Appendix 5: Regression analysis

Regression analysis

To illustrate the influence of rainfall on harvests, we regressed[1] total harvest per household (kg, year 2014) on age and years of schooling of the household head, planted area (ha), rainfall (mm, year 2014) and the number of dry spells (>7 days without rainfall in the rainy season January-May 2014[2] (Hanisch 2015).

Results

The overall model was most highly significant (P_{ANOVA} <0.001) and explained 65.4 % of the variance in harvest data. Toal planted area (P=<0.001), annual rainfall (P=0.005) and numbers of dry spells (P=0.002) were the only significant predictors of total harvest/yr/ha. Stepwise eliminating non-significant predictors yielded the same model. The standardised beta values of the coefficients confirmed that total planted area (0.717) had the strongest influence on harvests, followed by total rainfall in 2014 (0.166). The occurrence of dry spells had a significant negative effect on harvests (-0.155).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,808[a]	,654	,641	811,896055900230700

a. Predictors: (Constant), rainfall_mm, years of schooling, total planted area (ha), Age (HH head), dry

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	172809918,809	5	34561983,762	52,432	,000[b]
	Residual	91625353,577	139	659175,206		
	Total	264435272,386	144			

a. Dependent Variable: harvest (kg)

b. Predictors: (Constant), rainfall_mm, years of schooling, total planted area (ha), Age (HH head), dry spells (>7days)

[1] In multiple linear regression, the single predictors contribute in an additive fashion to predict the independent variable. Fundamental agricultural production theory (law of the minimum) requires a non-linear, mulitplicative model. In standard Cobb-Douglas production function analysis (Cobb and Douglas 1928), this is achieved by regressing the log of a harvest variable on the logs of socio-demographic, input and environmental variables. To illustrate the influence of rainfall on production in the most simple way, we opted for a simple regression analysis here following a reviewer comment.

[2] Concerning the 2013/2014 cropping season, the rainy season started in December 2013 (Hanisch 2015). Unfortunately we don´t have rainfall data for that month.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	309,694	449,922		,688	,492
	Age (HH head)	-3,125	5,041	-,033	-,620	,536
	Years of schooling	32,061	22,435	,074	1,429	,155
	Total planted area (ha)	357,132	26,511	,717	13,471	**,000**
	Dry spells (>7 days)	-171,977	60,208	-,155	-2,856	**,005**
	Rainfall (mm)	1,107	,359	,166	3,081	**,002**

a. Dependent Variable: harvest kg

LITERATURE CITED

Cobb, C. W., and P. H. Douglas. 1928. A Theory of Production. *American Economic Review* 18(1):139–165.

Hanisch, S. 2015. Improving cropping systems of semi-arid south-western Madagascar under multiple ecological and socio-economic constraints. Doctoral Dissertation. University of Witzenhausen.