Disentangling agronomic and economic yield gaps in Ethiopian wheat based systems for better targeting of development interventions (Yield Gap Wheat Ethiopia)

Report #1: Elaboration of work plans & data availability

Date: 2nd October 2018

Authors: João Vasco Silva, Tom Morley, Pytrik Reidsma and Martin K. van Ittersum

Wageningen University & Research

Collaborators: Frédéric Baudron, Kindie Tesfaye, Moti Jaleta

International Maize and Wheat Improvement Centre

1) Objectives of the project

The goal of this project is to decompose and understand wheat yield gaps in Ethiopia from a farm(ing) systems perspective in order to identify relevant management and technological interventions that have the potential to narrow yield gaps. Three objectives were defined at the beginning of the project (cf. project proposal) towards this goal:

Table 1. Objectives, timeframe and activities.

Project objective / result	Timeframe	Activity
1. Quantification and explanation of wheat yield gaps in Ethiopia	Six months	Development of Ethiopian wheat database from household survey data (LSMS) combined with the Global Yield Gap Atlas to quantify wheat yield gaps and associated policies.
2. Agronomically more detailed assessment of wheat yield gaps for data-rich areas. This includes identification of key TEDs, technologies and policy interventions to mitigate yield gaps, applicable to Ethiopian context.	Six months	Development of database for specific data-rich regions with detailed agronomic and socio-economic information on wheat production. Literature review of key scientific papers, reports and other available information on TEDs, technologies and in particular those most relevant for Ethiopia.
3. Synthesis: combination of previous steps	Three months	

The first two objectives require the quantification and explanation of wheat yield gaps in Ethiopia based on biophysical and agronomic information from the Global Yield Gap Atlas (GYGA) in combination with spatially explicit household survey data. To do so, we will build upon the framework presented by Silva et al. (2017) and van Dijk et al. (2017), which allows the decomposition of the overall yield gap into efficiency, resource (which can be further decomposed into allocative and economic) and technology yield gaps (see definitions below) and the identification of relevant technological and policy interventions corresponding to each of these yield gaps. An innovative aspect of this research is that technology extrapolation domains (TEDs) will be used for spatial upscaling of the results obtained in the previous step. The TEDs make it possible to identify areas beyond the specific survey areas which can benefit from identified technologies, or policy interventions, as well as more efficient targeting of new research and experimentation.

In the present report, we further elaborate the work plan for each of the objectives defined above as well as on the characteristics of the datasets compiled so far to be used within the project.

2) Problem definition

Ethiopia is the largest wheat producer in sub-Saharan Africa. Wheat is one of the major food security crops in Ethiopia and it is grown solely under rainfed conditions by ca. 4.7 million farmers on approx. 1.6 million ha. Currently, wheat yields in Ethiopia are about 24% of their water-limited potential (Yw), suggesting a considerable scope to increase wheat productivity in the country (www.yieldgap.org). Despite the large yield gap for wheat in Ethiopia, there has been considerable yield progress over the past years. According to official FAO statistics, wheat yields in Ethiopia increased at a rate of ca. 63 kg ha^{-1} yr^{-1} from 1.5 t/ha in 1993 to 2.7 t/ha in 2016. Nevertheless, this is still far below the water-limited yields that can be achieved under rainfed conditions.

Previous research showed that wheat yield gaps in Arsi, one of the wheat belts in Ethiopia, are largely attributed to the technology yield gap (the difference between Yw and the highest farmers' yields; Silva, under review). This means that technologies currently used do not allow farmers to approach Yw and that transformative changes are needed if yield gaps are to be narrowed. These may be related to the crop establishment method used by farmers (wheat is mostly broadcasted, not planted in rows; Nyssen et al.; 2011), type of varieties adopted (traditional vs improved varieties resistant against stripe rust; Shiferaw et al., 2014; Abro et al., 2017) and the low amounts of inputs applied, such as fertilizers (e.g., less than 75 kg N/ha; Habte et al., 2014; Silva et al., under review) and their interaction with hand-weeding (Tanner et al., 1993). Competition for labour during sowing, weeding and harvesting of wheat were also observed in Arsi area as labour peaks for other cereal and legume crops overlap with the ones for wheat. This results in potential trade-offs between crops at farm level and requires a deeper understanding of the farming system aspects in which wheat is currently cultivated.

The prospects of population growth in Ethiopia put pressure on the natural resource basis of current agricultural systems and require that yield gaps are narrowed if self-sufficiency is to be achieved at national level (van Ittersum et al., 2016). To do so, we need to understand the drivers of existing yield gaps and identify management and technological options contributing the intensification of farming systems. A recent econometric analysis combining census data with remote sensing provides insights into wheat yield gaps in Ethiopia (Mann and Warner, 2017). Here, we propose to complement this econometric approach with an integrated (biophysical and econometric) approach decomposing the yield gap into efficiency, resource (allocative and economic) and technology yield gaps and to identify relevant management options and policies that can help narrowing these gaps. Two streams of analyses are defined for this purpose, one at national level and one at local level.

3) Elaboration of work plans

The workflow defined in the beginning of the project (cf. Table 1) will be used as a guideline in the elaboration of concrete work plans. A Gantt chart (Table 2) of the main activities, and related tasks, and a description of the activities planned are provided below.

- **Gantt chart of planned activities**

Table 2. Gantt chart. TM = Tom Morley, MvL = Marloes van Loon, JVS = Joao Vasco Silva.

Activities	Contact person	2018					2019					
		Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
1) National Yg analysis												
a) Cleaning LSMS data	TM	X										
b) Analysing LSMS data	TM		X	X	X	X						
c) Writing progress report	TM						X					
2) Simulations of Yw												
a) Gathering weather data	MvL		X	X								
b) Cleaning weather data	MvL				X	X						
c) Crop model simulations	MvL						X					
3) Local Yg analysis												
a) Cleaning WAIS data	JVS				X	X						
b) Analysing WAIS data	JVS						X	X	X	X		
c) Up-scaling with TEDs	JVS									X	X	
d) Writing final report	JVS									X	X	X

- **Activity 1: National analysis of wheat yield gaps in Ethiopia**

The LSMS for Ethiopia contains crop cut data for a randomly selected subpopulation of wheat fields. Crop cuts are recorded by interviewers, according to a consistent protocol, in standard units at, or close to, the time of harvest. In comparison, farmer recall data is often reported in unreliable non-standard units and, in the case of Ethiopia, are recorded several months after harvest casting doubt on accuracy of the recollection.

The primary use of the crop cut wheat yields is to quantify the wheat yield gaps relative to the GYGA derived water limited potential wheat yields. The LSMS data is spatially explicit and defined by GPS household coordinates. This facilitates the quantification of sub national wheat yield gaps as well as the use of external spatial data on climate, soil and other factors which can be derived from GYGA and other sources. In addition, binary and categorical information on crop management variables such as labour, mechanization, seed types, and input use are recorded in the data at the field, parcel and household levels. These variables will be related to crop cut wheat yields through ordinary least squares and, potentially, stochastic frontiers estimation to identify the most productive farmers.

- **Activity 2: Simulation of water-limited yields (Yw)**

The protocols, data and tools developed and used in GYGA will be used simulate Yp and Yw of wheat in Ethiopia during the period 2012 - 2017. Currently, simulations of Yp and Yw for wheat in Ethiopia are available in GYGA for the period 1998 – 2011, while the household survey data compiled so far refer to more recent years (see below). Thus, it is important to update the simulations of Yp and Yw in order to define yield benchmarks for the same years the household surveys were conducted. As explained in Activity 3, Yp and Yw are crucial to quantify the magnitude of the (technology) yield gap and express the actual yields and the yield gap in relative terms (i.e., as a proportion of Yw).

The data needed to conduct this activity refers to the weather data (i.e., minimum temperature, maximum temperature, rainfall, relative humidity, wind speed and solar radiation) for the period 2012 – 2017 from the weather stations previously considered in GYGA. We have approached dr Kindie Tesfaye to access this information and to get guidance on the underlying agronomic details required to conduct the simulations.

- **Activity 3: Local analysis of wheat yield gaps in Ethiopia**

The local yield gap analysis will make use of frameworks described in Silva et al. (2017) and van Dijk et al. (2017). These frameworks differentiate at least four yield levels to decompose the yield gap. The **actual yields** (Ya) refer to the yields observed in farmers' fields which are recorded in household surveys. **Technically efficient yields** (Y_{TEx}) comprise the maximum yield that can be achieved for a given input level and they can be computed using methods of frontier analysis in combination with concepts of production ecology. The **highest farmers' yields** (Y_{HF}) refer to the maximum yields (average above the 90^{th} percentile of actual yields) observed in a sample of farmers sharing similar biophysical conditions (weather and soils) and similar technologies adopted (e.g., varieties). Differently from Y_{HF}, van Dijk et al. (2017) considers **economic yields** (Ye) and **feasible yields** (Yf): the former refers to yield level in which marginal costs are equal to marginal revenue and the latter refers to the maximum yield that can be reached with available technology ad best-practice management but without economic constraints. Finally, the **water-limited yield** (Yw) refers to the maximum yield that can be obtained under rainfed conditions in a well-defined biophysical environment; crop growth simulation models are generally used to estimate Yw (see Activity 2). For the local yield gap analysis, we will make use of the Wheat Adoption and Impact Surveys collected and the FACASI household survey by CIMMYT to estimate Ya, Y_{TEx} and Y_{HF} and of GYGA to estimate Yw. We requested access to these data to dr Moti Jaleta and the data were already shared with us.

Three intermediate yield gaps can be distinguished based on the aforementioned yield levels. The **efficiency yield gap** is defined as the difference between Y_{TEx} and Ya and it is explained by crop management imperfections related to time, form and/or space of the inputs applied. The **resource yield gap** is defined as the difference between Y_{HF} and Y_{TEx} and it captures the yield penalty due to a sub-optimal amount of inputs applied. According to van Dijk et al. (2017), the resource yield gap can be further decomposed into an **allocative yield gap** (Ye – Y_{TEx}) and into an **economic yield gap** (Yf – Ye). Finally, the **technology yield gap** is defined as the difference between Yw and Y_{HF} (Silva et al., 2017) or between Yw and Yf (van Dijk et al., 2017), which can be caused by resource yield gaps of specific inputs and/or the use of technologies in farmers' fields where Yw is not achieved. The framework of Silva et al. (2017) was applied for wheat in Arsi (in collaboration with dr Frédéric Baudron), where the efficiency, resource and technology yield gaps are 17, 5 and 52% of Yw, respectively.

Throughout the analysis focus will be given to important pillars of wheat agronomy research currently undertaken by CIMMYT. Where possible, these include 1) effectiveness of improved wheat varieties against diseases as means to narrow yield gaps; 2) implications of the crop establishment method for actual yields and labour use efficiency and/or 3) interactions between fertiliser use and weed control in relation to amounts and timing.

Another important pillar of this analysis is the up-scaling of the results using technology extrapolation domains (TEDs; Edreira et al., 2018). We envisage the use of TEDs as a robust framework for biophysical upscaling of the yields gaps and their drivers but how exactly these will be used will be elaborated at a later stage in the project.

4) Inventory of datasets available

Two different types of datasets were compiled so far: 1) spatially explicit water-limited and actual yields for wheat in Ethiopia during the period 2001 – 2012 from GYGA and 2) household surveys from diverse sources containing detailed information on crop yield, management practices and socio-economic conditions for a large of number of farms in the main wheat producing areas of Ethiopia.

- **Global Yield Gap Atlas**

GYGA is an important source of biophysical and agronomy data for yield gap analysis. Background information on the simulation of water-limited yields for cereals (including wheat) in Ethiopia can be found in http://www.yieldgap.org/en/web/guest/ethiopia. In short, water-limited yields (Yw) for a modern/improved wheat variety were simulated with the crop model WOFOST across the main wheat producing regions in Ethiopia (a total of 10 weather stations). In addition to Yw, estimates of the climatic potential yield (Yp) are also available for the same weather stations, which makes it possible to quantify the yield gap due to sub-optimal water supply (Yp – Yw). Currently, Yp and Yw are available per weather station / climate zone for the period 1998 – 2011. These data will be expanded for the more recent years (2012 – 2017) in collaboration with dr Kindie Tesfaye, who is helping us accessing the required weather data.

- **Household surveys**

Household surveys containing wheat yield, and associated management practices, for a large number of farms in Ethiopia, are available from at least three different sources. These include the Living Standards Measurement Survey (LSMS) collected by the World Bank; Wheat Adoption and Impact Surveys (WAIS) collected by CIMMYT and; detailed surveys on labour use also collected by CIMMYT within the FACASI project. The sample size and years of the available data are provided in Table 2.

Table 3. Datasets available.

Household survey	Sites covered	Years survey	No. of household	No. of wheat fields	Focus of the survey	Project activity
LSMS (World Bank)	National	2013 2015	282 389	291 455	Crop cut surveys conducted as part of the LSMS	Activity 1: National Yg analysis
WAIS (CIMMYT)	National	2011 2014	2069 1921	2096 2655	Documenting wheat variety adoption dynamics	Activity 2: Regional Yg analysis
FACASI (CIMMYT)	Regional (Arsi)	2012	100	97	Mechanization and labour use at crop and farm level	Activity 2: Regional Yg analysis

Background information on the methodology and data available in the **LSMS survey** can be found in http://microdata.worldbank.org/index.php/catalog/2783. The LSMS is a nationally representative panel dataset, which was conducted in three waves: 2011, 2013 and 2015. Crop cut yields are available for all three waves but preliminary analysis of the data indicated significantly higher wheat yields in 2011. This appears to be the result of a change in survey protocol from crop cuts on a 4m2 plot in 2011 to a 16m2 plot in 2013 and 2015. As such, our analysis will focus on the second and third waves only. In total, yield data based on crop cuts is available for 282 and 389 farms and for 291 and 455 wheat fields in 2013 and 2015, respectively. These provide a reliable basis for the national yield gap analysis.

The **WAIS survey** was conducted for the purpose of tracking varietal change and of assessing the impact of genetic improvement for wheat in Ethiopia. Two empirical studies were previously conducted using these data: Shiferaw et al. (2014) investigated the drivers of adoption of improved wheat varieties in Ethiopia while Abro et al. (2017) studied the impact of wheat varieties with differing resistance against stripe rust on wheat yields. The WAIS survey is a panel of households and was conducted in two rounds covering the growing seasons of 2011 and 2014. As described by Abro et al. (2017), the sampling frame comprised the selection of 148 major wheat growing districts of Ethiopia, followed by a random selection of farmers' associations (communities) within these districts and by a random selection of 15 to 18 households within each farmers' association. This resulted in a sample of representative farmers in the major wheat-growing areas of Ethiopia. The survey includes a wide range of farm and farmer characteristics as well as detailed information on the types and quantities of inputs used and crop yields obtained in all fields of each farm. The large sample size and national coverage makes this survey highly suitable for the purpose of this project.

The **FACASI survey** was conducted in 2012 with the purpose of mapping the potential demand for mechanisation in Eastern and Southern Africa. A total of 200 farms were interviewed in Southern Ethiopia across two contrasting farming systems: 100 farms in the wheat-based farming system around Asella and 100 farms in the maize-based farming system around Hawassa. The survey was conducted using a semi-structured questionnaire requesting detailed information on labour use at crop and farm level. Household selection was done using a systematic procedure based on transect routes across each village in which every fourth household on alternate sides of the track was sampled. These data were previously used by Silva et al. (under review) to understand the role of labour in explaining yield gaps in the aforementioned cereal-based farming systems in Southern Ethiopia. The detailed information available in this survey on labour use for wheat and other crops makes it an interesting complement to the WIAS survey, which lacks such detail, for contextualizing yield gaps at the farm level.

5) Methodological and data concerns

So far, we have not been able to identify a robust set of on-farm / on-station wheat trials to benchmark farm data and validate Yw. Differently from maize, literature on indigenous nutrients for wheat is scant in Ethiopia and it is unclear what is the yield range when no nutrients / N are applied. Similarly, we envisage strong differences between direct-seeded and broadcasted crops but literature and data are equally scant for wheat in Ethiopia.

The quality of household survey data is varied and surveys are prone to error. Recall error, non-standard units and non-response have been well documented in the literature (Carletto et al., 2015). In addition, many datasets suffer from discernible systematic errors due to data entry or misunderstandings on the part of survey enumerators with respect to quantitative variables. Recent research has suggested that these errors can distort analysis if not identified carefully. This is the case in the Ethiopia LSMS survey where fertilizer variables are misreported by specific enumerators by a factor of ten or one hundred and household land holdings reflect a distinct, and suspicious, lack of consistency across waves. Where necessary these data quality issues will be highlighted and the decision to use, or not use, a particular variable will be justified. This will have implications regarding the (econometric) methods which can be used to analyse these data and it is unclear to which extent we will be able to follow the frameworks of Silva et al. (2017) and van Dijk et al. (2017). As for the WAIS data, one of its limitation is the lack of detailed information on the timing of the inputs applied as well as on labour use, which will limit the identification of the drivers behind the efficiency yield gap and the development of labour calendars at farm level.

6) References

Abro, Z.A.; Jaleta, M.; Qaim, M. (2017). Yield effects of rust-resistant wheat varieties in Ethiopia. *Food security*, 1343 – 1357.

Carletto, C.; Jolliffe, D.; Banerjee, R. (2015). From tragedy to renaissance: Improving agricultural data for better policies. *The Journal of Development Studies*, 51, 133 – 148.

Edreira, J.I.R.; Cassman, K.G.; Hochman, Z.; van Ittersum, M.K.; Bussel, L.G.J.; Claessens, L.; Grassini, P. (2018) Beyond the plot: technology extrapolation domains for scaling out agronomic science. *Environmental Research Letters*, 13, 054027.

Habte, D.; Tadesse, K.; Admasu, W.; Desalegn, T.; Mekonen, A. (2014). Agronomic and economic evaluation of the N and P response of bread wheat growing in the moist and humid midhighland vertisols areas of Arsi zone, Ethiopia. African Journal of Agricultural Research, 10, 89 – 99.

Mann, M.L.; Warner, J.M. (2017). Ethiopian wheat yield and yield gap estimation: A spatially explicit small area integrated data approach. *Field Crops Research*, 201, 60 – 74.

Nyssen, J.; Govaerts, B.; Araya, T.; Cornelis, W.M.; Bauer, H.; Haile, M.; Sayre, K.; Deckers, J. (2011). The use of the marasha ard plough for conservation agriculture in Northern Ethiopia. *Agronomy for Sustainable Development*, 31, 287 – 297.

Silva, J.V.; Baudron, F.; Reidsma, P.; Giller, K.E. (under review). Is labour a major determinant of yield gaps in sub-Saharan Africa? A study for cereal-based production systems in Southern Ethiopia. *Agricultural Systems*.

Silva, J.V.; Reidsma, P; Laborte, A.G.; van Ittersum, M.K. (2017). Explaining rice yields and yield gaps in Central Luzon, Philippines: An application of stochastic frontier analysis and crop modelling. *European Journal of Agronomy*, 82 Part B, 223 – 241.

Shiferaw, B.; Kassie, M.; Jaleta, M.; Yirga, C. (2014). Adoption of improved wheat varieties and impacts on household food security in Ethiopia. *Food policy*, 44, 272 – 284.

Tanner, D.G.; Gorfu, A.; Taa, A. (1993). Fertiliser effects on sustainability in the wheat-based small-holder farming systems of southeastern Ethiopia. *Field Crops Research*, 33, 235 – 248.

van Dijk, M.; Morley, T.; Jongeneel, R.; van Ittersum, M.K.; Reidsma, P.; Ruben, R. (2017). Disentangling agronomic and economic yield gaps: An integrated framework and application. *Agricultural Systems*, 154, 90 – 99.

van Ittersum, M.K.; van Bussel, L.G.J.; Wolf, J.; Grassini, P.; van Wart, J.; Guilpart, N.; Claessens, L.; de Groot, H.; Wiebe, K.; Mason-D'Croz, D.; Yang, H.; Boogaard, H.; van Oort, P.A.J.; van Loon, M.P.; Saito, K.; Adimo, O.; Adjei-Nsiah, S.; Agali, A.; Bala, A.; Chikowo, R.; Kaizzi, K.; Kouressy, M.; Makoi, J.H.J.R.; Ouattara, K.; Tesfaye, K.; Cassman, K.G. (2016). Can sub-Saharan Africa feed itself? *Proceedings of the National Academy of Sciences*, 113, 14964 – 14969.