Field Crops Research 197 (2016) 97–106

Contents lists available at ScienceDirect

Field Crops Research

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

Data rich yield gap analysis of wheat in Australia

Zvi Hochman[a,*], David Gobbett[b], Heidi Horan[a], Javier Navarro Garcia[c]

[a] CSIRO Agriculture and Food, 306 Carmody Road, St Lucia, Queensland, 4076 Australia
[b] CSIRO Agriculture and Food, PMB 2, Glen Osmond, South Australia, 5064 Australia
[c] CSIRO Land and Water, PO Box 2583, Brisbane, Queensland, 4001 Australia

ARTICLE INFO

Article history:
Received 21 March 2016
Received in revised form 10 August 2016
Accepted 11 August 2016
Available online 17 August 2016

Keywords:
Yield gap
Global Yield Gap and Water Productivity Atlas (GYGA)
Agricultural Production System sIMulator (APSIM)
Water-limited yield (Yw)
Spatial framework

ABSTRACT

Closing the gap between yields currently achieved on farms and those that can be achieved with best practice and current technology in a given environment is a key strategy for increasing grain production on existing cropland. In this study, the yield gap of rainfed wheat in Australia was derived by a data rich (DR) analysis method made possible by the high density of data available in the Australian grain zone. At the national scale, the analysis revealed that Australia's 15 year (1996–2010) average actual wheat yield (Ya) was 1.7 t/ha, while the simulated water-limited yield potential (Yw) was 3.5 t/ha; thus the yield gap (Yg = Yw-Ya) was 1.8 t/ha; and the relative yield ($Y_{\%}$ = 100 x Ya/Yw) was 49%. The fifteen year average Ya and Yw values for each of the 245 statistical local areas (SLAs) in Australia's wheat zone were strongly and positively correlated (Ya = 0.456 x Yw + 0.15; R^2 = 0.65; RMSE = 0.291; $p < 0.001$) such that the yield gap (Yw-Ya) tended to be larger in SLAs with higher Yw values. Mean Y% values vary among SLAs from a low of 34% to a high of 69% (standard deviation = 7.7%). The extremely tight correlation, at the national scale (Ya = 0.488 x Yw + 0.021; R^2 = 0.95; RMSE = 0.087; $p < 0.001$), between the methodologically independent estimates of Yw and Ya values points to the consistency of both estimates. It also indicates the potential for using updated Yw estimates to help derive more reliable national crop yield forecasts in real time. The results of this analysis were compared with those of two alternative analyses of Australia's wheat yield gaps using (1) EarthStat, a statistically based global yield gap analysis methodology and (2) the Global Yield Gap Atlas (GYGA) designed to achieve locally credible assessments of yield gaps for developing regions such as sub-Saharan Africa (SSA) and South Asia (SA) with the limited data typically available in such countries. At the national level all three methods gave similar results. However, at the sub-national level (based on agroecological zones) differences emerged between the water-limited yield potential values of EarthStat on the one hand and those of the DR and GYGA methods on the other. Given the high spatial environmental variability within the Australian wheat zone the remarkable level of agreement between results from these two methodologies provides evidence of the robustness of the GYGA protocols. The advantage of the DR method is in its higher resolution and the subsequent local relevance of results.

1. Introduction

Wheat is an important crop in Australia. It grows on 55% of the total cropland and averaged 12.6 million ha between 1998 and 2011 (ABARES, 2012). Australia has a significant role to play in global food security because, as a significant wheat exporter (Australia contributed 12% of global wheat exports in 2005–2012), it can help compensate for seasonal fluctuations in other global regions.

Future global food security depends on producing enough nutritious food for a world population expected to peak at over 9 billion by 2050. Achieving this in an environmentally sustainable manner clearly depends on realising the highest possible yields on existing farm land to protect our carbon-rich and bio-diverse forests, wetlands, and grasslands. One promising strategy to increase grain production is to close the gap between yields currently achieved on farms and those that can be achieved by using the best adapted crop varieties with the best current crop and land management practices for a given environment (van Ittersum et al., 2013). The first step towards closing the yield gap is to quantify its size and distribution. Other pathways to future food security, including increasing potential yields (Fischer et al., 2014), reducing pressure on food demand and avoiding losses in current or future production potential (Keating et al., 2014) must also be addressed.

* Corresponding author.
E-mail address: zvi.hochman@csiro.au (Z. Hochman).

http://dx.doi.org/10.1016/j.fcr.2016.08.017

The case for a global estimate of yield gaps, based on locally relevant and credible assessment of the yield gaps of the world's major crops, was made by van Ittersum et al. (2013) who outlined the methodology for developing the Global Yield Gap Atlas (GYGA). This "bottom-up" methodology is based on the 'GYGA protocol' for generating national yield gap estimates using a climate zonation scheme (Van Wart et al., 2013) to group agro-climatically similar areas for analysis. The protocol outlines guidelines for the selection of key climate zones (CZs), and then for the selection of reference weather stations (RWS) to represent these zones. Water-limited yield potential (Yw; yield that can be achieved under current best practice with well adapted commercial varieties and known technologies) is simulated using a locally validated crop model parameterized for local agronomic and soil information for a 100 km buffer zone around those RWS. Actual yields (Ya) are sourced from reliable survey data and both Yw and Ya are first estimated at RWS scale and then scaled up from RWS to CZs and then to national scale using cropland area-weighted averages (van Bussel et al., 2015). Tiered data selection methodologies (Grassini et al., 2015) ensure the highest quality data and locally relevant expertise are used first, but where necessary, progressively lower quality data may be incorporated into the analysis. Readers are referred to these publications for a more detailed description and justification of these protocols. At the time of writing the yields of 9 field crops in 32 countries (between one and five crops per country) had been mapped using these protocols and work was underway in another 12 countries (www.yieldgap.org; last accessed 10-08-2016).

Key aspects of the GYGA methodology such as the climate zonation scheme, minimum area covered and the minimum number of weather stations required to achieve consistent national estimates of yield gaps have been validated (Van Wart et al., 2013; van Bussel et al., 2015; Grassini et al., 2015). However, some elements such as the assumed size of the zone of influence of the RWS and the importance of adequate representation of soils have not been formally evaluated. Importantly, for countries such as Australia with its highly variable climate, van Bussel et al. (2015) concluded that for "semi-arid areas with large variability in rainfall . . . scaled-up water-limited yield gap estimates can be prone to errors, especially if only a limited number of weather stations is available per climate zone". Hence, there is still a need to validate the GYGA methodology in its entirety for a whole country, and especially for a country like Australia with its semi-arid cropping areas and large spatial and temporal variability in rainfall.

The GYGA protocols provide estimates of yield gaps at three scales: 1. 100 km buffer zones around selected RWS; 2. a number of CZs at a sub national scale; 3. National scale. This is valuable information for prioritising research and development effort, especially for international aid donors, but is less than ideal for directing efforts at a sub-national level. One shortcoming is that because 50% coverage of the harvested area is considered sufficient, it provides no specific information about the rest of the cropped areas which may well be distributed in many smaller CZs. There is also no way of determining whether yield gaps are uniform within the CZs or the 100 km RWS buffer zones. Further, there is no reason to assume that they are uniform since yield gaps are just as likely to be determined by socio-economic factors as they are by the bio-physical factors that define these buffer zones.

For these reasons it is worth investigating whether a more detailed and spatially more complete methodology can be developed to provide a fuller and more detailed picture of yield gaps and their distribution in countries with access to more detailed data on climate and soils. While such an analysis may not apply to many developing countries, it could well apply to many countries with greater data resources than the minimum required for the GYGA protocols.

In this study we adapted a data rich (DR) method developed initially for estimating the wheat yield gap in the Victoria Mallee region of Australia (Hochman et al., 2012) to the whole Australian wheat zone. We used the results of this analysis to investigate the relationship between yield gaps and climate variability and to conduct a comparison with the results of a recently completed analysis of dryland wheat yield gaps in Australia that deployed the GYGA protocols over the same time period (Gobbett et al., 2016).

An alternative methodology to both the DR and GYGA protocols is the "top-down" method (e.g. Monfreda et al., 2008; Mueller et al., 2012) which is often deployed at the global scale. To compare and contrast the results of the DR analysis with these two alternative methodologies for determining yield gaps at multiple scales we also compared the results obtained from the DR method to the results previously reported from such a top down methodology. For this purpose we used data provided for Australian wheat by the global EarthStat maps (Monfreda et al., 2008; Mueller et al., 2012).

2. Methods

This analysis is based on adapting, for the whole Australian wheat zone, a yield gap analysis method developed initially to estimate the wheat yield gap in the Wimmera region of Victoria, Australia (Hochman et al., 2012). This method exploits a large set of available soil, climate and crop data in the expectation that more detail will lead to greater accuracy and local relevance. The first step in calculating the yield gap was to determine where the crop is grown. For wheat in Australia, the National Land Use of Australia version 4 (2005-6) (ABARE-BRS, 2010) data set provides data for mapping a 'cereals' land use class at approximately 1 km pixel size. The nature of crop rotations in Australia is such that land used for other cereals in a given year will be used for wheat in other years.

Next we mapped actual annual wheat grain yields (t/ha) obtained by farmers between 1996 and 2010 (Ya) onto the land use map. For this we used the national agricultural data collated by the Australian Bureau of Statistics (ABS) at the level of statistical division (SD) annually, and at the finer scale of statistical local area (SLA) every five years when a census is carried out (Walcott et al., 2013). SDs are relatively uniform regions with boundaries determined from socioeconomic criteria. SLAs are subdivisions of SDs (see SLA and SD boundaries in Supplementry Appendix A, Fig. 1). SLAs are the smallest administrative unit at which national crop yield data are available. However, since 1996 data at SLA resolution has only been published at five-yearly intervals in census years (1996, 2001 and 2006). For intervening years, only SD level data were available. Data on annual crop harvested area and average yields for the years 1996–2010 were sourced from ABS (2012). To derive SLA level estimates for each year, linear regressions were fitted to yield (t) and crop area (ha) data for each of the SLAs that grew more than 1000 ha of wheat from the 17 past census years from 1982 to 2010 (method described in detail in Supplementry Appendix D of Gobbett et al., 2016). SLAs for which regression between SLA yield and SD yield was not significant at $p < 0.1$ were omitted from the analysis.

To determine Yw we deployed the APSIM (Version 7.4) wheat model (Keating et al., 2003; Holzworth et al., 2014) which is well validated for wheat in Australia (e.g. Asseng et al., 1998; Wang et al., 2003; Hochman et al., 2007; Brown et al., 2014) to simulate Yw, using 30 years of weather data from 3,912 SILO Patched Point data weather stations (Jeffrey et al., 2001) covering the grain zone at a median distance apart of 17 km. We assumed a 20 km radius as the nominal zone of relevance of each weather station and chose up to three dominant soil types per weather station using the ASRIS soil map (Johnston et al., 2003) to determine the proportional areas of the most relevant soil types covering the cereal land use area in each 20 km radius zone. Typical soil profiles for each soil type were deter-