Agriculture, Ecosystems and Environment 214 (2015) 1–9

Contents lists available at ScienceDirect

Agriculture, Ecosystems and Environment

journal homepage: www.elsevier.com/locate/agee

The importance of local factors and management in determining wheat yield variability in on-farm experimentation in Tigray, northern Ethiopia

R. Kraaijvanger[a,b,*], A. Veldkamp[c]

[a] Soil Geography and Landscape group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands
[b] College of Dryland Agriculture, Mekelle University, P.O. Box 231, Mekelle, Ethiopia
[c] Faculty of Geo-Information Science and Earth Observation (ITC), Twente University, P.O. Box 217, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 1 August 2015
Accepted 3 August 2015
Available online 27 August 2015

Keywords:
Soil properties
Multiple linear regression
Farm management
Fertilizer application
Best practices

ABSTRACT

Low crop yield in Tigray is one of the causes of food insecurity. Intervention work to increase yields, however, had only limited success and farmers often hesitated to adopt recommended practices. Considering this, we used participatory on-farm experimentation to arrive at best practices matching local preferences, complexity and context. Outcomes were evaluated at meta level and at site level, respectively to identify major sources of yield variability and direct relationships between yield and treatment, location and soil properties. About 56% of the total grain yield variability in our experiments was explained by a linear regression model with management, altitude and N-fertilizer input. When management was excluded, still 49% of the grain yield variability was explained by altitude, N-fertilizer input, N-total, organic-C, rainfall and K-exchangeable of the soil. This indicated that grain yield was very location specific and related, next to treatment effects, to local climate and soil conditions. Excluding management, straw yield variability was explained for approximately 38% by including N-fertilizer input, the soil stoniness, soil P-content and the slope of the field as predictors. This indicated strong location dependent variability. Again excluding management, fertilizer responses were mainly explained by soil characteristics, which together with the inputs explained almost half of the total response variability. Focusing specifically on the relation between soil properties (N-total, P-available and K-exchangeable) and response to recommended fertilizer application we found this relation indeterminate, except for N. Differences in yields between recommended application and farmer managed fields were limited and non-significant. We concluded that (1) defining best practices is a location specific and tailor-made task which requires the involvement of farmers to deal with local preferences and context and (2) on-farm experimentation includes such local environment and farmer-related variability. Our participatory approach using on-farm experimentation demonstrated why a one-size-fits-all strategy, i.e. blanket recommendations, will not work unconditionally in Tigray. Both grain and straw yield were determined by the complex local interplay of farmer management, soil properties, landscape and fertilizer input.

1. Introduction

1.1. Back ground

Traditionally agronomists often use on-station experimentation to design and test novel technologies that aim at sustainable increase of crop yield. In order to achieve reliability, replications and controls are always included in the experimental design. Experimental lay-out in most cases follows complete randomized block designs. The location variability is kept to a minimum by selecting flat non-shaded locations with uniform and usually deep well-drained soils. Local management is usually high tech. As a consequence of these choices, outcomes of on-station trials in Africa tend to be quite different compared to the actual situation in most farmer fields. Mugwe et al. (2009), for example, reported for maize, in response to specific treatments, an up to 50% higher yield on-station as compared to on-farm experiments. Due to its standardized conditions and procedures on-station research is considered more scientific and able to identify causal relationships (Johnston et al., 2003).

On-farm experimentation is, more than on-station experimentation, seen as an appropriate way to inform farmers about novel

---

* Corresponding author.
E-mail address: richard.kraaijvanger@wur.nl (R. Kraaijvanger).

http://dx.doi.org/10.1016/j.agee.2015.08.003
0167-8809/

technologies (Chambers and Jiggins, 1987). Objectives of on-farm experimentation relate to demonstration, testing or fine-tuning of novel technologies such as new crop varieties, fertilizer application or pest management. In many cases in on-farm experimentation, pre-defined technologies are evaluated to determine their suitability in a specific local context. Consequently, the reverse question, i.e. what technology is required in a specific context, is often ignored.

In on-farm experiments conditions are typically less uniform than in the case of on-station field trials. Consequently, substantial observed variability is controlled by local diversity in environmental conditions and farmer practices (Raman et al., 2011). If local variability in conditions and management is known, its relevance can be quantified and studied, allowing, in this way, an evaluation of the proposed technology. An important disadvantage of on-farm experiments in relation to intervention work is that outcomes are locally specific and, therefore, non-transferrable to other locations and conditions (Johnston et al., 2003).

1.2. Context

In our study area, Tigray in northern Ethiopia, crop yield in general is low (Vancampenhout et al., 2006; Habtegebrial et al., 2007; Tsegay, 2012; Kraaijvanger and Veldkamp, 2014) and is often not sufficient to sustain rural families. Since alternative livelihood options are scarce, this often leads to food insecurity. This lack of food security is counteracted by food-aid in the form of Food-For-Work-programs (FFW) or direct aid. In 2011 about 40% of the rural households depended for 1 month or more on such external support. Given this dependency, increasing crop yield is considered an important option to achieve sustainable development of these rural communities. In the densely populated Tigray increased crop production can only be achieved by attaining higher crop yields. One of the main identified yield constraints in Tigray is soil fertility, which resulted in an extensive promotion of fertilizer use (Kassie et al., 2009) by the regional and woreda level Bureaus of Agriculture and Rural Development (BoARD) and different NGO’s.

Within the framework of our research on participatory farmer experimentation we focused on soil fertility because the farmers involved identified this as a major opportunity to improve crop yield (Kraaijvanger et al., 2014). In a participatory process we facilitated, by using focus group discussion, farmer groups to design different experiments. This resulted in a series of experiments that were conducted on-farm in four different areas in Tigray. Consequently, farm management and environmental factors, like climate, soil type and topography, and related to that yield potential were different (for details see Kraaijvanger and Veldkamp (2014) and online supplementary file). Our involvement secured that control experiments and replications were included and that the experimentation followed standardized procedures. Fertilizer quality and size of the experimental plots were constant and all fields were relatively flat. All measurements were done by the scientific team. In total four years of on-farm experimentation resulted in an extensive data-set on achieved yield, responses to fertilizer treatments and local environmental and farm management characteristics. In this paper only experiments with documented inputs were included.

1.3. Research questions

Different treatments in our experimentation were expected to give differences in yields. Besides that, the low level of control in on-farm experimentation likely contributed to outcome variability by allowing environmental factors to become explicit. Constraints, as identified by the farmers involved were found to be depending on location (Kraaijvanger et al., 2015). Previous research (Veldkamp et al., 2001) indicated that a significant part of yield variability can be explained by local field (i.e. site) and farm (i.e. management) variability. In line with this we hypothesized that at meta level, in addition to treatment effects, a substantial part of the yield variability in on-farm experimentation can be explained by local environmental and management factors.

Our research questions:

1. To what extent can, at meta level, on-farm yield variability be explained by treatment effects, by environmental factors and by management factors.
2. How does, at site level, yield achieved in on-farm experimentation, relate to treatment, location and soil properties.

2. Materials and methods

2.1. Field experimentation

In total 16 farmer groups were involved, coming from four administrative units (*woredas*) in Tigray. These *woredas* were Weri Lekhe, Hawzen, Ahforom and Dogua Tembien. Our experimental sites were located nearby the administrative centres of these *woredas* (respectively Edaga Arbi, Hawzen, Intcho and Hagere Selam). In this paper we refer to these administrative centres. Important environmental differences between our experimental sites in these *woredas* related to altitude (Hagere Selam above 2300 m, others around 2000 m), climate (annual precipitation Hagere Selam 850 mm, others around 600 mm), parent material (Hawzen sandstone and shale, others mainly basalt) and soil type (Hawzen mainly Cambisols, others Luvisols and Vertisols (FAO-IUSS, 2006)). The sites also varied in crops grown (in all sites wheat and teff were important, in Inticho also sorghum). For additional details about the locations we refer to (Kraaijvanger and Veldkamp, 2014) and our online supplementary file. In our joint experimentation programme farmer groups were challenged, for four years, to design experiments with the objective to achieve improved crop yield. In addition to these farmer experiments, science based experiments were included. In this experimentation three crops were involved: wheat (Triticum aestivum, teff (Eragrostis tef) and hanfets. Hanfets is a traditional local mixture of wheat and barley (Hordeum vulgare). In this paper we focused on wheat, which was included in the experiments by about 60% of the farmer groups. Farmer based experiments were extremely diverse and involved, for example, combinations of organic and mineral fertilizers. Science based experiments involved recommended application of urea and DAP, additional supply of potassium and sowing in rows (see Table 1). In contrast to the farmer experiments, these science based experiments in most cases were replicated and included controls (see Table 2). In this paper the analysis relates mainly to two science based treatments with wheat: (1) recommended application of urea and Diammonium Phosphate (DAP, containing N and P) and (2) recommended application of urea and DAP plus additional potassium fertilizer (containing N, P and K). In addition, yields of controls, treatments with documented inputs (e.g. single

Table 1
Characterization different treatments.

Code	Description
C	Control, no application of fertilizer
FF	Farmer field hosting the experimental block
R	Recommended application of fertilizer (100 kg/ha urea + 100 kg/ha DAP for Edaga Arbi, Inticho and Hagere Selam, 50 kg/ha urea + 100 kg/ha DAP for Hawzen)
R + K	Recommended application + 94 kg/ha KCL

Table 2
Overview of replications per site involved in on-farm experimentation for four years.

Period	controls	Recommended application of urea and DAP	Recommended application of urea and DAP and additional potassium
February 2010–November 2010	3	3	3
May 2011–November 2011	2	2	0–3
May 2012–November 2012	2	1–3	0–3
May 2013–November 2013	2	1–3	0–3

application of urea) and hosting farmer fields (FF) were included. In this evaluation only the first year of experimentation in a specific site was considered to avoid residual effects of fertilizer and manure applications.

2.2. Experimental management

The experiments were conducted on-farm in fields selected by the farmer groups. All fields were terraced, which is a common practice in Tigray. Within these fields conventional experimental blocks containing the different treatments were positioned central (Fig. 1). In most cases experimental blocks were composed of 15 plots in 3 rows of 5 plots, with the long rows along the contour lines. The plot size was 9.0 m^2 (3.0 × 3.0 m). In 2010 the farmer-based treatments were positioned in the centre row for demonstration purposes; scientist-based treatments were distributed random over the remaining rows. In 2011–2013 all treatments were distributed randomly over the three rows. In the case of replications we considered lower, middle and higher positions in order to deal with fertility gradients in terraced fields (Vancampenhout et al., 2006). Main treatments (control, recommended application of NP and application of NPK-fertilizer) were replicated to deal with variability within the experimental block (see Table 2). This variability was assumed to be high, even in small fields, and related to processes like terracing, water redistribution and the resulting erosion and sedimentation (Vancampenhout et al., 2006).

A few weeks before sowing, composite samples of the topsoil (0–15 cm) within the experimental block were collected and analysed for total N (Kjeldahl method (van Reeuwijk, 2002)), available P (Olsen method (van Reeuwijk, 2002)), exchangeable K (ammonium acetate extraction (Okalebo et al., 2002)), organic-C (based on loss on ignition (NEN, 2014)), clay content (estimated using reference samples) and stone content (2 mm sieve). At the same time we delineated the extremities of the block and requested the responsible farmer not to apply fertilizers, manure or compost within this block. At sowing-time farmers broadcasted the seeds before final ploughing. Immediately after ploughing we applied, for each of the treatments, the required amounts of mineral fertilizers (with uniform composition) and incorporated them into the top soil. In most cases about 1 month after sowing, additional urea was applied, according to BoARD-recommendations, as a top dressing. All other management of the fields, like weeding and crop protection was done by the farmers, in most cases the owner of the field.

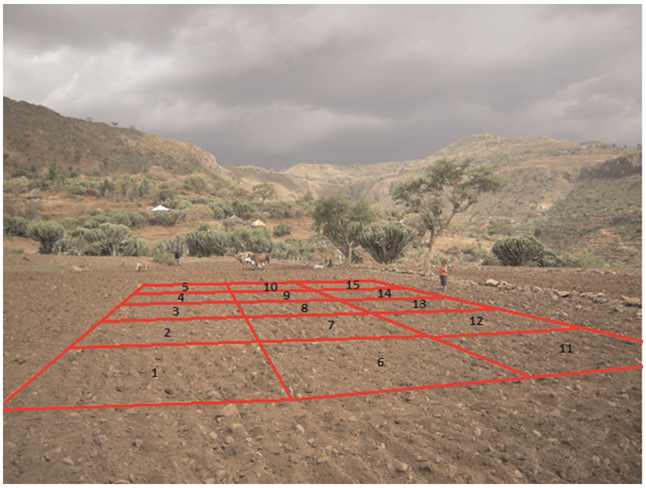

Fig. 1. Experimental block within the (hosting) farmer field in Inticho.

Harvesting was done by the research team, taking two random 1.0 m^2 samples within each plot. Representative samples (in duplo) from the hosting farmer field (FF) were taken adjacent to the experimental block. Harvesting was done manually. The crop was first cut and then weighed immediately to determine total biomass. After that the grains were separated by hand and also weighed. Chaff was removed in the traditional way by wind. Composite samples were taken to determine moisture content for grain and straw. Three experimental sites were excluded from the analysis due to damage caused by hail and flooding.

2.3. Analysis of outcomes

2.3.1. Calculation of yield and responses

Yield (in kg/ha) was calculated using measured dry matter yield of grain and straw based on the individual plots.

Responses were calculated using the following ratio:

$$\text{Response} = 100 \times (Y_T - Y_C)/Y_C$$

Y_T = yield treatment (kg dry weight/ha)

Y_C = yield control (kg dry weight/ha)

2.3.2. Overall variability assessment

Yield and response variability were statistically related to treatments, site and management. Treatments were different in N, P and K inputs. Site characteristics were divided into 3 main groups: soil, climate and landscape (see Table 3). Soil properties we considered were soil organic carbon, stone content, clay content, nitrogen content, phosphorus content and potassium content (all of the top soil). Included climate factors were annual rainfall, mean maximum temperature (based on tabia, *woreda* and literature data) and length of growing period (based on experimental data). With respect to landscape we considered altitude (GPS measured) and field slope. To assess the potential impact of farmer management we classified the management level, based on observations concerning weeding and soil protection (terracing), into 3 levels.

Statistical analysis was done on a plot based data-set, containing 128 data-points for wheat grain yield, 119 data-points for wheat straw yield, 102 data points for grain response and 93 data points for straw response. Variables were tested for normal distribution by using Q–Q plots. Only the normally distributed

Table 3
Variables used in the multiple linear regression models.

Variable	Unit	Range	Source
N-input	kg N/ha	0–75.5	experimental
P-input	kg P/ha	0–24	experimental
K-input	kg K/ha	0–49.1	experimental
N-total	mg/kg	360–2380	laboratory
P-available	mg/kg	3.67–83.3	laboratory
K-exchangeable	mg/kg	31–858.9	laboratory
Organic-C topsoil	%	0.5–4.7	laboratory
Clay topsoil	%	8–55	estimation
Stoniness topsoil	%	0.3–46.9	laboratory
Altitude	m asl	1966–2639	GPS
Slope	%	1–12	clinometer, estimation
Temperature (mean annual maximum)	°C	23–27	literature (Gebrehiwot and Van der Veen, 2013) for 2008.
Rainfall (annual)	mm	535–850	*tabia*- and *woreda* data (averages for the years 2005–2008 of longer)
Length of growing period	days	94–128	field observations
Management- level		1–3	field observations

variables were used in multiple linear regression analyses to estimate significant regression models.

Assuming linearity the following general relationship can be provided to predict yield:

$$\mu_y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \text{ .... } + \beta_n x_n$$

μ_{y} = yield of plot y

β_0 = constant

$\beta_{1\text{-}n}$ = coefficient for variable x

$x_{1\text{-}n}$ = variables

In our case we considered produced grains and straw (expressed as dry matter) and response (grain and straw) as indicators for yield. For the regression model of wheat response we excluded control treatments from the data-set. Estimations for β_0–β_n were obtained by calculating multiple linear regression models using SPSS. For our exploration we used stepwise and backward elimination techniques. The model fit of the regression-equations obtained was expressed by the coefficient of determination (R^2). This R^2 was used to indicate the total variability as explained by the included variables (=predictors). Standardized Beta-values and semi-partial correlation coefficients were calculated to indicate the relative contribution of individual and clusters of predictors. Standardized Beta-values informed us about the relative importance of a specific variable in the regression-equation; semi-partial correlation coefficients informed us about the unique contribution of a specific variable to the explanation of variability.

2.3.3. Yield per treatment and location

To evaluate the impact of specific treatments on crop yield we used both treatment and location as a unit of analysis. Average grain and total biomass yield of wheat were calculated based on site averages for a specific year.

2.3.4. Responses and soil properties

The responses to recommended application of urea and DAP (100 kg DAP/100 kg urea/ha, except for Hawzen 100 kg DAP/ 50 kg urea/ha) were related to different soil properties (N-total, P-av, K-exch and organic-C) and analysed for direct correlation. Soil nutrient content (N-total, P-av and K-exch) was evaluated using the ratings as provided by Landon (1991) (Table 4). In addition, based on these ratings, we defined three soil quality groups (low, medium and high) that matched our data range (Table 4). To analyse differences in responses between these soil quality groups we used ANOVA.

3. Results and Discussion

3.1. Yield and response variability

3.1.1. Overall explanation of wheat yield variability

Including all variables in the regression allowed us to estimate the contribution of the 5 main group of variables (inputs, soil, landscape, climate and management level) in explaining grain yield (Tables 5 and 6). For this purpose semi-partial correlation coefficients were used to estimate unique contributions of specific variables to the explanation of crop yield variability.

When comparing the contributions of the different clusters of variables we observed that wheat yield is mainly related to environmental factors (soil and landscape), management and input (treatment). The unique contribution of climate related variables is relatively limited. Overall, this means that farmer-impact (management) by about 45% was most important in explaining yield variability, environmental factors came second by about 31% and treatment factors third by about 24%.

3.1.2. Significant relationships

Five different significant multiple linear regression models were derived from our data-set (Table 7): (1) grain yield, (2) grain yield (including management-level), (3) straw yield, (4) response of grain to fertilizer inputs and (5) response of straw to fertilizer inputs.

Table 4
Ratings and corresponding soil quality groups for soil nutrient content.

Rating[*]	Very low	Low	Medium	High
N-total (mg/kg)	<1000	1000-2000	2000-5000	>5000
P-available (mg/kg)	–	<10	10-20	>20
K-exchangeable (mg/kg)	–	<60	60-120	>120
Soil quality group	Group I (low)	Group I (low)	Group II (medium)	Group III (high)

[*] Ratings are based on (Landon, 1991).

Table 5
Outcomes of linear regression using all variables.

Dependent	Grain yield dm ($R^2 = 0.616$)		
Predictors	Beta	Standardized Beta	Semi-partial coefficient of correlation
Constant	-4.30×10^3		
N-input*	11.8	0.347	0.238
P-input	0.705	0.010	0.007
K-input	0.650	0.015	0.014
N-total	0.220	0.131	0.080
P-av	−8.65	−0.168	−0.094
K-exch	0.687	0.175	0.101
Organic-C	51.5	0.057	0.036
Clay	−10.3	−0.189	−0.087
Stone%	3.67	0.062	0.042
Altitude*	1.80	0.447	0.148
Slope	−54.7	−0.208	−0.116
Rainfall	0.457	0.067	0.016
Temperature	−4.68	−0.009	−0.004
Growing period	6.18	0.064	0.025
Management level*	509	0.433	0.324

* Significant ($p = 0.05$).

Table 6
Unique explanation of predictors (all being included).

All predictors included	Included predictors	% Of total unique variance*
Sum inputs	N-input, P-input, K-input	24.3
Sum soil	stone%, clay, organic-C, N-total, P-av, K-exch	15.4
Sum landscape	altitude, slope	15.1
Sum climate	rainfall, temperature, growing period	0.4
Management-level	management- level	44.8

* Unique variances were calculated as the sum of squares of the semi-partial correlation coefficients (SSSP) of the predictors involved. To determine% of total unique variance we calculated the ratio of the cluster-based unique variance and the total unique variance (based on all predictors included in the regression model).

The multiple linear regression techniques applied resulted in the following regression equations:

$$\text{Wheat yield (grain)} = -3.66 \times 10^3 + 2.99 \times \text{altitude} + 12.7 \times \text{N-input} + 3.60 \times 10^{-1} \times \text{N-total} - 3.81 \times \text{rainfall} + 210 \times \text{organic-C} + 6.52 \times 10^{-1} \times \text{K-exch}\ (R^2 = 0.487) \tag{1}$$

The regression model indicated that grain yield variability was explained for almost 50% by six variables. Standardized B-coefficients and semi-partial correlation coefficients indicated that in the case of the regression model for grain yield, altitude was the most important explanatory factor, N-input came second, followed by 3 soil variables (N-total, organic-C and K-exch). Rainfall had a negative contribution on yield within the regression model. It appeared that of the treatment factors especially N-input was relevant in explaining yield variability, of the site-factors both landscape and soil factors were important. It is clear that N-input positively contributed to grain yield.

$$\text{Wheat yield (grain) including management} = -3.49 \times 10^3 + 1.74 \times \text{altitude} + 12.3 \times \text{N-input} + 536 \times \text{management-level}\ (R^2 = 0.560) \tag{2}$$

When the wheat grain yield model included the estimated management level, a small increase in total explained variability was observed. In the model only 3 explanatory factors remained: altitude, management level and N-fertilizer input. The strong contribution of management-level in this regression was striking. This was supported by our field observation that the visually best managed fields (no weeds) usually were also the most productive ones. This might be either due to the effort of the farmer or due to the fact that farmers, being constrained in labour, invested less time in less productive fields.

$$\text{Wheat yield (straw)} = 2.37 \times 10^3 + 29.2 \times \text{N-input} - 48.2 \times \text{stone\%} + 42.4 \times \text{P-av} + 159 \times \text{slope}\ (R^2 = 0.383) \tag{3}$$

The regression model for straw yield performed less well than that for grain yield. However, still more than 38% of the total variability was explained by the model. N-input was the most important explanatory variable, followed by two soil variables (P-content and stoniness) and one site variable (slope). In contrast to grain yield, N-total did not influence straw yield.

$$\text{Wheat yield response (grain)} = -112 + 7.04 \times 10^{-1} \times \text{N-input} - 5.50 \times 10^{-2} \times \text{N-total} + 13.3 \times \text{organic-C} + 1.05 \times \text{clay}\ (R^2 = 0.411) \tag{4}$$

The regression model for wheat grain response explained 41% of the total variability. In this regression model N-total content of the topsoil was an important (negative) predictor: this meant that responses to fertilizer application (in our case in the form of urea, DAP and KCl) on the more fertile soils were likely to be smaller than on the poorer soils. The yield response was almost proportional to N-input; other important positive predictors were the topsoil properties organic-C and clay content.

$$\text{Wheat yield response (straw)} = 33.8 + 1.09 \times \text{N-input} - 8.55 \times 10^{-1} \times \text{stone\%} + 1.34 \times \text{clay} - 5.00 \times 10^{-2} \times \text{N-total} + 5.50 \times 10^{-2} \times \text{K-exch}\ (R^2 = 0.552) \tag{5}$$

The model for wheat straw response explained more than 55% of the total variability. This model demonstrated the importance of fertilizer input and soil factors. It seemed that using N-inputs on the "better soils" (high clay and high potassium content) straw responses were higher. This result contrasted with the model for grain responses. The contribution of N-input and clay content to grain response was less and N-total even a more negative factor.

Table 7
Results of the multiple linear regressions.

Dependent	Predictors	Beta	Standardi-zed Beta	Semi-partial coefficient of correlation
Grain yield dm[a] ($R^2 = 0.487$)				
	Constant	$-3.66 \times 10^{3*}$		
	Altitude	2.99^{*}	0.745	0.436
	N-input	12.7^{*}	0.374	0.358
	N-total	$3.60 \times 10^{-1*}$	0.215	0.151
	Rainfall	-3.81^{*}	−0.556	−0.272
	Organic-C	210^{*}	0.234	0.184
	K-exch	$6.52 \times 10^{-1*}$	0.166	0.134
Grain yield dm (incl. management)[a] ($R^2 = 0.560$)				
	Constant	$3.49 \times 10^{3*}$		
	Altitude	1.74^{*}	0.434	0.430
	Management-level	$5.36 \times 10^{2*}$	0.456	0.456
	N-input	12.3^{*}	0.363	0.360
Straw yield dm[a] ($R^2 = 0.383$)				
	Constant	$2.37 \times 10^{3*}$		
	N-input	29.2^{*}	0.435	0.429
	Stone%	-48.2^{*}	−0.403	−0.346
	P-av	42.4^{*}	0.406	0.378
	Slope	159^{*}	0.298	0.262
Response grain dm[b] ($R^2 = 0.411$)				
	Constant	$-1.12 \times 10^{2*}$		
	N-input	$7.04 \times 10^{-1*}$	0.268	0.242
	N-total	$-5.50 \times 10^{-2*}$	−0.495	−0.411
	Organic-C	13.3^{*}	0.226	0.166
	Clay	1.05^{*}	0.298	0.204
Response straw dm[a] ($R^2 = 0.552$)				
	Constant	33.8^{*}		
	N-input	1.09^{*}	0.458	0.420
	Stone%	$-8.55 \times 10^{-1*}$	−0.236	−0.217
	Clay	1.34^{*}	0.413	0.330
	N-total	$-5.00 \times 10^{-2*}$	−0.493	−0.352
	K-exch	$5.50 \times 10^{-2*}$	0.242	0.200

[*] Significant ($p = 0.05$).
[a] Stepwise regression.
[b] Backward elimination.

This outcome probably related to the use of high-straw yielding varieties. Such varieties respond to N-input primarily by producing straw rather than grains. Despite the availability of improved short straw varieties the traditional long straw varieties are often more appreciated by the farmers.

3.2. Soil properties

Based on the ratings defined (see Table 4) N-total content over the whole range of sites pointed to a limited availability (Table 8). Only in 2 sites in Hagere Selam and 1 in Edaga Arbi nitrogen was having a medium availability. P-av and K-exch appeared adequate in most cases. Still, in all *woredas* soils were found that were low in P. A possible cause for this might be long-term depletion. Potassium availability was limited in three sites in Hawzen. The specific sandstone and shale parent materials found in Hawzen, are reported to be responsible for the low availability of potassium (Murphy, 1959).

3.3. Impact of treatment on yield

Grain yield of wheat, compared to the controls, increased from about 1500 to about 2200 kg/ha in case of recommended application of urea and DAP (Fig. 2). In case additional KCl was applied the increase of grain yield was even up to about 2500 kg/ha. However, differences between the grain yield of farmer fields (FF) and recommended application (R) and application of NPK (R + K) were not significant ($p = 0.05$). In most cases farmers applied

Table 8
Overview of nutrient composition of the top soil for all experimental sites.

Location	Total sites	Org-C average (range) %	N-total average (range) mg/kg	P-avail average (range) mg/kg	K-exch average (range) mg/kg	Limitations				
						No	N	NP	NK	NPK
Edaga Arbi	12	2.4 (1.3–3.5)	1114 (550–2380)	28 (8–83)	341 (66–859)	1	10	1	–	–
Hawzen	12	1.4 (0.5–2.2)	788 (400–1610)	18 (4–38)	170 (31–626)	–	8	1	1	2
Inticho	12	3.4 (2.1–4.7)	1075 (690–1460)	29 (6–48)	281 (73–597)	–	11	1	–	–
Hagere Selam	9	3 (2.2–4.4)	1623 (970–2040)	27 (6–40)	328 (113–587)	2	6	1	–	–

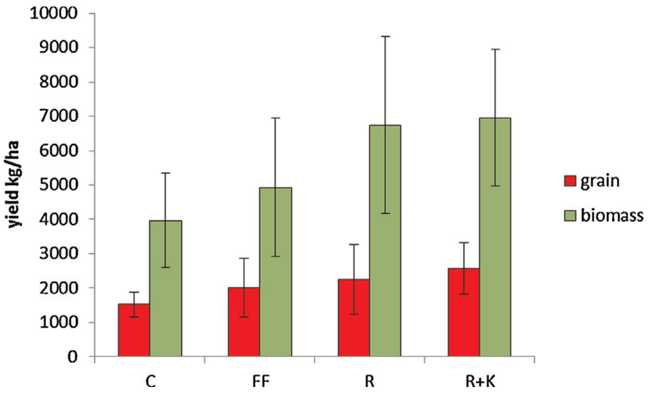

Fig. 2. General overview of average yield for wheat for 4 years (on site basis). Standard deviation is indicated by error bars. (C = control, FF = farmer field, R = urea + DAP, R + K = urea + DAP + KCl)

combinations of organic fertilizers and limited amounts of urea and DAP on their fields; these combinations appeared to be quite effective.

Recommended application (R) and application of NPK (R + K) had considerable effect on biomass yield, which almost doubled. Differences in biomass yield between FF and R and FF and P were significant ($p = 0.05$).

The application of NPK (R + K) provided results comparable to recommended application of urea and DAP (R), and consequently, did not result in significant differences in grain or biomass yield ($p = 0.05$).

3.4. The impact of location on yield

Highest grain yield in the control plots was achieved in Hagere Selam, lowest in Inticho (Fig. 3). Hagere Selam is a highland area with an altitude of around 2300 m with lower average temperatures and a higher amount of rainfall. These conditions lead under natural conditions to soils with higher organic matter content, which again leads to a higher natural soil fertility and a higher yield. In the case of both Inticho and Edaga Arbi altitude and soil type were similar. Still, soils in Inticho were less productive for all treatments. This might be a consequence of a more intensive land use and the frequent inclusion of sorghum in the rotation which might have resulted in nutrient mining (Kraaijvanger and Veldkamp (2014); personal observations first author). The low N-content of many of the Inticho soils also pointed to depletion (see Table 8).

Edaga Arbi and Hawzen achieved comparable yields for recommended application of DAP and urea. This was surprising since recommended application of urea in Hawzen was only half of that in Edaga Arbi. The location with a lower recommended fertilizer application (Hawzen) showed, as one would expect, a lower, but non-significant ($p = 0.05$), response to this application.

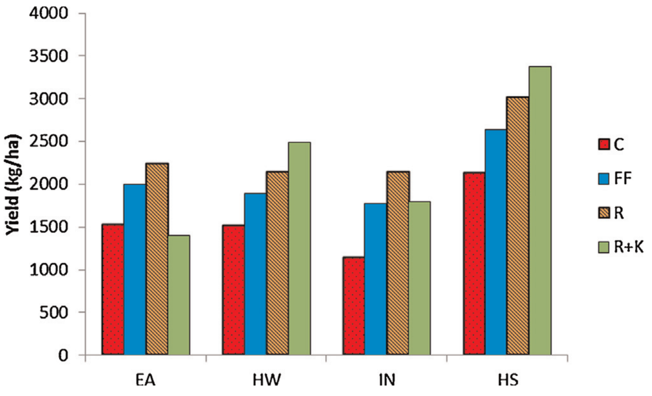

Fig. 3. Overview of grain yield for 4 locations (on site basis) for wheat. (EA = Edaga Arbi, HW = Hawzen, IN = Inticho, HS = Hagere Selam, C = control, FF = farmer field, R = urea + DAP, R + K = urea + DAP + KCl)

The application of potassium, in combination with urea and DAP, led to a higher grain yield of wheat in the case of Hawzen and Hagere Selam. For Hawzen this matched with the soils that had a somewhat lower content of K-exch. The good response to potassium in the case of Hagere Selam probably related to the incidence of higher achieved productivities, which required a more adequate supply of K. In addition, potassium often has a positive effect on the uptake of nitrogen (Mengel and Kirkby, 1987).

3.5. The impact of soil properties on response to fertilizer application

In this correlation only the sites that received a recommended application of 100 kg/ha urea and 100 kg/ha DAP were included. Both available P and exchangeable K did not show significant correlation with response to recommended application of urea and DAP. Only in the case of N-total (Fig. 4), correlation between response to recommended application of 100 kg/ha DAP and 100 kg/ha urea showed a significant quadratic trend ($R^2 = 0.332$, $p = 0.05$). Low N-soils apparently benefited much more from the addition of N-fertilizer. High-N soils were productive in most cases and demonstrated limited response to application of urea and DAP.

Assuming that in Tigray total N will be mostly based on organic N, the availability of this organic N will depend on mineralization in the rainy season (Bartholomew and Clark, 1965). However, this mineralization will take some time and will be proportional to the content of organic N in the soil. Given the short growing period in Tigray (3 months), mineral fertilizers therefore will be effective to supply nitrogen at the start of the growing period and prevent delay in crop development, especially in the case of low N-soils.

This correlation between total N and crop response contrasted with the opinion that, in general, total N is not considered a good indicator for availability of N (Landon, 1991). However, correlation between availability and total N is assumed to improve when the soils considered fall within a relatively small range of soil fertility and overall local conditions are comparable (Page and Dinauer, 1982), which did apply in our case.

For K-exchangeable a weak quadratic trend ($R^2 = 0.1711$) was observed for the response to 100 kg/ha DAP and 100 kg/ha urea (Fig. 5). This appears to provide evidence for the concept of non-responsive poor soils and non-responsive fertile soils (Tittonell et al., 2008). In this case high potassium soils pointed to fertile soils, that were no longer responsive to the application of NP-fertilizers.

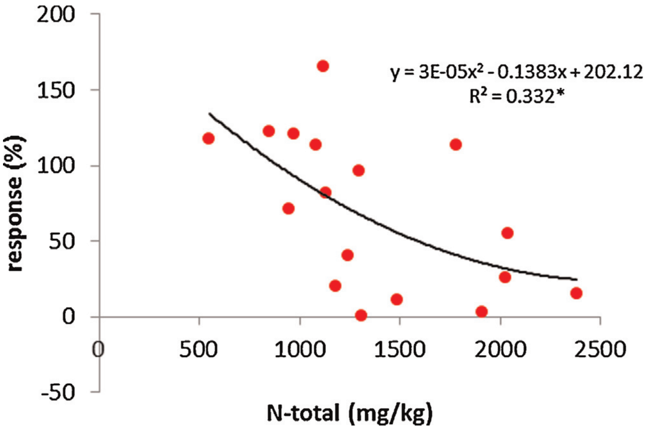

Fig. 4. Grain responses of wheat to recommended application of fertilizer (100 kg urea/ha and 100 kg DAP/ha) versus soil N-total. (*significant at $p = 0.05$)

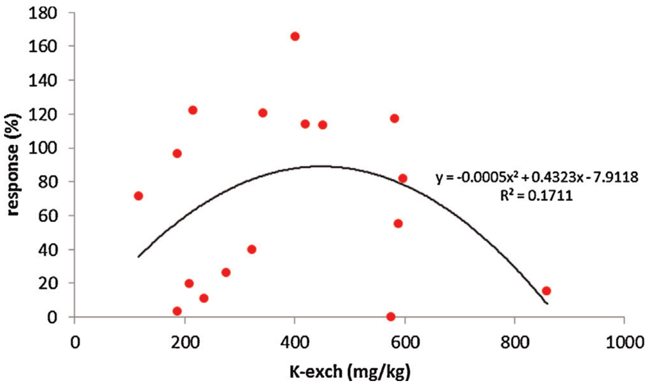

Fig. 5. Grain responses of wheat to recommended application of fertilizer (100 kg urea/ha and 100 kg DAP/ha) versus K-exch (soil).

3.6. Evaluating overall impact of treatment and environmental factors on yield variability

The regression model for grain yield (without management) demonstrated that fertilizer input and local environmental factors (climate, landscape and soil) were both important in explaining 49% of the total yield variability. N-input was the most important input factor, altitude the most important environmental factor. When management-level was also included in the grain yield model, the total explained variability increased to about 56 %.

Comparing unique contribution of the individual predictors, confirmed the picture of different variables having impact on yield variability. Management-level (45%) and environmental characteristics (31%) both accounted largely for the explanation of the grain yield variability. Variability explained by treatment factors appeared to have less explanatory weight (24%). This indicated that a large proportion of wheat yield variability in the studied farms was determined by local factors that were location (landscape and soil) and farmer (management) related (together 76%). Compared to this the impact of fertilizer inputs on yield variability was relatively limited.

3.7. Arriving at recommendations

One of the objectives of our joint experimentation was to identify relevant best practices to improve crop yield in the context of the farmers involved. Interpreting outcomes of experimentation we observed that yield responses to recommended application of DAP and urea were highly variable and differences with farmer fields were limited. The same held for the application of potassium. The small difference in response between farmer fields and recommended application of urea and DAP is important in relation to (1) the trade-off between fertilizer cost and yield increase and (2) farmer field and management (as in FF) being an important point of reference for farmers.

Most soils responded to NP-application, especially soils very low in N. However, responses were not convincing in all cases. This suggested that possibly other (micro) nutrients were a limiting factor. P and K are unlikely candidates because we observed in our experiments that available P and exchangeable K of the soil were medium or high in most cases. In such soils, high in P and K, one would expect an effective response to the application of limiting nitrogen. However, outcomes of our experiments did not confirm this. With respect to other (micro) nutrients, Habtegebrial Habtemichial et al., (2007) previously suggested that in Tigray, application of S might improve performance of legumes. Next to nutrients also other effects appeared to be important, notably, management and crop factors. A simple straightforward measurement of only NPK-status of the soil, therefore, does not seem sufficient to support recommendations.

Consequently, clear recommendations with respect to best practices cannot be provided based on our findings so far. The complexity of the agricultural system demanded more detailed research on different interactions. Particularly the combination of organic and mineral fertilizers, that is often applied successfully by farmers in Tigray, deserves due attention. The NUANCES-project (Rufino et al., 2007) and Integrated Soil Fertility Management (ISFM) approaches (Vanlauwe et al., 2010), for example, illustrate the importance of these combinations.

The disappointing response of nutrient inputs, as compared to actual practice, indicated that at farm-level probably other factors were more significant in increasing yield. Our analysis pointed to local management as a key factor. Total biomass yield responded better to recommended application of urea and DAP. Technically this offers the possibility to increase grain yield by using short-straw varieties. It is doubtful if farmers will follow such a recommendation because straw is considered important as a fodder for livestock in their mixed farming system (Kraaijvanger et al., 2015).

3.8. Yield variability in relation to on-farm experimentation

On-farm experimentation in our case resulted in highly variable outcomes. However, this variability was only to some extent caused by the low level of control in on-farm experimentation. Our findings demonstrated that treatment-factors, environmental characteristics and management were able to explain over 60% of the observed yield variability. This means that part of on-farm experimentation variability could be attributed to local site- and farmer characteristics. This closes the circle since on-farm experimentation is motivated by paying tribute to exactly these factors that represent the local context. The apparent lack of control in on-farm experimentation can be counteracted by quantifying the control by the local environment and by farmer management. This allowed us to gain more insight into the farming system at hand than a normal on-station trial could provide. Environmental and farmer characteristics, therefore, need to have a more pronounced position in the evaluation of outcomes of on-farm experimentation. Different tools to support such analysis and even the identification of causal relationships are available. Examples of such tools are the use of aggregated indices like the Environmental Index in Modified Stability Analysis, multiple linear regression or advanced statistical models (Hildebrand et al., 1993; Riley and Alexander, 1997; Raman et al., 2011)

4. Conclusion

Linear regression of the outcomes of on-farm fertilizer experiments with wheat in Tigray indicated that 56% of the grain yield variability was explained by local management, local environmental characteristics and treatment-effects. This implied that with on-farm experimentation also environment and farmer definitely mattered and need to be accounted for.

Outcome variability in on-farm experiments in different locations was high and no simple clear and relevant relationships could be identified. Differences between the various introduced treatments and farmer fields were limited and non-significant. The main limiting nutrient was N. Responses to recommended (NP) fertilizer application demonstrated significant negative correlation with N-content of the soil. The correlation between P- and K-content of the soil and response to recommended fertilizer application proved to be negligible and not very helpful in coming to general valid recommendations.

We concluded that defining best practices is a location specific and tailor-made task which requires the involvement of farmers and their fields to deal with local preferences and context. Local management and local environmental characteristics matter. On-farm experimentation involving farmers demonstrated why a one-size-fits-all strategy, i.e. blanket recommendations, will not work to solve all yield problems in Tigray. Yield was determined by the complex local interplay and interface of farmer management, soil properties, landscape and fertilizer input. This complexity is not likely to be addressed in traditional on-station research and its outcomes.

Acknowledgements

We thank all involved farmers and staff from Mekelle University and BoARD for their support. In addition, we thank Zemen Mesfin, Mohammed Nur and Semere Gashaw for their field assistance and Sylvia de Jager for supporting laboratory analysis. Also we want to express our gratitude to two anonymous reviewers for their thoughtful comments which helped us to upgrade the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.agee.2015.08.003.

References

Bartholomew, W.V., Clark, F.E., 1965. Soil Nitrogen. American Society of Agronomy, Madison.

Chambers, R., Jiggins, J., 1987. Agricultural research for resource-poor farmers Part I: transfer-of-technology and farming systems research. Agric. Adm. Extension 27, 35–52.

FAO-IUSS, 2006. World reference base for soil resources 2006: A framework for international classification, correlation and communication. FAO, Rome.

Gebrehiwot, T., Van der Veen, A., 2013. Assessing the evidence of climate variability in the northern part of Ethiopia. J. Dev. Agric. Econ. 5, 104–119.

Habtegebrial Habtemichial, K., Ram Singh, B., Aune, J.B., 2007. Wheat response to N2 fixed by faba bean (*Vicia faba* L.) as affected by sulfur fertilization and rhizobial inoculation in semi-arid Northern Ethiopia. J. Plant Nutr. Soil Sci. 170, 412–418.

Habtegebrial, K., Singh, B.R., Haile, M., 2007. Impact of tillage and nitrogen fertilization on yield, nitrogen use efficiency of tef (Eragrostis tef (Zucc.) Trotter) and soil properties. Soil Tillage Res. 94, 55–63.

Hildebrand, P.E., Singh, B.K., Bellows, B.C., Campbell, E.P., Jama, B.A., 1993. Farming systems research for agroforestry extension. Agrofor. Syst. 23, 219–237.

Johnston, L.J., Renteria, A., Hannon, M.R., 2003. Improving validity of on-farm research. J. Swine Health Prod. 11, 240–246.

Kassie, M., Zikhali, P., Manjur, K., Edwards, S., 2009. Adoption of sustainable agriculture practices: Evidence from a semi-arid region of Ethiopia. Natural Resources Forum. Wiley Online Library, pp. 189–198.

Kraaijvanger R, Almekinders C.J.M., Veldkamp, A., 2014. Identifying constraints and opportunities: a case study with farmers from Tigray. Submitted to NJAS—Wageningen Journal of Life Sciences.

Kraaijvanger, R., Sonneveld, M., Almekinders, C., Veldkamp, T., 2015. Comparison of methods to identify crop productivity constraints in developing countries. A review. Agron. Sustainable Dev. 35, 625–637.

Kraaijvanger, R., Veldkamp, T., 2014. Grain productivity, fertilizer response and nutrient balance of farming systems in Tigray, Ethiopia: a multi-perspective view in relation to soil fertility degradtion. Land Degrad Dev, in press.

Landon, J.R., 1991. Booker Tropical Soil Manual: a Handbook for Soil Survey and Agricultural Land Evaluation in the Tropics and Subtropics. Longman, Harlow.

Mengel, K., Kirkby, E.A., 1987. Principles of Plant Nutrition. International Potash Institute, Worblaufen-Bern.

Mugwe, J., Mugendi, D., Kungu, J., Muna, M.M., 2009. Maize yields response to application of organic and inorganic input under on-station and on-farm experiments in central Kenya. Exp. Agric. 45, 47–59.

Murphy, H.F., 1959. A report on the fertility status of some soils of Ethiopia. Experimental Station Bulletin No. 1. Imperial Ethiopian College of Agriculture and Mechanical Arts, Alemaya, Ethiopia.

NEN, 2014. NEN 5754Calculation of organic matter content on a mass basis as loss-on-ignition (in dutch). National Committee for Environmental Quality, Delft.

Okalebo, J., Gathua, K., Woomer, P., 2002. Laboratory Methods of Plant and Soil Analysis: a Working Manual. Tropical Soil Biology and Fertility Programme, Nairobi.

Page, A.L., Dinauer, R.C., 1982. Methods of Soil Analysis: Chemical and Microbiological Properties. American Society of Agronomy, Madison.

Raman, A., Ladha, J.K., Kumar, V., Sharma, S., Piepho, H.P., 2011. Stability analysis of farmer participatory trials for conservation agriculture using mixed models. Field Crops Res. 121, 450–459.

Riley, J., Alexander, C.J., 1997. Statistical literature for participatory on-farm research. Exp. Agric. 33, 73–82.

Rufino, M.C., Tittonell, P., van Wijk, M.T., Castellanos-Navarrete, A., Delve, R.J., de Ridder, N., Giller, K.E., 2007. Manure as a key resource within smallholder farming systems: analysing farm-scale nutrient cycling efficiencies with the NUANCES framework. Livestock Sci. 112, 273–287.

Tittonell, P., Vanlauwe, B., Corbeels, M., Giller, K.E., 2008. Yield gaps, nutrient use efficiencies and response to fertilisers by maize across heterogeneous smallholder farms of western Kenya. Plant Soil 313, 19–37.

Tsegay, A., 2012. Improving crop production by field management strategies using crop water productivity modeling: Case study of tef (Eragrostis tef (Zucc.) Trotter) production in Tigray, Ethiopia. Faculty of Bioscience Engineering. Katholieke Universiteit Leuven, Leuven.

van Reeuwijk, L.P., 2002. Procedures for Soil Analysis, 6th ed. ISRIC, Wageningen.

Vancampenhout, K., Nyssen, J., Gebremichael, D., Deckers, J., Poesen, J., Haile, M., Moeyersons, J., 2006. Stone bunds for soil conservation in the northern Ethiopian highlands: impacts on soil fertility and crop yield. Soil Tillage Res. 90, 1–15.

Vanlauwe, B., Bationo, A., Chianu, J., Giller, K.E., Merckx, R., Mokwunye, U., Ohiokpehai, O., Pypers, P., Tabo, R., Shepherd, K.D., Smaling, E.M.A., Woomer, P.L., Sanginga, N., 2010. Integrated soil fertility management: operational definition and consequences for implementation and dissemination. Outlook Agric. 39, 17–24.

Veldkamp, A., Kok, K., De Koning, G.H.J., Schoorl, J.M., Sonneveld, M.P.W., Verburg, P.H., 2001. Multi-scale system approaches in agronomic research at the landscape level. Soil Tillage Res. 58, 129–140.