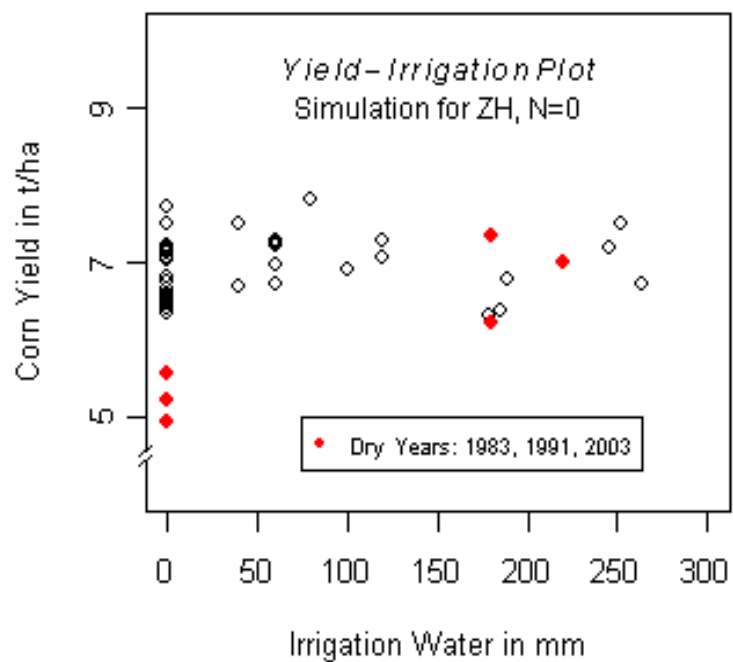


Extreme weather Events

- Weather affects crop yields! & OLS estimation (and further inference) is vulnerable to exceptional observations (i.e. outliers)
 - Identification, analysis and ‘treatment’ of outliers necessary
- Different observation years are currently ‘treated’ using dummy variables
- (However,...) Weather conditions also affect the input-output relationships

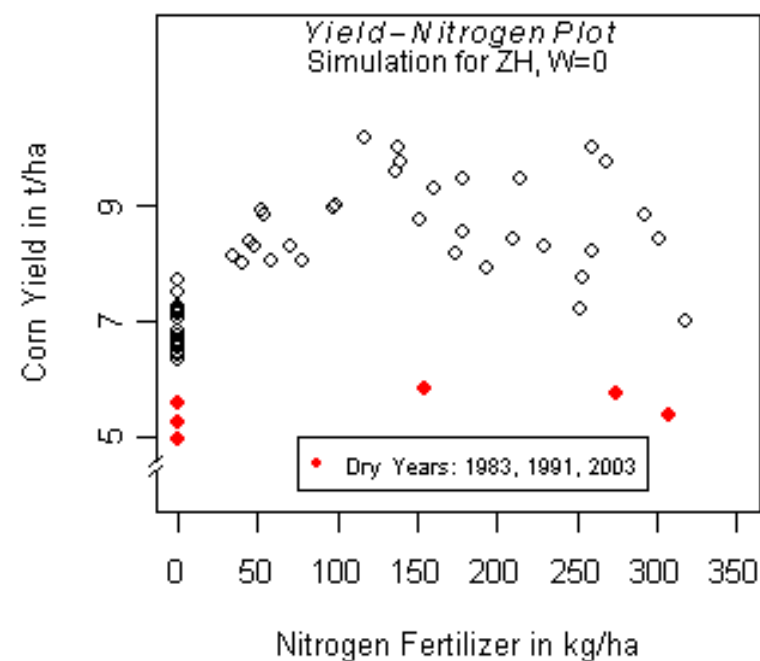
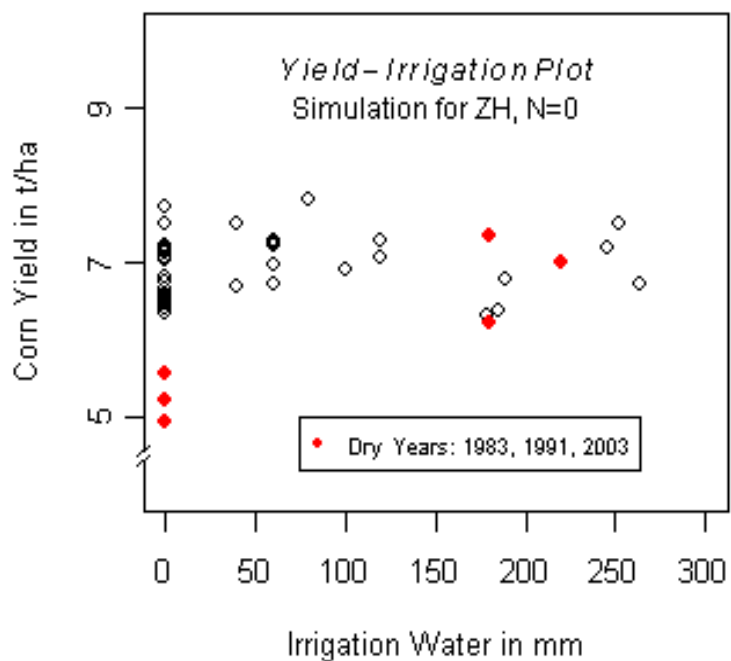
In exceptionally hot and dry years:

- Higher (yield) response to irrigation



In exceptionally hot and dry years:

- Higher (yield) response to irrigation
- Smaller (yield) response to nitrogen



Extreme weather Events

- Weather affects crop yields! & OLS estimation (and further inference) is vulnerable to exceptional observations (i.e. outliers)
 - Identification, analysis and ‘treatment’ of outliers necessary
- Different observation years are currently ‘treated’ using dummy variables
- Weather conditions also affect the input-output relationships
 - single ‘year’ dummy is not sufficient
- Dummies for weather extremes? – (multivariate) distribution of variables matters

Robust Regression

- Goals:
 - Efficient estimation in the presence of outliers
 - Identification of the 'true' underlying relationship
 - Identification & Interpretation of outliers

Robust Regression

- Estimation of quasi-linear functions: **Reweighted Least Squares (RLS)**
 - Least Trimmed Squares Regression (Rousseeuw & Leroy, 1987)
 - Outlier identification: standardized robust residual $> 2.5 \rightarrow$ outlier
 - LS estimation without outliers
- Non-linear function: **Iteratively Reweighted Least Squares (IRLS)**
 - Iteration until convergence
 - robust (down-)weighting of outlying observations at each iteration step (M-estimation, Tukey's biweight)
- Estimation: SAS (ROBUSTREG, NLIN)

Results

- Robust regression leads, generally, to higher goodness of fit than OLS
- 7-8% of the observations are identified as outliers
- most outliers are identified in the years 1983, 1991 and 2003
 - hot and dry summers (vegetation periods)

→ An example of coefficient estimates...

Estimation Results for Corn – Square Root Function

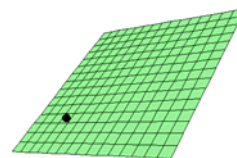
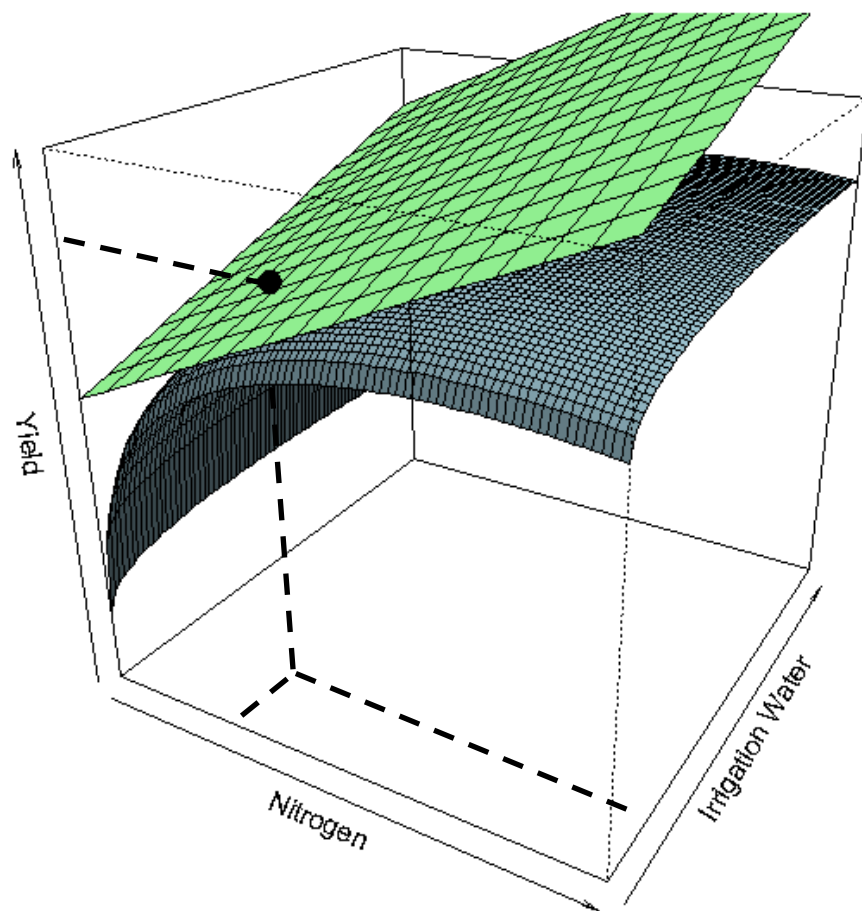
Variable	Estimation Method	
	OLS	RLS
Intercept	6589.997 (155.02)**	6601.924 (162.13)**
$N^{1/2}$	297.1821 (12.42)**	313.0936 (16.34)**
$W^{1/2}$	75.09137 (4.26)**	67.1385 (4.17)**
N	-11.2156 (6.88)**	-10.544 (8.15)**
W	-3.03419 (2.40)*	-2.49922 (2.17)*
$(NW)^{1/2}$	1.46442 (1.43)	0.364377 (0.45)
adj. R^2	0.5834	0.7330

Note: Statistics in parentheses are t statistics
 (**) – indicates significance at the 1% level
 (*) – indicates significance at the 5% level

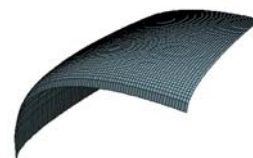
$$Y = \alpha_0 + \alpha_1 \cdot N^{1/2} + \alpha_2 \cdot W^{1/2} + \alpha_3 \cdot N + \alpha_4 \cdot W + \alpha_5 \cdot (N \cdot W)^{1/2}$$

→ Coefficient estimates are used for profit maximization...

Profit Maximization



*Input-Output Price Plane
with tangent point*



Production Function

$$\frac{\partial f(W^*, N)}{\partial W} = P_{\text{Water}} / P_{\text{Corn}} \quad \& \quad \frac{\partial f(W, N^*)}{\partial N} = P_{\text{Nitrogen}} / P_{\text{Corn}}$$

Results

- Profit maximizing input and output levels, quasi-rents:

Functional Form-Estimation Method	Optimal amount of Nitrogen applied (kg/ha)	Optimal amount of irrigation Water applied (mm)	Optimal yield (kg/ha)	Maximum net return (CHF/ha)
Quadratic-OLS	172.8	179.6	9695	5840.32
Square Root-OLS	131.3	133.9	9180	5602.82
Mitscherlich-Baule-OLS	111.2	61.3	9078	5613.55
Quadratic-RLS	177.4	163.8	9859	5947.68
Square Root-RLS	147.7	108.6	9324	5684.56
Mitscherlich-Baule-IRLS	124.9	116.7	9286	5691.51

Results

- Robust regression leads to smaller ‘differences’ between the functional forms
→ approaches real underlying relationship
- Economic analysis: relative costs of misspecification – goal: function/estimation technique that minimizes maximum losses if incorrectly applied
→ Robust estimations lead to smaller costs of misspecification – even if ‘true’ function is based on OLS regression
- Square root function estimated with RLS is the best specification for corn yields at the Swiss Plateau

Conclusions

- Econometric AND economic point of view:
 - recommendation of application of robust regression to production function estimation (besides OLS estimation)
 - valuable tool for agricultural (economic) modeling
- Extreme weather events can lead to inefficient OLS estimation
- Square root function estimated with RLS is the best specification for corn yields on the Swiss Plateau

Outlook

- Yield variability estimation is based on regression residuals of production function estimation
 - improved by robust regression (no masking of exceptional observations)
- Integration of biophysical and economic modeling to assess impacts of climate change on managed agricultural systems
 - More extreme events in future: need for robust analysis methods!



Thank you very much for your attention

Cost of Misspecification

- Evaluation of functional forms - concept of relative costs of misspecification:
- The decrease in quasi-rents if optimal input levels of an incorrect function are used instead of the real underlying production function
- Potential loss due to a wrong functional form – and – due to an improper estimation technique

Relative Costs of Misspecification

Cost of using optimal input levels based on:						
<u>When the true function is:</u>	Quadratic-OLS	Square Root-OLS	Mitscherlich-Baule-OLS	Quadratic-RLS	Square Root-RLS	Mitscherlich-Baule-IRLS
Quadratic-OLS	0	93.01	297.88	4.23	77.85	135.18
Square Root-OLS	30.61	0	39.83	32.13	8.41	2.01
Mitscherlich-Baule-OLS	113.22	41.38	0	109.97	41.86	27.34
Quadratic-RLS	3.77	104.65	296.39	0	68.59	145.23
Square Root-RLS	7.18	27.08	35.49	8.45	0	23.14
Mitscherlich – Baule-IRLS	57.52	54.08	3.11	51.85	9.86	0