LETTER • OPEN ACCESS

Simplicity on the far side of complexity: optimizing nitrogen for wheat in increasingly variable rainfall environments

To cite this article: Z Hochman and F Waldner 2020 *Environ. Res. Lett.* **15** 114060

View the article online for updates and enhancements.

ENVIRONMENTAL RESEARCH
LETTERS

CORRIGENDUM

Corrigendum: Simplicity on the far side of complexity: optimizing nitrogen for wheat in increasingly variable rainfall environments (2020 *Environ. Res. Lett.* **15** 114060)
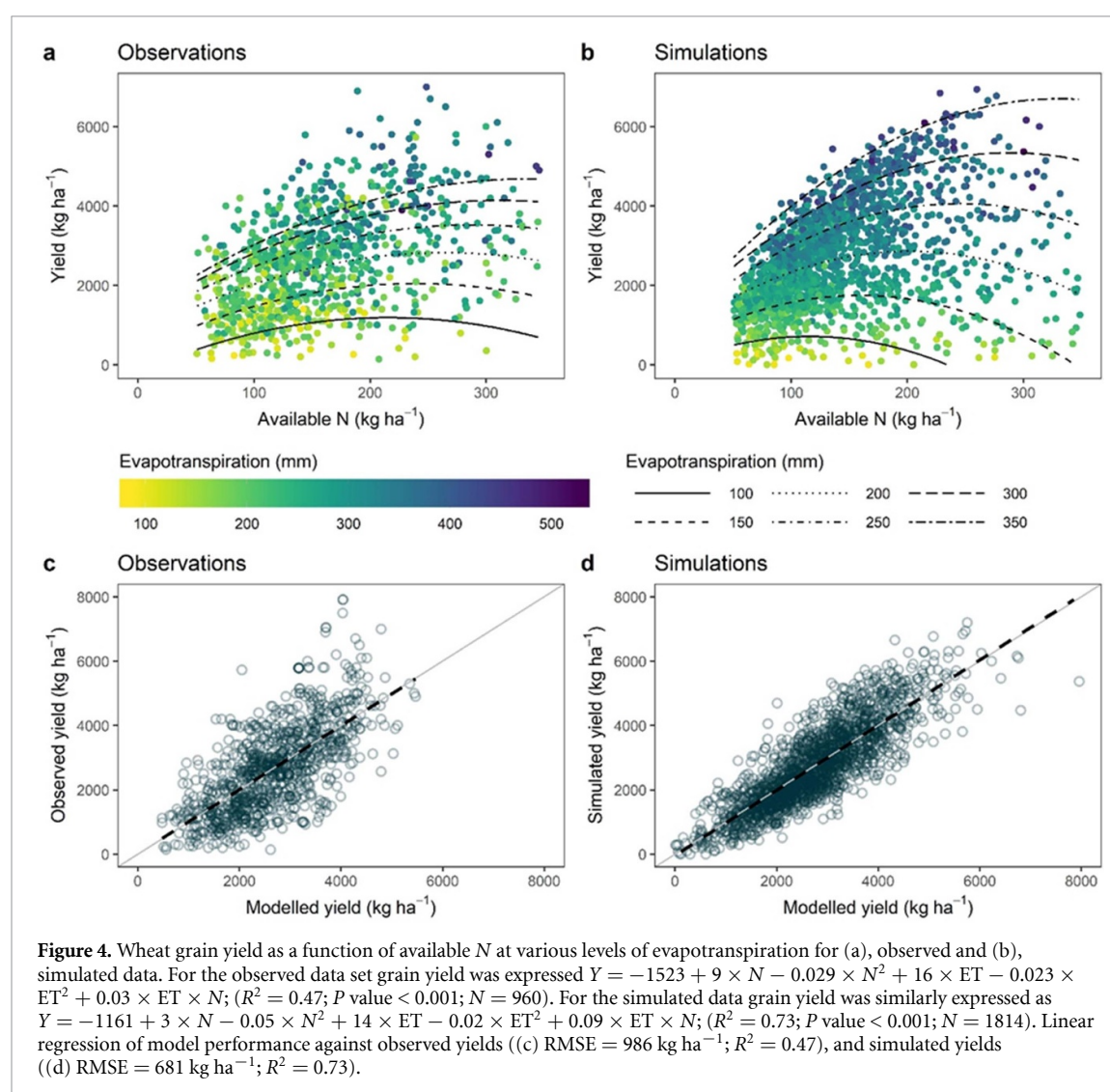
Z Hochman[*] and F Waldner

CSIRO Agriculture and Food, Queensland Bioscience Precinct, 306 Carmody Road, St. Lucia, QLD 4067, Australia
[*] Author to whom any correspondence should be addressed.

E-mail: zvi.hochman@csiro.au

We made errors in transcribing the formulae provided in the titles of figures 4 and 5. These errors were made after the statistical analysis was complete and the figures were generated. Hence, the originally published figures and statistical values (R^2, RMSE and P values) are correct.

Figure 4. Wheat grain yield as a function of available N at various levels of evapotranspiration for (a), observed and (b), simulated data. For the observed data set grain yield was expressed $Y = -1523 + 9 \times N - 0.029 \times N^2 + 16 \times ET - 0.023 \times ET^2 + 0.03 \times ET \times N$; ($R^2 = 0.47$; P value < 0.001; $N = 960$). For the simulated data grain yield was similarly expressed as $Y = -1161 + 3 \times N - 0.05 \times N^2 + 14 \times ET - 0.02 \times ET^2 + 0.09 \times ET \times N$; ($R^2 = 0.73$; P value < 0.001; $N = 1814$). Linear regression of model performance against observed yields ((c) RMSE = 986 kg ha^{-1}; $R^2 = 0.47$), and simulated yields ((d) RMSE = 681 kg ha^{-1}; $R^2 = 0.73$).

Figure 5. Wheat grain yield as a boundary function of available *N* at various levels of evapotranspiration for (a), observed and (b), simulated data. For the observed data set grain yield was expressed as $Y = -590 + 18 \times N - 0.04 \times N^2 + 14 \times \mathrm{ET} - 0.02 \times \mathrm{ET}^2 + 0.03 \times \mathrm{ET} \times N$. For the simulated data grain yield was similarly expressed as $Y = -785 + 21 \times N - 0.10 \times N^2 + 10 \times \mathrm{ET} - 0.03 \times \mathrm{ET}^2 + 0.11 \times \mathrm{ET} \times N$. A graphic illustration of the rate of *N* required to achieve water-limited yield (100%) as well as 90% and 80% of water-limited yield for (c), the observed data set and for (d), the simulated data set.

ORCID iDs

Z Hochman ⓘ https://orcid.org/0000-0002-6217-5231
F Waldner ⓘ https://orcid.org/0000-0002-5599-7456

Environmental Research Letters

LETTER

Simplicity on the far side of complexity: optimizing nitrogen for wheat in increasingly variable rainfall environments

Z Hochman and F Waldner

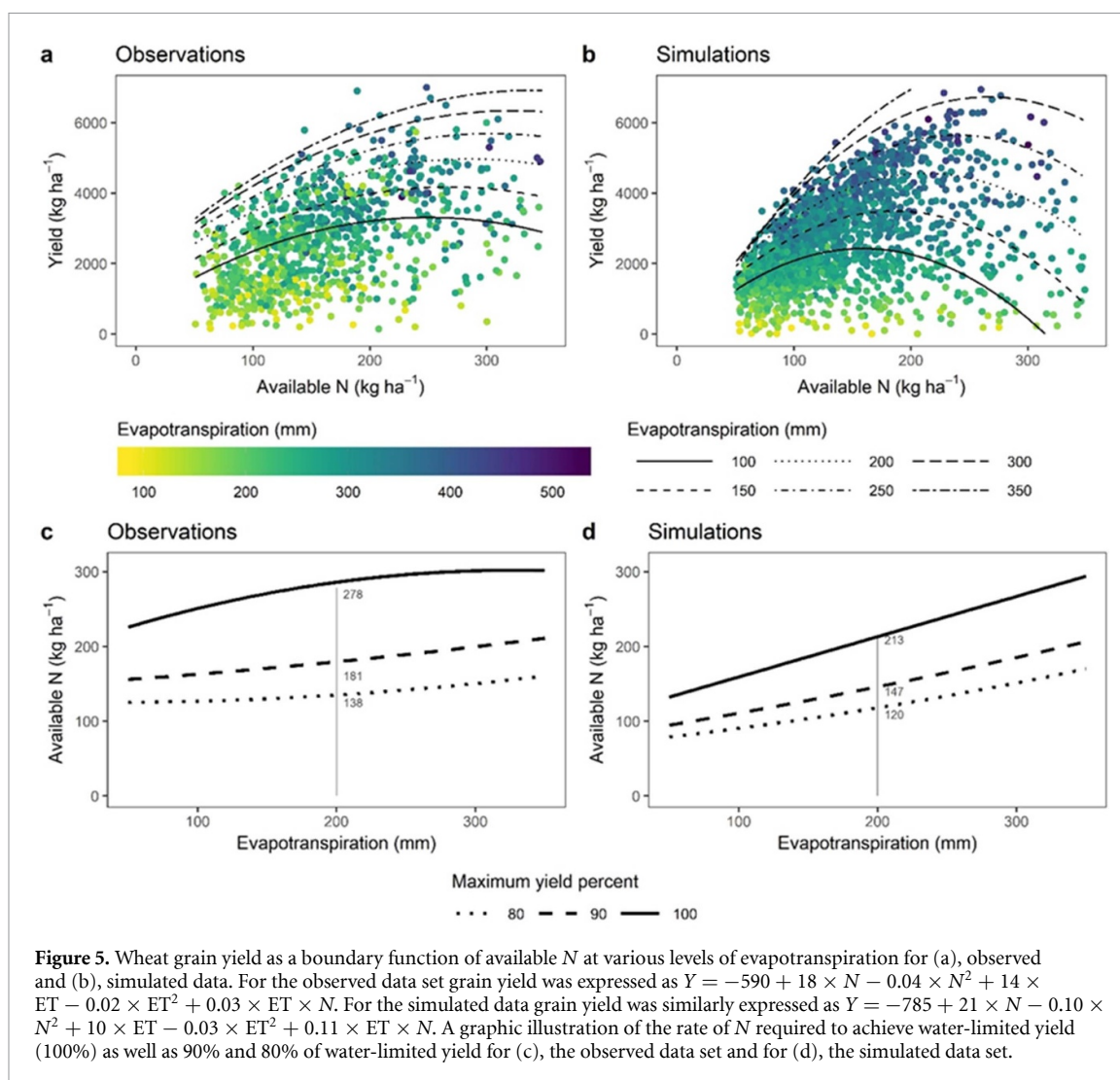CSIRO Agriculture and Food, Queensland Bioscience Precinct, 306 Carmody Road, St Lucia, QLD 4067, Australia

E-mail: zvi.hochman@csiro.au

Abstract

The increasingly chaotic nature of rainfall in semi-arid climates challenges crop growers to balance nitrogen fertiliser inputs for both food security and environmental imperatives. Too little nitrogen restricts yields and runs down soil organic carbon, while too much nitrogen is economically wasteful and environmentally harmful. The degree to which crop-water and crop-nitrogen processes combine to drive yields of rainfed wheat crops is not well understood or quantified. Here we investigate two comprehensive Australia-wide data sets, one from commercial wheat growers' fields and the other from systematic simulation of 50 sites by 15 years using a comprehensive mechanistic cropping system model. From these data, we derived a simple model combining water use with available nitrogen and their interaction. The model accounted for 73% of the variation in the simulated yield data and 46% of the variation in the growers' yield data. We demonstrate how the simple model developed here can be deployed as a tool to aid growers' in-crop nitrogen application decisions.

1. Introduction

The rise in the world's population and per capita food consumption challenges agricultural science to rapidly increase food production. The most limiting factors affecting cereal production in rainfed environments are the amount of water (expressed as seasonal evapotranspiration; ET) and nitrogen (N) available to these crops (Fischer 1981, Smith and Harris 1981, Sinclair and Horie 1989, Grindlay 1997, Passioura 2002, Sadras *et al* 2016). Knowledge of the factors governing supply and demand of N is essential to predict the needs of crops under a wide range of field situations so that growers can be given more reliable fertilizer recommendations. This is important as risks to the environment can arise from the over-application of N fertilizers resulting in nitrate leaching, nitrous oxide (N_2O and NO_x) emissions and soil acidification (Addiscott *et al* 1991, Galloway *et al* 2003, Vitousek *et al*, Sutton *et al* 2013, Carberry *et al* 2013) where less than half of the reactive N added to croplands is converted into harvested products (Lassaletta *et al* 2014). Yet, under-fertilization causes soil degradation, low yields and poverty (Sutton *et al* 2013, Tittonell and Giller 2013)

in smallholder agriculture and is the most important single factor explaining large yield gaps in Australia (Hochman and Horan 2018). These considerations highlight the need for quantitative models to reliably predict N-yield interactions for a wide variety of crop-climate situations.

Studies of ET and N effects on wheat yields have been conducted with regularity (Sharma *et al* 1990, Latiri-Souki *et al* 1998, Karam *et al* 2009). However, these have usually produced N response functions at specific levels of ET supply, or ET response functions at specific levels of N, for a single or a few seasons and/or locations with limited applicability to other environments. The current generation of crop simulation models (e.g. van Keulen and Seligman 1987, Brisson *et al* 1998, Jones *et al* 2003, Holzworth *et al* 2014) can simulate crop growth and yield in response to limited ET and available N and may be used to simulate yield under a wide range of N and water supply conditions. APSIM and similar process-based models capture key aspects of the interactions between water and N. Available mineral N in soil is updated daily by the soil N module, which simulates soil N processes including mineralisation, immobilisation, nitrification, denitrification, movement in soil

and leaching (Probert *et al* 1998). Suboptimal water and N in soils lead to an imbalance between crop demand and the soil's ability to meet this demand. The shortfall is expressed as stress functions and these are updated daily. These stress functions are then used to reduce the environmentally determined rate of leaf area expansion, root front exploration, biomass growth and other processes. Where water limits growth, reduced biomass reduces demand for N. Similarly, N stress reduces leaf area and hence crop evaporative demand. Reduced leaf area also affects the partitioning of daily water use between transpiration and soil evaporation. Water and N interactions are also captured in the water-driven uptake of N by mass flow and diffusion and in the water-driven fate of soil N processes such as mineralisation and leaching (Sadras *et al* 2016).

The APSIM wheat model has been evaluated in numerous publications, especially its grain yield response to a wide range of N fertiliser applications and water supply conditions (11 relevant reports were cited by Keating *et al* 2003). In the most comprehensive such study, the APSIM Nwheat model was incorporated into the DSSAT modelling platform and was evaluated using more than 1000 observations from field experiments of 65 treatments, which included a wide range of N fertilizer applications in diverse climatic regions that represented the main wheat-growing areas of the world. The study found that model reproduced the observed grain yields well with an overall root mean square deviation (RMSD) of 0.89 t ha^{-1} (13%). N applications, water supply, and planting dates had large effects on observed biomass and grain yields, and the authors found that the model reproduced these crop responses well (Kassie *et al* 2016).

However, these models require a significant amount of parameterization, especially of soil water holding characteristics, soil organic mater, plant available soil water and mineral N status of the whole profile to the depth of maximum root penetration. This requirement puts the practical application of these models beyond the reach of most grain growers and their advisors.

At the opposite end of the complexity scale, there are rules of thumb, whereby target yields are multiplied by a constant value to determine a crop's N requirement. For example, in Yield Prophet Lite a target yield is determined using a simple water use efficiency formula (Sadras and Angus 2006; see equation (1) section 2.3) and this target yield (in t ha^{-1}) is then multiplied by 40 to determine the rate of available N (kgN ha^{-1}) required to achieve the target yield (www.yieldprophet.com.au/yplite/). There do not seem to be any intermediate tools to assist farmers decisions on matching N fertiliser application to seasonal conditions.

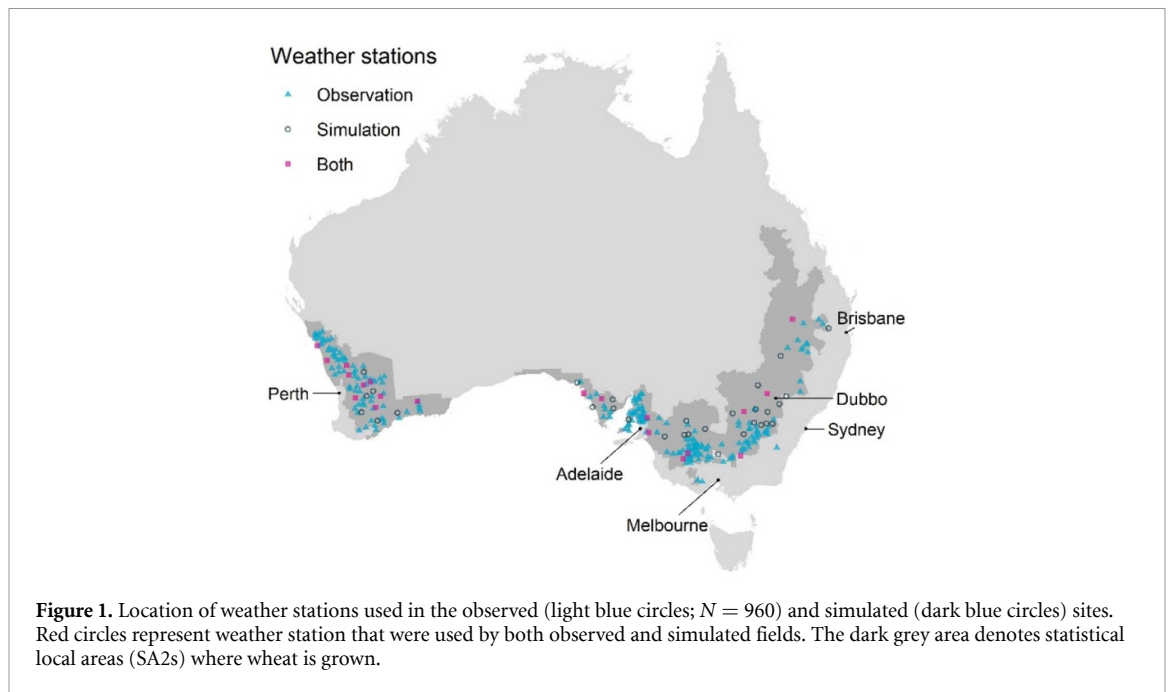To empower farmers to make informed decisions about N fertilization we require a simple, yet not simplistic, model of how crop yields respond to N fertilizer in a variable rainfall environment. Remarkably, we were unable to find any published attempts to use these comprehensive crop models to generalize the relationship between the limitations imposed by various combinations of ET by N availability and yield for a wide range of environments. Such a situation is reminiscent of a quote attributed to Oliver Wendell Holmes

> *'I would not give a fig for simplicity this side of complexity, but I would give my life for simplicity the other side of complexity.'*

The aim of this research was to establish the optimal rate of N fertilizer to be applied to rainfed crops across a wide range of available soil N and seasonal ET conditions by an investigation of two independent data sets from Australia's cropping zone. The first analysis involves a comprehensive data set of wheat growers' commercial fields, distributed throughout the Australian grain zone over 11 seasons. The second data set consists of simulated yields from 50 sites over 15 years. For each data set, we derived an equation of the combined effects of ET and Available N and their interaction (ET × Available N) on wheat grain yield. We also re-purposed this combined equation to calculate the Available N values required to achieve a crops' water-limited yield (Yw), as well as 90% and 80% of Yw for any amount of ET, thus proposing a new decision tool to guide growers on crop N requirements in variable rainfall environments. These models should have wide relevance to dryland cropping globally as Australia's cropping zone shares agroclimatc zones with cropping regions in southern Africa, South America, southeastern USA, Mexico, Middle and Near Eastern as well as southern European countries (van Wart *et al* 2013).

2. Methods

The data analyzed included observed and simulated data sets. Both data sets were geographically well distributed throughout the Australian cropping zone (figure 1) across 11 agroclimatic zones (van Wart *et al* 2013) which together cover 90% of the winter cereals cropping areas (ABARE-BRS 2010). In an earlier study (Hochman *et al* 2017) we showed that the number of sites per zone correlates strongly with the proportion of national winter cereals cropping area within the agro-ecological zones ($R^2 = 0.92$) showing that the 50 sites capture a representative range of cropping environments across the Australian continent. The major soil types (Isbell 1996) on which winter grain crops are grown (Hochman *et al* 2016) were also well represented.

Figure 1. Location of weather stations used in the observed (light blue circles; $N = 960$) and simulated (dark blue circles) sites. Red circles represent weather station that were used by both observed and simulated fields. The dark grey area denotes statistical local areas (SA2s) where wheat is grown.

2.1. Observed data

The observed data were sourced from the Yield Prophet® (Hochman *et al* 2009b) data base. It contains grower supplied field level data on grain yield (kg ha^{-1}), soil characterization data (including crop lower limit, drained upper limit, bulk density and soil organic carbon) pre-sowing soil mineral N, pre-sowing soil water content, weather data recorded on farm or from the nearest weather station as well as management information including N fertilizer input, time of sowing, crop type and crop variety. The data analyzed include 960 fields from the years 2005–2015. Recorded yields averaged 2667 kg ha^{-1} with a range of 140–7910 kg ha^{-1}.

ET was calculated as the sum of in-crop rainfall plus the difference between soil water measured pre-sowing and soil water at harvest (simulated in Yield Prophet®). The mean ET across all sites and years was 229 mm with a range from 80 to 526 mm. Available N was calculated as the sum of mineral N measured pre-sowing and fertilizer N applied to the crop. This simplified calculation of available N does not take into account in-crop N cycling processes such as mineralisation, immobilisation, leaching and denitrification. While these processes contribute to available N, significant losses of N from Australian cropping systems are infrequent and at low intensities (Carberry *et al* 2013) and they cannot be measured by growers or their consultants and are therefore excluded. The mean Available N measured across all sites and years was 160 kgN ha^{-1} with a range of 25–346 kgN ha^{-1}.

2.2. Simulated data

The simulation data used in this analysis are a subset of the simulations produced to investigate the causes of wheat yield gaps in Australia (Hochman and Horan 2018). The Agricultural Production Systems Simulator (APSIM v.7.8; Holzworth *et al* 2014) was used to model water and N-limited wheat grain yield over the 2001–2015 growing seasons using the climate files of 50 sites and soil characterization data representative of the dominant soil type in winter cropping land use within a 20 km radius of the weather station. This spread of sites and years was chosen to ensure that the range of seasonal conditions encountered over the Australian cropping zone is more than adequately captured.

In this research we used the same APSIM management rules as those used to simulate water-limited yields except that annual fertilizer N applications were limited to 22.5, 30, 45, and 90 kgN ha^{-1} in various sites and treatments in order to create a highly diverse set of ET and N-limited situations.

2.2.1. Water-limited yield (Yw)

Yw represents the yield that can be achieved by rainfed crops when grown with best management practices under current technology, with nutrients non-limiting and biotic stress effectively controlled. Under conditions that can achieve Yw, crop growth rate is determined only by available water, solar radiation, temperature, atmospheric CO_2 and genetic traits that govern the length of the growing period and light interception by the crop canopy. Yw is location-specific because of the climate and soil properties that govern soil water availability based on available water storage capacity, rooting depth and soil constraints such as salinity or physical barriers to root proliferation (van Ittersum *et al* 2013). All simulations in the current research were based on Yw practices but were also potentially limited by Available N via the variable N fertilizer treatments.

2.2.2. Sowing rules

All sites north of latitude −32.244 (Dubbo, NSW) were classed as northern sites and used the northern sowing rule; all other sites used the southern sowing rule:

- Northern sowing rule: sow if rain ⩾15 mm over 3 d and PAW ⩾30 mm from 26 April–15 July.
- Southern sowing rule: sow if rain ⩾15 mm over 3 d regardless of soil moisture from 26 April–15 July.

In both cases, the crop is sown on 15 July if criteria are not met during the sowing window. Other key sowing rules include: sowing density = 150 plants m^{-2}, row spacing = 250 mm, and sowing depth = 30 mm.

2.2.3. Soil initialization and annual parameter reset rules

Because initial soil moisture is an important but unmeasured parameter at the start of the simulation period of interest, initial soil water was arbitrarily set to 10% PAWC 15 years before the start date of the simulation in order to allow soil water to find its correct level at the start of the simulation period (the first 15 years of data are then discarded). Soil organic carbon is initiated as per soil profile data. Initial soil NO$_3$ is set to 25 kgN ha^{-1} for each metre depth of soil, initial soil ammonium (NH$_4$) is set to 5 kg ha^{-1} for each metre depth of soil. Initial surface organic matter is set to 100 kg ha^{-1} with the C:N ratio set at 80. Surface organic matter; soil organic matter; soil water; soil NO$_3$ and NH$_4$ are not reset at any time.

2.2.4. N fertilizer treatments

The N fertilizer treatments are set at 22.5, 30, 45 and 90 kgN ha^{-1} and were applied to each of the 50 sites × 15 years. For each treatment, the set rate of N is added annually at sowing. While the N45 and N90 treatments were added to all sites and years, the N22.5 and N30 treatments were applied only to the 15 sites with the lowest average annual yields to reflect grower practices in the lower yielding areas. Additionally, as the soil parameters NO$_3$ and NH$_4$ are not reset at crop maturity, the effects of over or under fertilization in one season are carried over to the next.

2.3. Statistical analysis

The wheat yield response to seasonal ET and Available N was analyzed using linear and quadratic regression models, respectively. Individual models were fitted for the observed and simulated data sets. The significance of model parameters was assessed with t tests and their associated *P* values and the goodness of fit of these models was evaluated with the adjusted coefficient of determination (R^2), which corrects for the degrees of freedom.

Maximum boundary functions were historically established by drawing arbitrary lines along the upper boundary of data to manually parametrize equations in light of known biophysical principles. The equations take the general form
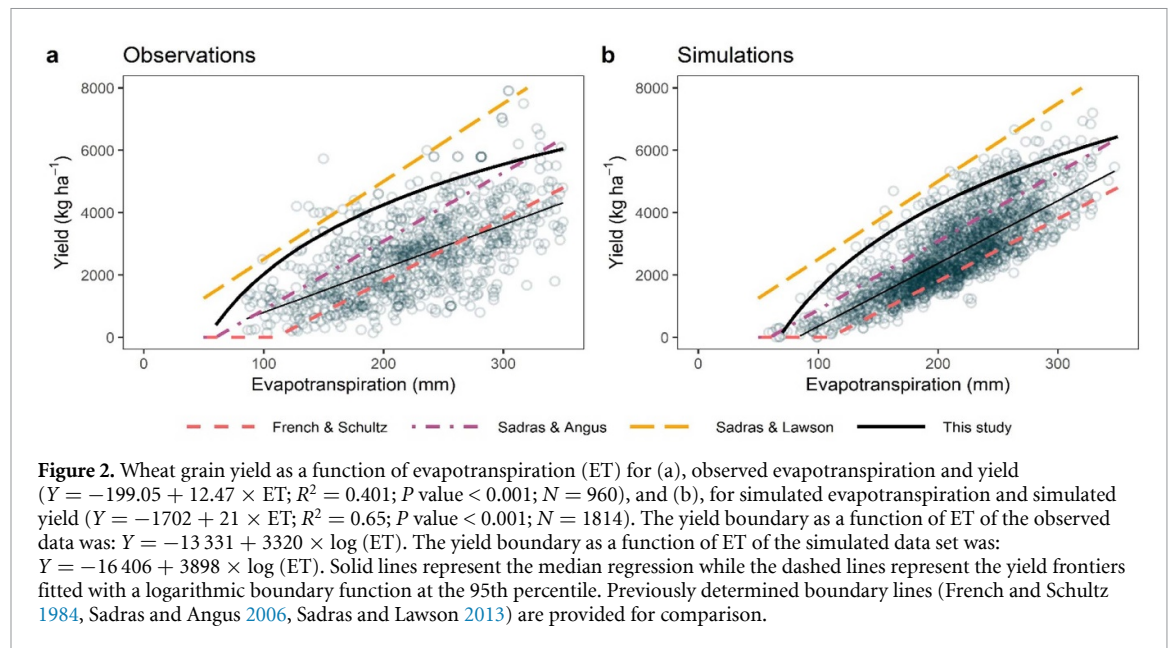
$$\text{Yield} = a + b \, (\text{Water Use}) \tag{1}$$

where the *x*–intercept (a) represents water lost to soil evaporation, and the slope (b) represents a crop's maximum transpiration efficiency (French and Schultz 1984, Sadras and Angus 2006, Hochman *et al* 2009a). Recently, more objective, i.e. data-driven, methods such as quantile regression (Cade and Noon 2003) or production frontiers (Aigner *et al* 1977) have been applied to derive the maximum bound for yield (Phillips *et al* 2006, Grassini *et al* 2009, Muller *et al* 2014, Long *et al* 2017). Here, boundary functions at the 95th percentile were fitted to identify the maximum yield values attainable for given values of ET and Available N for both the observed and simulated data sets. Logistic and quadratic functions were fitted to model the maximum yield response for ET and Available N, respectively. Boundary functions had the same form for the observed and simulated data sets. A response surface methodology was followed to identify the appropriate model form of the median yield response to ET and to Available N. Based on the analysis of variance, second-order models with an interaction term were selected because these terms contributed significantly to the model. To minimize bias, median models were calibrated to maximize Lin's concordance correlation coefficient (Lin 1989), which measures the relationship between two variables in terms of their deviation from a 1:1 ratio. The significance of each term was assessed using *P* values for the t test statistic. The prediction accuracy was assessed by computing the R^2 and the root mean square error (RMSE) between the modelled yields and the observed and simulated yields using a five-fold cross-validation approach (Hastie *et al* 2001). Multivariate yield frontier models were then developed with Available N and ET as predictor variables. The models had the same functional form as the average models (second order with an interaction term) and were fitted on the 95th percentile of the observed and simulated data sets. All analyses were performed in R (R Development Core Team 2009) with the following packages: quantreg (Koenker 2019), rsm (Lenth 2009), epiR (Stevenson *et al* 2020) and caret (Kuhn 2008).

3. Results

3.1. Yield as a function of ET

A significant linear correlation of Yield and ET was obtained for both the observed and simulated data sets (figure 2). The average yield response to total in-crop ET of the observed data was 12 kg grain mm^{-1} ET ha^{-1} with a threshold (*x*–intercept value) of 16 mm while the simulated average yield response to

Figure 2. Wheat grain yield as a function of evapotranspiration (ET) for (a), observed evapotranspiration and yield ($Y = -199.05 + 12.47 \times$ ET; $R^2 = 0.401$; *P* value < 0.001; $N = 960$), and (b), for simulated evapotranspiration and simulated yield ($Y = -1702 + 21 \times$ ET; $R^2 = 0.65$; *P* value < 0.001; $N = 1814$). The yield boundary as a function of ET of the observed data was: $Y = -13\,331 + 3320 \times \log$ (ET). The yield boundary as a function of ET of the simulated data set was: $Y = -16\,406 + 3898 \times \log$ (ET). Solid lines represent the median regression while the dashed lines represent the yield frontiers fitted with a logarithmic boundary function at the 95th percentile. Previously determined boundary lines (French and Schultz 1984, Sadras and Angus 2006, Sadras and Lawson 2013) are provided for comparison.

ET was 21 kg grain mm^{-1} ET ha^{-1} with a threshold of 81 mm. The relationship between grain yield and ET is commonly described as a boundary function where the boundary is postulated to represent the physiological limit of water use efficiency and water productivity (French and Schultz 1984, Sadras and Angus 2006, Grassini *et al* 2009, Hochman *et al* 2009a). The best boundary model for both the observed (figure 2(a)) and simulated (figure 2(b)) data sets was found to be a logarithmic function.
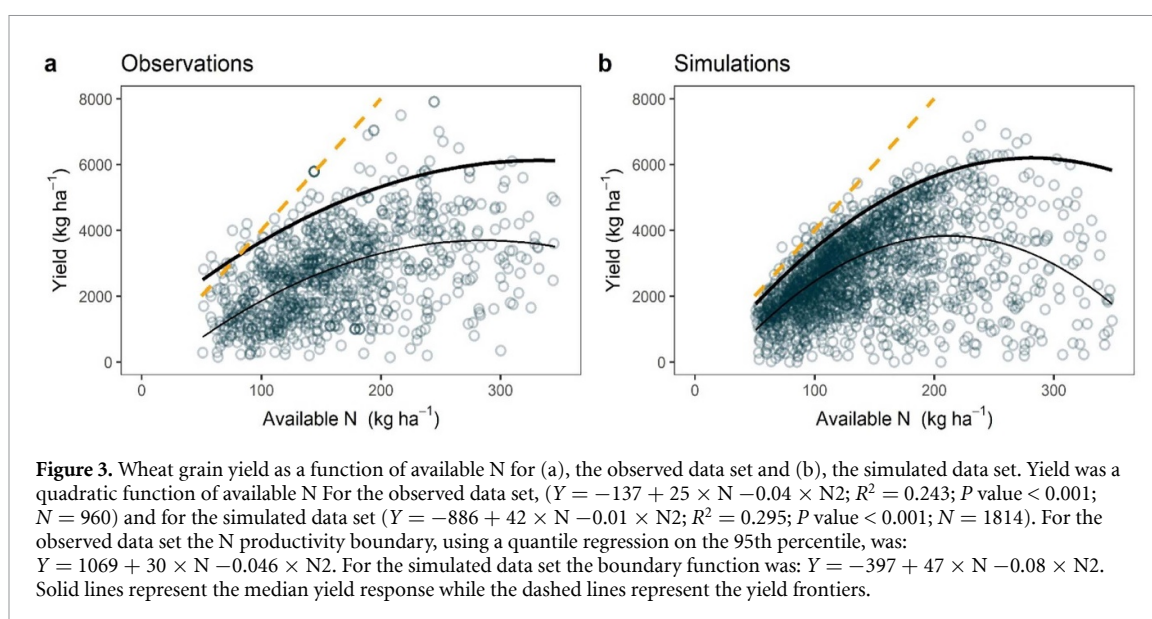
Both observed and simulated ET—Yield slope values are within the range of previously observed boundary functions for wheat crops (French and Schultz 1984, Sadras and Angus 2006, Hochman *et al* 2009a). The predictive power of the linear regressions ($R^2 = 0.4$) for the observed data set is moderate and lower than for the simulated data set ($R^2 = 0.65$). In contrast with the simple linear (Yield = (ET)) boundary models described by previous research, the boundary functions derived in this study are best described by logistic functions. This implies that maximum ET efficiency is higher than previously calculated in dry seasons and lower for higher seasonal ET values. It is possible that, for both data sets, the logistic function may reflect that factors such as Available N become more limiting as ET becomes less limiting. The considerably lower average water productivity calculated of observed data relative to the simulated data reflects the 50% yield gap calculated for wheat in Australia (Hochman *et al* 2016). The similarity between the boundary functions of the observed and simulated data sets demonstrates that, while a large yield gap exists on average, some wheat growers do achieve yields that are at or near their water-limited yield at least in some seasons (van Rees *et al* 2014, Lollato *et al* 2019).

A comparison between the boundary functions derived in this research and previously defined boundary functions for wheat crops (figure 2) demonstrates that two of the earlier defined functions (French and Schultz 1984, Sadras and Angus 2006), are too conservative when compared to more contemporary data and simulation results. The more recent study (Sadras and Lawson 2013) matches more closely the results presented here but the linear form of its boundary function is too optimistic at both low and high ET values.

3.2. Yield as a function of available N

In summary, the simulation treatments provided data from 50 sites × 15 years × various N treatment combinations or a total of 1814 yield, ET and Available N data sets. Simulated yields averaged 2725 kg ha^{-1} with a range of 200–7197 kg ha^{-1}. The mean simulated ET across all treatments, sites and years was 212 mm with a range from 59 to 390 mm. The mean simulated Available N across all treatments, sites and years was 139 kgN ha^{-1} with a range of 44–349 kgN ha^{-1}.

Grain yields in both the observed and simulated data sets were significantly correlated with Available N (figure 3). As with water productivity, the limits of N use efficiency can be described as boundary functions (Cassman *et al* 2002, Grassini *et al* 2009, Grassini and Cassman 2012, Hochman *et al* 2013). The average response of wheat grain yield to Available N (Yield = f (Available N)) was described as a quadratic function for both the observed and simulated data. The simulated average response curve peaked at 225 kgN ha^{-1} with a grain yield of 3818 kg ha^{-1}. The observed response curve was not as steep and peaked at 335 kgN ha^{-1} with a similar grain yield of 3801 kg ha^{-1}. The different average responses suggest that, especially with high N supply, observed fields

Figure 3. Wheat grain yield as a function of available N for (a), the observed data set and (b), the simulated data set. Yield was a quadratic function of available N For the observed data set, ($Y = -137 + 25 \times N - 0.04 \times N2$; $R^2 = 0.243$; P value < 0.001; $N = 960$) and for the simulated data set ($Y = -886 + 42 \times N - 0.01 \times N2$; $R^2 = 0.295$; P value < 0.001; $N = 1814$). For the observed data set the N productivity boundary, using a quantile regression on the 95th percentile, was: $Y = 1069 + 30 \times N - 0.046 \times N2$. For the simulated data set the boundary function was: $Y = -397 + 47 \times N - 0.08 \times N2$. Solid lines represent the median yield response while the dashed lines represent the yield frontiers.

were less responsive to available N than simulated fields.

The N productivity boundaries, using a quantile regression on the 95th percentile, of both the observed (figure 3(a)) and the simulated (figure 3(b)) data sets were also best described as quadratic functions. As with the average function the observed N response function was not as steep as the simulated function and did not peak within the range of observed data (the yield boundary was 6121 kg ha^{-1} at 350 kgN ha^{-1}) while the simulated boundary function peaked at about 280 kgN ha^{-1} with a yield of 7330 kg ha^{-1}. The average and boundary functions derived from the observed data in this research are similar if somewhat more conservative than those previously estimated from an earlier subset of these data (Hochman *et al* 2009a). They are also much lower than the 57 kg grain ha^{-1} achieved per applied kgN ha^{-1} for irrigated maize in the USA (Grassini *et al* 2009).

In the absence of a clear calculation method to determine the marginal value of applying N fertiliser to rainfed wheat crops, agronomic advisers and growers tend to apply rules of thumb. In Yield Prophet Lite (www.yieldprophet.com.au/yplite/), for example, a Water Use Efficiency formula is used to convert likely ET into yield potential (in t ha^{-1}) and this value is multiplied by 40 to determine the rate of available N (kgN ha^{-1}) required to achieve that yield. This rule of thumb is represented by a linear function (figure 3) that closely matches the boundary functions derived from this research for available N values up to about 100 kgN ha^{-1} but over-estimates the yield response where available N exceeds 100 kgN ha^{-1}.
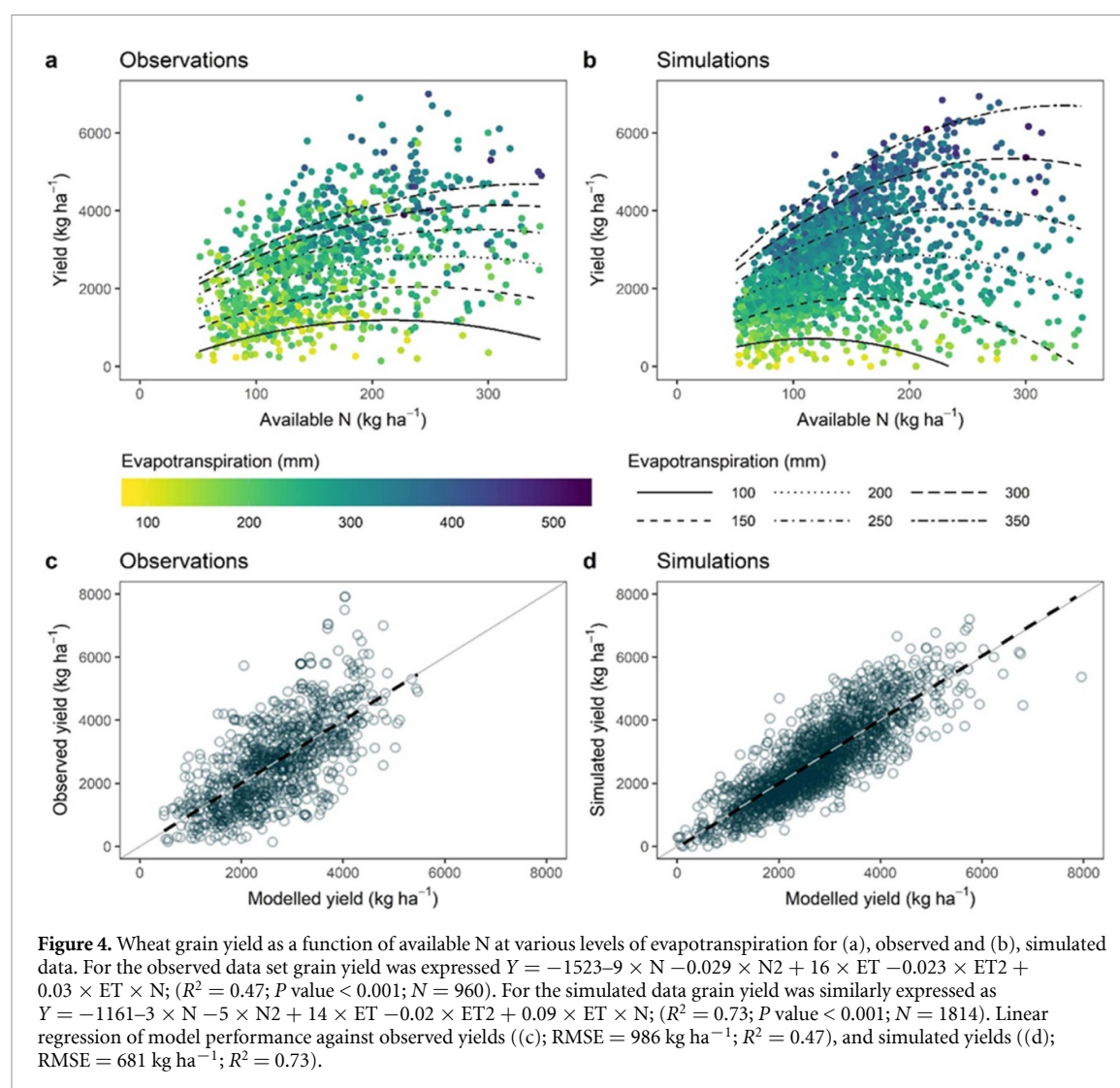
3.3. Yield as a function of ET and available N
Grain yields in both the observed and simulated data sets were expressed as polynomial functions

with respect to the Available N, ET and an Available N $\times$ ET interaction term (figure 4). All terms of the model were statistically significant (P value < 0.05) in both the observed (figure 4(a)) and simulated (figure 4(b)) data sets. This combined ET-Available N model accounted for more of the yield variability than either ET or Available N alone for both the observed and simulated data sets.

A cross-validation of modelled predictions for observed or simulated yield data provides an assessment of the remaining uncertainty of yield predictions using the combined models. With the observed data, there is considerable uncertainty and a saturation effect for yields >4500 kg ha^{-1} (figure 4(c)). A better fit is observed for the simulated data set which has less uncertainty around predicted yields and does not saturate at higher yields (figure 4(d)). The combined (Yield $= f$ (ET, Available N, ET $\times$ Available N)) model accounted for 47% of the yield variance in the observed data and 73% of the yield variance in the simulated data. With both the observed and simulated data models, all parameters, including the interaction term, were statistically significant.

We further developed yield frontier models based on ET and Available N by fitting quadratic models with interaction terms to the 95th percentile of the observed and simulated data sets (figures 5(a) and (b)). All terms were significant (P values < 0.05) for the simulated data. However, in the model obtained with the observed data set the intercept (P value $= 0.528$), the second-order term for ET (P value $= 0.063$), and the interaction (P value $= 0.154$) terms were not significant. These combined models may be regarded as a first step to develop a simple tool to aid commercial wheat growers' decisions about in-season N fertiliser application rates (figures 5(c) and (d)) in response to likely seasonal ET.

Figure 4. Wheat grain yield as a function of available N at various levels of evapotranspiration for (a), observed and (b), simulated data. For the observed data set grain yield was expressed $Y = -1523 - 9 \times N - 0.029 \times N2 + 16 \times ET - 0.023 \times ET2 + 0.03 \times ET \times N$; ($R^2 = 0.47$; P value < 0.001; $N = 960$). For the simulated data grain yield was similarly expressed as $Y = -1161 - 3 \times N - 5 \times N2 + 14 \times ET - 0.02 \times ET2 + 0.09 \times ET \times N$; ($R^2 = 0.73$; P value < 0.001; $N = 1814$). Linear regression of model performance against observed yields ((c); RMSE $= 986$ kg ha^{-1}; $R^2 = 0.47$), and simulated yields ((d); RMSE $= 681$ kg ha^{-1}; $R^2 = 0.73$).
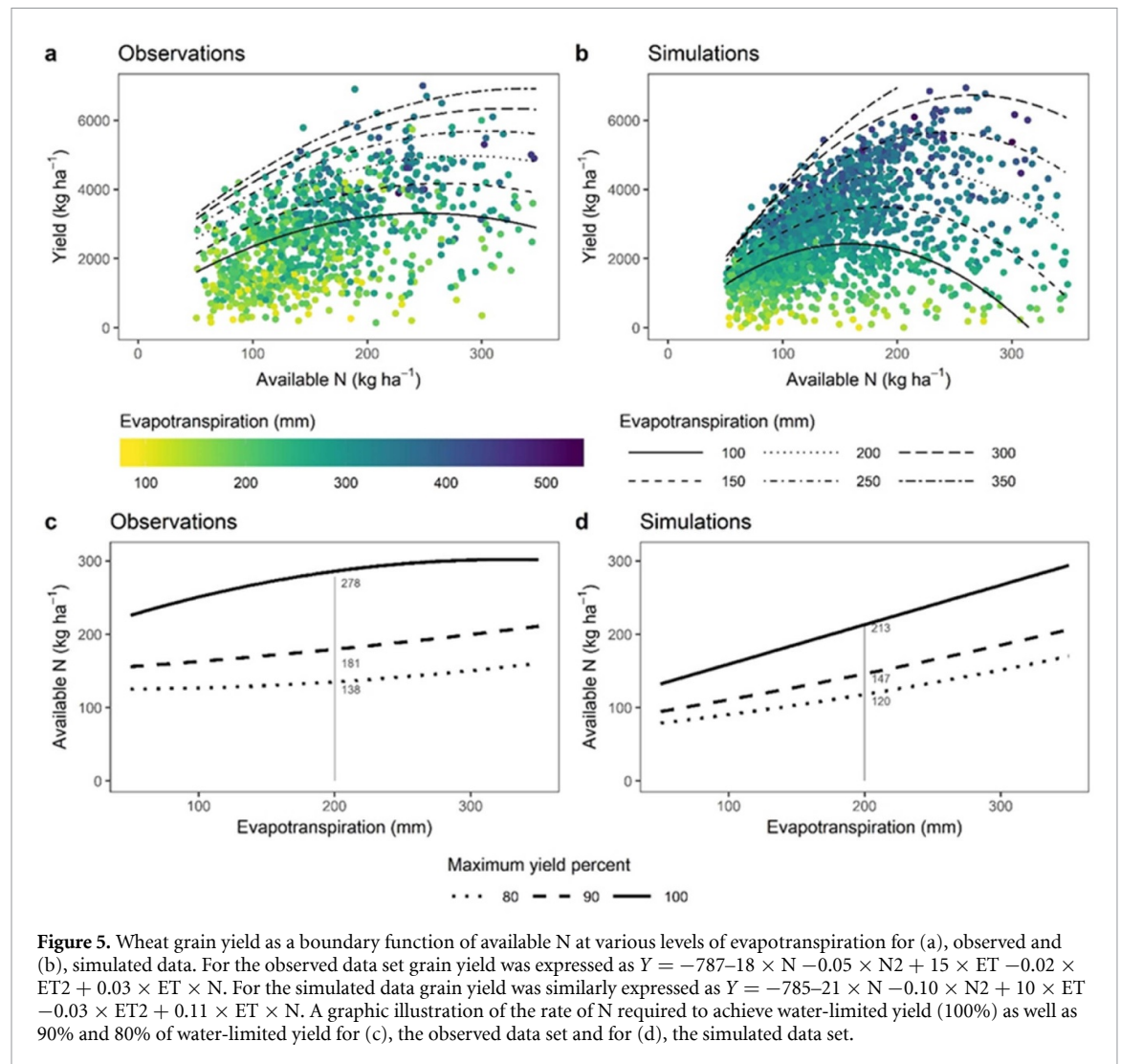
4. Discussion

The empirical model developed here, using just two variables, ET and Available N, accounts for 46% of the variability in wheat growers' yields across highly variable environments throughout the Australian grain zone. This is remarkable, considering the multitude of other factors that can influence grain yield such as solar radiation, temperature, soil properties, extreme climate events, agronomic practices, wheat cultivars and biotic stresses.

In the combined yield response model both ET and Available N are described as quadratic functions with additive effects and a positive interaction term. This differs from all previous characterisations of the combined impacts of ET and Available N limitations on yield (Bloom *et al* 1985, Grimm *et al* 1987, Sinclair and Park 1993, Sadras 2005, Cossani *et al* 2010). The combined response function shows that both factors are simultaneously limiting and hence rejects the application of von Liebig's law of the minimum (Grimm *et al* 1987) to this situation. The interaction term supports the co-limitation hypothesis (Sadras

2005, Cossani *et al* 2010) since it indicates that, as ET and Available N increase, they augment the effectiveness of each factor acting independently. However, this is not a large overall influence on yield, especially under less favourable conditions. More influential in describing the impact of the yield response is the diminishing returns to added ET and Available N implied by the quadratic functions used to describe their impacts on yield. Hence, the assumptions that are built into current rules of thumb for N recommendation, i.e. of von Liebig's Law, and linear response functions to ET and Available N, are not supported by this analysis.

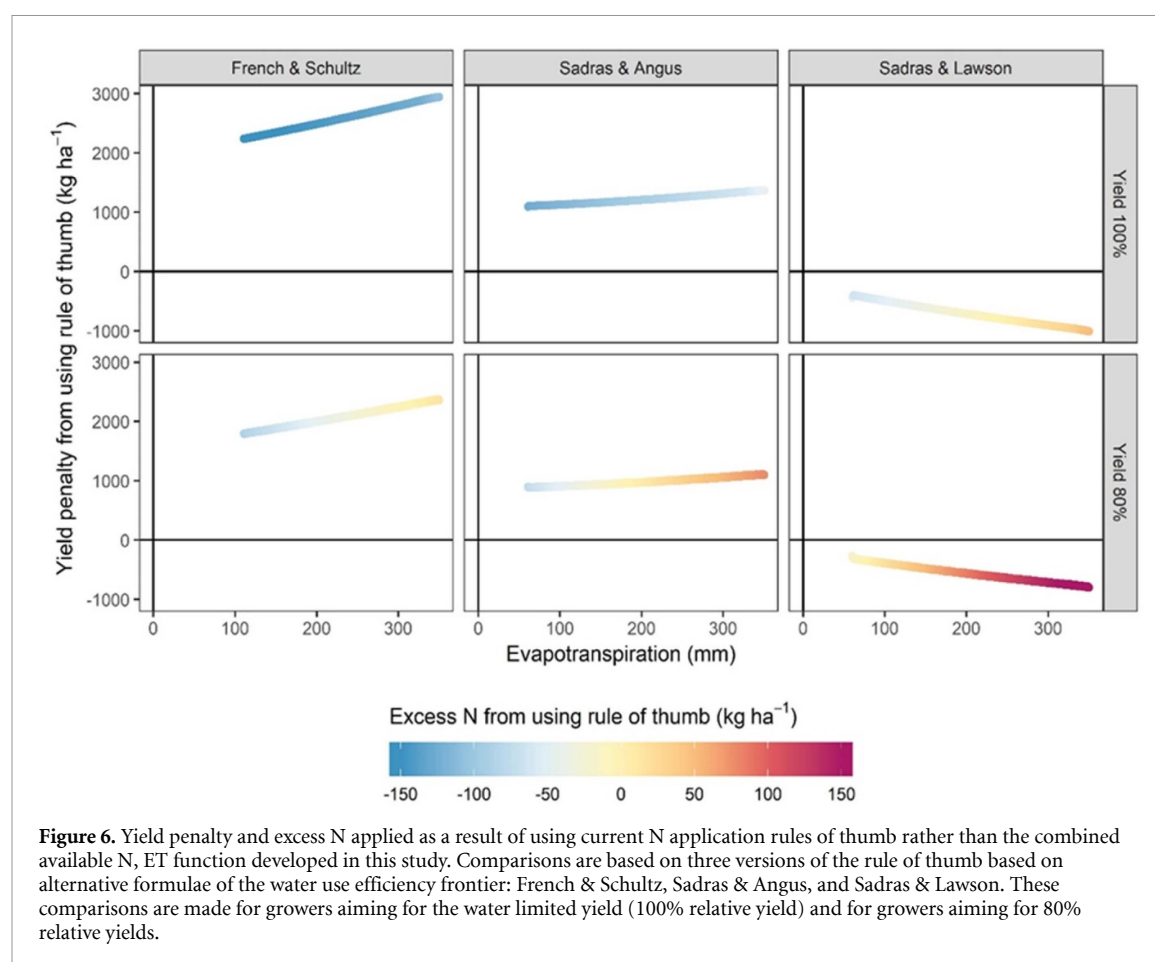### 4.1. Application of the yield function to support N application decisions

Growers need to identify the minimum rate of N required to maximise profit for a given season (or expected ET). Given that the ideal time for in-crop N application is about 3 months before crop maturity, ET can only be estimated with considerable uncertainty. With such uncertainty on top of model uncertainty, risk averse growers may choose to aim for

Figure 5. Wheat grain yield as a boundary function of available N at various levels of evapotranspiration for (a), observed and (b), simulated data. For the observed data set grain yield was expressed as $Y = -787 - 18 \times N - 0.05 \times N2 + 15 \times ET - 0.02 \times ET2 + 0.03 \times ET \times N$. For the simulated data grain yield was similarly expressed as $Y = -785 - 21 \times N - 0.10 \times N2 + 10 \times ET - 0.03 \times ET2 + 0.11 \times ET \times N$. A graphic illustration of the rate of N required to achieve water-limited yield (100%) as well as 90% and 80% of water-limited yield for (c), the observed data set and for (d), the simulated data set.

90% or 80% of the expected water-limited yield (Yw). If, for example, a grower is expecting in-crop ET to reach 200 mm, then according to the model obtained with observed data (figure 5(a)), achieving Yw would require 278 kgN ha^{-1}, while 90% Yw would require 181 kgN ha^{-1} and achieving 80% Yw would only require 138 kgN ha^{-1} (figure 5(c)). Similar calculations can be made with the simulated data (e.g. for 200 mm ET the Available N targets are 213, 147 and 120 kgN ha^{-1} for 100%, 90% and 80% of Yw respectively (figure 5(d)).

The practical implication for wheat growers and their agronomic advisers is that the combined formula could be developed as a decision tool that can be used to fine-tune their in-crop N application decisions. We propose that they first estimate the likely ET for their crop, based on ET to date plus an estimate of ET derived from either historic records or from seasonal climate forecasts, and then choose the level (%) of Yw that they intend to pursue. With these two numbers, they can apply the relationships obtained in figure 5(d) to determine their N rate.

Matching the right amount of N fertiliser to the water-limited yield is critical for farmers' income and for the environment. Too little fertiliser results in yield losses, too much fertiliser results in economic waste and environmental harm. It is thus instructive to compare the outcome of results of applying the model expressed in figure 5(b) against the current rules of thumb as this would impact on the yield potential expected for any ET forecast and on the amount of mineral N recommended. Here we present a comparison of the current model outputs against the Yield Prophet Lite rule of thumb for three alternative wheat water-use efficiency (WUE) boundary functions (French and Schultz 1984, Sadras and Angus 2006, Sadras and Lawson 2013) when aiming for either 100% of Yw, or a more risk averse 80% (figure 6; methodological details on how to construct this figure are presented in the supplemental information (available online at https://stacks.iop.org/ERL/15/114060/mmedia)). Growers aiming for the water-limited yield (the top three panels of figure 6) and using the French and Schultz (1984) WUE frontier to estimate it, will

Figure 6. Yield penalty and excess N applied as a result of using current N application rules of thumb rather than the combined available N, ET function developed in this study. Comparisons are based on three versions of the rule of thumb based on alternative formulae of the water use efficiency frontier: French & Schultz, Sadras & Angus, and Sadras & Lawson. These comparisons are made for growers aiming for the water limited yield (100% relative yield) and for growers aiming for 80% relative yields.

underestimate their Yw by 2.2–3.0 t ha^{-1} as ET rises from 150 to 350 mm. This will potentially lead growers to apply 110–150 less kgN ha^{-1} than required to achieve Yw. A similar, though less dramatic, result was observed for growers using the Sadras and Angus (2006) WUE frontier formula. Here Yw is underestimated by 1.1–1.4 t ha^{-1} as ET rises from 150 to 350 mm, potentially leading growers to apply 40–130 less kgN ha^{-1} than required to achieve Yw. Conversely, using the Sadras and Lawson (2013) frontier to estimate Yw would lead to an overestimate of 0.4–1.0 t ha^{-1} as ET rises from 150 to 350 mm, potentially leading growers to apply a deficit of 70 to an excess of 50 kgN ha^{-1} relative to the amount of N required to achieve Yw. A qualitatively similar set of results was observed when aiming for 80% of Yw, though in this case, excessive N application rates were more common.

4.2. Reflection on the similarities and differences in analysis of the observed and simulated data
With the three yield models explored (ET, Available N and combined) the same functions emerged for both the observed and the simulated data sets albeit with differences in parameter values. In all cases, the average functions fitted to the observed data were less responsive to Available N and ET than the corresponding functions fitted to the simulated data. With the boundary functions, the yield response to lower input levels was steeper for the simulated data but as ET and Available N increased, the boundary functions were more inclined to peak at lower input levels than for the observed data.

The higher predictive power of the model derived for simulated data is not surprising given that the main determinants of the simulated results are those that affect Available N and ET. Other yield determining factors that vary by year and location and may influence the simulated yields are temperature and solar radiation, as well as the daily distribution (not just the in-crop totals) of Available N and ET throughout the season. Observed yields are also subject to the abovementioned factors but may be further constrained by factors that are not accounted for in the simulations. For example, a multitude of suggested yield-limiting factors including: extremes of temperature, weeds, pests and diseases, agronomic deficiencies, tillage practices, late sowing, seeding density, nutrients other than N, subsoil chemical constraints, and other soil chemical and physical properties including soil organic carbon and plant-available water capacity (Hochman *et al* 2009a). Hence, with the important exception of Available N, the factors that account for the wheat yield gap are the factors that contribute to lower predictive power and the less responsive ET and Available N parameters of the observed data model.

5. Conclusions

Addressing the need for crop growers to match the amount of N fertiliser applied to achieve a crop's water-limited yield in the face of increasingly chaotic rainfall in semi-arid environments is important because while too little fertiliser restricts yields and runs down soil organic carbon, too much fertiliser is economically wasteful and environmentally harmful. Complex simulation models can relate wheat grain yield to available water and available N. Yet, their detailed data input requirements make them inaccessible to most growers who instead rely on simplistic rules of thumb. This is the first paper to develop a simple quantitative model that relates wheat grain yield to available water and available N. The model was developed from analysis of two independent and comprehensive data sets: measured data from growers' fields and data simulated by a well-validated crop simulation model. Both data sets cover the full range of environments and seasonal conditions likely to be encountered in the highly variable Australian grain zone. This simple model differs from other published models and from currently adopted rules of thumb that rely on a linear WUE function in combination with a constant NUE value and the assumption, in accordance with von Liebig's law of the minimum, that either ET or available N limit yield. The models developed here predicts yield in response to both ET and Available N which are described as quadratic functions with additive effects and a positive interaction term. They account for 73% of the variation in the simulated yield data and 46% of the variation in the growers' yield data. Thus, we have used a complex model to derive a relatively simple one that is more sophisticated than the current rules of thumb. The proposed model is simple to use, requiring only readily available input data—simplicity on the far side of complexity.

We propose to incorporate the model developed here into the Yield Prophet Lite decision support tool by replacing its current algorithm without requiring any additional inputs. We expect that this more reliable yet simple decision support tool, when compared to currently used rules of thumb, will lead to fertilizer recommendations that can support better management of the competing food security and environmental imperatives.

Acknowledgments

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Z Hochman ⊙ https://orcid.org/0000-0002-6217-5231
F Waldner ⊙ https://orcid.org/0000-0002-5599-7456

References

ABARE-BRS 2010 *Land Use of Australia, Version 4, 2005–06* (Canberra, Australia: Australian Bureau of Agricultural and Resource Economics—Bureau of Rural Sciences (ABARE-BRS). Economics—Bureau of Rural Sciences (ABARE-BRS)) (https://data.gov.au/data/dataset/a18dbf2b-c0fe-4a26-b39a-e553bf6c39b5)

Addiscott T M, Whitmore A P and Powlson D S 1991 *Farming, Fertilizers and the Nitrate Problem* (Wallingford, UK: CAB International (CABI))

Aigner D, Lovell K and Schmit P 1977 Formulation and estimation of stochastic frontier production models *J. Econom.* **6** 21–37

Bloom A J, Chapin F S I and Mooney H A 1985 Resource limitation in plants—an economic analogy *Annu. Rev. Ecol. Syst.* **16** 363–92

Brisson N *et al* 1998 STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn *Agronomie* **18** 311–46

Cade B S and Noon B R 2003 A gentle introduction to quantile regression for ecologists *Front. Ecol. Environ.* **1** 412–20

Carberry P S, Liang W-L, Twomlow S, Holzworth D P, Dimes J P, Mcclelland T, Huth N I, Chen F, Hochman Z and Keating B A 2013 Scope for improved eco-efficiency varies among diverse cropping systems *Proc. Natl Acad. Sci. USA* **110** 8381–6

Cassman K G, Dobermann A and Walters D T 2002 Agroecosystems, nitrogen-use efficiency, and nitrogen management *Ambio* **31** 132–40

Cossani C M, Slafer G A and Savin R 2010 Co-limitation of nitrogen and water, and yield and resource-use efficiencies of wheat and barley *Crop Pasture Sci.* **61** 844–51

Fischer R A 1981 Optimizing the use of water and nitrogen through breeding of crops *Plant Soil* **58** 249–78

French R J and Schultz J E 1984 Water use efficiency of wheat in a Mediterranean type 414 environment. I. The relation between yield, water use and climate *Aust. J. Agric. Res.* **35** 743–64

Galloway J N, Aber J D, Erisman J W, Seitzinger S P, Howarth R W, Cowling E B and Cosby B J 2003 The nitrogen cascade *BioScience* **53** 341–56

Grassini P and Cassman K G 2012 High-yield maize with large net energy yield and small global warming intensity *Proc. Natl Acad. Sci. USA* **109** 1074–9

Grassini P, Yang H and Cassman K G 2009 Limits to maize productivity in Western Corn-Belt: a simulation analysis for

fully irrigated and rainfed conditions *Agric. For. Meteorol.* **149** 1254–65

Grimm S S, Paris Q and Williams W A 1987 A von Liebig model for water and nitrogen crop response *West. J. Agric. Econ.* **12** 182–92

Grindlay D J C 1997 Towards an explanation of crop nitrogen demand based on the optimization of leaf nitrogen per unit leaf area *J. Agric. Sci. Camb.* **128** 377–96

Hastie T, Tibshirani R and Friedman J 2001 *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Berlin: Springer)

Hochman Z *et al* 2009b Re-inventing model-based decision support: 4. Yield Prophet®, an internet-enabled simulation-based system for assisting farmers to manage and monitor crops in climatically variable environments *Crop Pasture Sci.* **60** 1057–70

Hochman Z, Carberry P S, Robertson M J, Gaydon D S, Bell L W and Mcintosh P C 2013 Prospects for ecological intensification of Australian agriculture *Eur. J. Agron.* **44** 109–23

Hochman Z, Gobbett D L and Horan H 2017 Climate trends account for stalled wheat yields in Australia since 1990 *Glob. Change Biol.* **23** 2071–81

Hochman Z, Gobbett D, Horan H and Navarro Garcia J 2016 Data rich yield gap analysis of wheat in Australia *Field Crops Res.* **197** 97–106

Hochman Z, Holzworth D P and Hunt J R 2009a Potential to improve on-farm wheat yield and water-use efficiency in Australia *Crop Pasture Sci.* **60** 708–16

Hochman Z and Horan H 2018 Causes of wheat yield gaps and opportunities to advance the water-limited yield frontier in Australia *Field Crops Res.* **288** 20–30

Holzworth D P *et al* 2014 APSIM–evolution towards a new generation of agricultural systems simulation *Environ. Model. Softw.* **62** 327–50

Isbell R F 1996 *The Australian Soil Classification* (Collingwood, VIC: CSIRO Publishing)

Jones J W, Hoogenboom G, Porter C H, Boote K J, Batchelor W D, Hunt L A, Wilkens P W, Singh U, Gijsman A J and Ritchie J T 2003 The DSSAT cropping system model *Eur. J. Agron.* **18** 235–65

Karam F, Kabalan R, Breidi J, Rouphael Y and Oweis T 2009 Yield and water-production functions of two durum wheat cultivars grown under different irrigation and nitrogen regimes *Agric. Water Manag.* **96** 603–15

Kassie B T, Asseng S, Porter C H and Royce F S 2016 Performance of DSSAT-Nwheat across a wide range of current and future growing conditions *Eur. J. Agron.* **81** 27–36

Keating B *et al* 2003 An overview of APSIM, a model designed for farming systems simulation *Eur. J. Agron.* **18** 267–88

Koenker R 2019 Quantreg: quantile regression. R package version 5.54 (available at: https://CRAN.R-project.org/package=quantreg)

Kuhn M 2008 Building predictive models in R using the caret package *J. Stat. Softw.* **28** 1–26

Lassaletta L, Billen G, Grizzetti B, Anglade J and Garnier J 2014 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland *Environ. Res. Lett.* **9** 105011

Latiri-Souki K, Nortcliff S and Lawlor D W 1998 Nitrogen fertilizer can increase dry matter, grain production and radiation and water use efficiencies for durum wheat under semi-arid conditions *Eur. J. Agron.* **9** 21–34

Lenth R V 2009 Response-surface methods in R, using RSM *J. Stat. Softw.* **32** 1–17

Lin L I-K 1989 A concordance correlation coefficient to evaluate reproducibility *Biometrics* **45** 255–68

Lollato R P *et al* 2019 Agronomic practices for reducing wheat yield gaps: a quantitative appraisal of progressive producers *Crop Sci.* **59** 333–50

Long N V, Assefa Y, Schwalbert R and Ciampitti I A 2017 Maize yield and planting date relationship: a synthesis-analysis for US high-yielding contest-winner and field research data *Front. Plant Sci.* **8** 2106

Mueller N D, West P C, Gerber J S, Macdonald G K, Polasky S and Foley J A 2014 A tradeoff frontier for global nitrogen use and cereal production *Environ. Res. Lett.* **9** 054002

Passioura J B 2002 Environmental biology and crop improvement *Funct. Plant Biol.* **29** 537–46

Phillips S J, Anderson R P and Schapire R E 2006 Maximum entropy modelling of species geographic distributions *Ecol. Modell.* **190** 231–59

Probert M E, Dimes J P, Keating B A, Dalal R C and Strong W M 1998 APSIM's water and nitrogen modules and simulation of the dynamics of water and nitrogen in fallow systems *Agric. Syst.* **56** 1–28

R Development Core Team 2009 R: a language and environment for statistical computing R Foundation for Statistical Computing, Vienna

Sadras V O 2005 A quantitative top-down view of interactions between stresses: theory and analysis of nitrogen-water co-limitation in Mediterranean agroecosystems *Aust. J. Agric. Res.* **56** 1151–7

Sadras V O and Angus J F 2006 Benchmarking water use efficiency of rainfed wheat in dry environments *Aust. J. Agric. Res.* **57** 847–56

Sadras V O, Hayman P T, Rodriguez D, Monjardino M, Bielich M, Unkovich M, Mudge B and Wang E 2016 Interactions between water and nitrogen in Australian cropping systems: physiological, agronomic, economic, breeding and modelling perspectives *Crop Pasture Sci.* **67** 1019–53

Sadras V O and Lawson C 2013 Nitrogen and water-use efficiency of Australian wheat varieties released between 1958 and 2007 *Eur. J. Agron.* **46** 34–41

Sharma B D, Kar S and Cheema S S 1990 Yield, water use and nitrogen uptake for different water and N levels in winter wheat *Fertil. Res.* **22** 119–27

Sinclair T R and Horie T 1989 Leaf nitrogen, photosynthesis, and crop radiation use efficiency: a review *Crop Sci.* **29** 90–98

Sinclair T R and Park W I 1993 Inadequacy of the Liebig limiting-factor paradigm for explaining varying crop yields *Agronomy J.* **85** 742–6

Smith R C G and Harris H C 1981 Environmental resources and restraints to agricultural production in a Mediterranean type environment *Plant Soil* **58** 31–57

Stevenson M *et al* 2020 epiR: tools for the analysis of epidemiological data. R package version 1.0-14 (available at: https://CRAN.R-project.org/package=epiR)

Sutton M A *et al* 2013 Our nutrient world: the challenge to produce more food and energy with less pollution NERC/Centre for Ecology and Hydrology, Edinburgh

Tittonell P and Giller K E 2013 When yield gaps are poverty traps: the paradigm of ecological intensification in African smallholder agriculture *Field Crops Res.* **143** 76–90

van Ittersum M K, Cassman K G, Grassini P, Wolf J, Tittonell P and Hochman Z 2013 Yield gap analysis with local to global relevance—a review *Field Crops Res.* **143** 4–17

van Keulen H and Seligman N G 1987 *Simulation of Water Use, Nitrogen Nutrition and Growth of a Spring Wheat Crop. Simulation Monograph* (Wageningen: PUDOC)

van Rees H, Mcclelland T, Hochman Z, Carberry P, Hunt J, Huth N and Holzworth D 2014 Leading farmers in South East Australia have closed the exploitable wheat yield gap: prospects for further improvement *Field Crops Res.* **164** 1–11

van Wart J *et al* 2013 Use of agro-climatic zones to upscale simulated crop yield potential *Field Crops Res.* **143** 44–55

Vitousek P M *et al* 2009 Nutrient imbalances in agricultural development *Science* **324** 1519–20