

RESEARCH

Open Access

Building capacity for assessing spatial-based sustainability metrics in agriculture

Louis Kouadio^{1,2*} and Nathaniel K Newlands¹

* Correspondence:

Amani.Kouadio@usq.edu.au

¹Lethbridge Research Centre,
Science and Technology Branch,
Agriculture and Agri-Food Canada,
5403 – 1st Avenue South, P.O. Box
3000, Lethbridge, AB, T1J 4B1, Canada

²Present address: International
Centre for Applied Climate Sciences,
University of Southern Queensland,
West Street, Toowoomba, QLD,
4350, Australia

Abstract

Crop yield is influenced over time and space, namely, by a wide range of variables linked with crop genetics, agronomic management practices and the environment under which the crop dynamically responds to maximize growth potential and survival. Such variability can pose substantial uncertainty and risks in the use of agricultural sustainability decision-making frameworks that include crop yield as a leading metric. Here, decision analytics can play a vital role by guiding the use of statistical-based analytics to build in a higher degree of intelligence to enable better predictive (i.e., crop yield forecasting both over the growing season and inter-annually) and prescriptive (optimization across crop areas and subdivisions) approaches. While inter-annual variability in yield can be modelled based on a deterministic trend with stochastic variation, quantifying the variability of yield and how it changes across different spatial resolutions remains a major knowledge gap. To better understand how yield scales spatially, we integrate in this study, for the first time, multi-scale crop yield of spring wheat and its variance (i.e., field to district to region) obtained within the major wheat growing region of the Canadian Prairies (Western Canada). We found large differences between the mean and variance from field to district to regional scales, from which we determined spatially-dependent (i.e., site specific) scaling factors for the mean and variance of crop yield. From our analysis, we provide several key recommendations for building capacity in assessing agricultural sustainability using spatial-based metrics. In the future, the use of such metrics may broaden the adoption and consistent implementation of new sustainable management protocols and practices under a precautionary, adaptive management approach.

Keywords: Agriculture; Metrics; Scaling; Sustainability; Uncertainty

Background

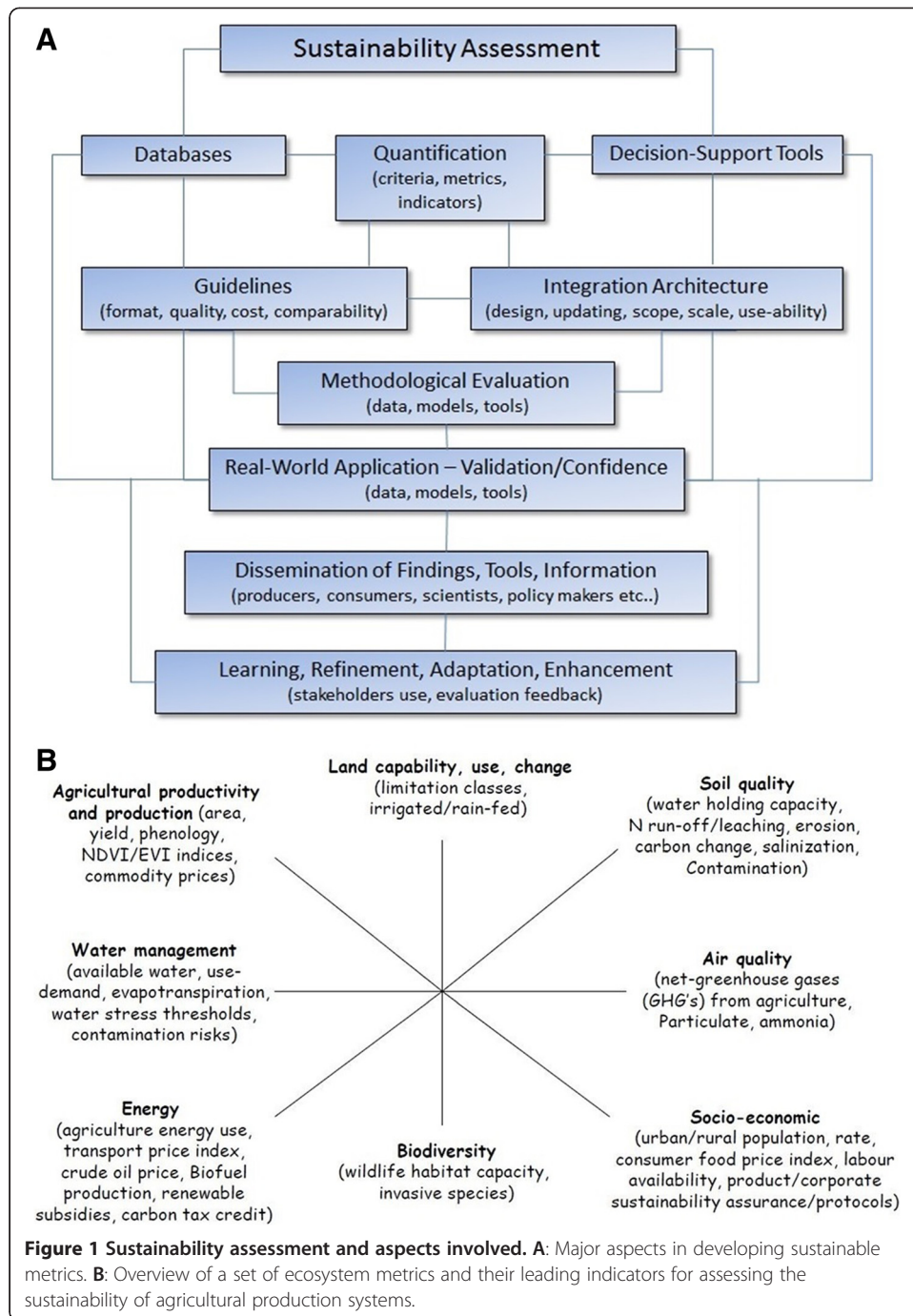
Current issues and challenges in agriculture sustainability assessment

Agricultural production systems integrate data and knowledge on climate, soils, crops and livestock, alongside economic elements, such as material, labour, energy inputs, and food and services outputs. Often, such information is not always available within a study region, and so proxy data or information from other similar regions is utilized. This process can be very subjective. In many real-world applications, there is a need to upscale and/or downscale information in time and/or space to align different sources of data and information. While there are a wide range of geospatial-based scaling techniques, there is no clear consensus. Each technique, whether deterministic or stochastic, can have very different assumptions and levels of suitability and reliability/confidence

associated with a given application. As highlighted by Popp et al. (2004), economists, policy analysts and other agricultural decision-makers, in the case of crop yield risk, often have a variety of different sources of information available to deal with. This includes experimental plot data from agricultural field research plots, agroecosystem model predictions, survey farm-level sampling data, or regional- and national-scale estimates from various governmental or private agencies. However, agricultural production systems are affected by socio-economic and cultural processes at local, regional, national, and international scales, including markets and trade, policies, trends in rural/urban population, and technological development (de Vries 2013; INFASA 2006; Tubiello and Rosenzweig 2008). For this reason, when using and integrating environmental and economic information for assessing the sustainability of agricultural production systems, the relative dependence of data or information to a given scale needs to be determined and the valid range over which it can be reliably scaled in time and space shall be determined.

Building capacity of development is the process by which individuals, groups, organizations, institutions or societies increase their abilities to perform core functions, solve problems, define and achieve objectives, and understand and deal with their development needs in a broad context and in a sustainable manner (De Grauwe 2009). To build capacity for assessing the sustainability of agricultural production systems, the scaling behaviour of selected sustainability metrics must be statistically assessed and validated in time and space. Different metrics for agricultural production systems (Figure 1) are selected to span multiple spatial and temporal scales, i.e., from local to regional, national and global, and short-term and long-term information on impacts. A metric is a quantitative measure of the degree that a system, component, or process can be measured by a given attribute. Because major aspects are involved in building sustainability frameworks (Figure 1), better integrating a sufficient amount of knowledge and updating it so that it remains relevant and accurate to current day is of great importance. However, there is a lack of consensus on how best to integrate and utilize existing types of agricultural data and agroecosystem model outputs in terms of what combinations of knowledge are best suited to accurately assess agricultural production and processing systems, and enable compliance with emerging sustainability requirements. The existing methodologies (e.g., the IRIS (<https://iris.thegiin.org/>) catalog of the Global Impact Investing Network) are generally based on different sets of assumptions, sustainability measures (i.e., criteria/sets of indicators), and accuracy of input data sets. Simply adding more indicators (an indicator is defined as a variable, measured or derived, that attains a given process state or condition) is a misguided goal. The indicator choice and weighting are inherently very subjective processes, depending on value and importance judgments (Morse et al. 2001). Often there is very little consistency between different frameworks and they are framed based more on what is available than what is needed to reliably assess sustainability across all industrial sectors. Understanding interconnections, risks and beneficial trade-offs is key to ensuring greater efficiency and equitability of agricultural resource supply, alongside long-term sustainability and resilience of agricultural systems.

Agricultural production depends upon various and interconnected factors (e.g., crop genetics, agronomic management practices, the environment under which it dynamically responds to maximize crop growth potential and survival). Recent findings from the inter-comparison and validation of agricultural crop models indicate that the



prediction of crop yield over time and space is highly variable and, in turn, the relative responses to climate shocks predicted by models can widely disagree (Nelson et al. 2014). Such variability pose substantial uncertainty and risks in the use of agricultural sustainability decision-making frameworks that include crop yield as a leading metric. To address this problem, decision analytics can play a vital role by guiding the use of statistical-based analytics to build in a higher degree of intelligence to enable better predictive (i.e., crop yield forecasting both over the growing season and inter-annually) and prescriptive (optimization across crop areas and subdivisions) approaches. While

inter-annual variability in yield can be modelled under various deterministic trends with stochastic variation, quantifying the variability of yield and how it changes across different spatial resolutions remains a major knowledge gap. In some studies, field data is available and is used to train and validate a crop model. In other applications, regional estimates of crop yield mean and variance is measured by sampling a randomly determined subset of fields within a region, without sampling all fields.

Research objectives

To better understand how yield scales spatially, we integrate, for the first time, multi-scale crop yield of spring wheat and its variance (i.e., field to district to region) obtained within the major wheat growing region of the Canadian Prairies (Western Canada). The paper aims to build knowledge capacity for improving the design and development of a sustainability assessment framework for agricultural production systems and spatial-based sustainability metrics. We report selected findings from a multi-scale analysis of spring wheat yield as a leading metric of the sustainability performance of agricultural production systems (crop yield being typically included within existing national and regional sustainability development frameworks). Thus, we explore multi-scales scaling behaviour of observed crop yield from field to sub-regional to regional scales and highlight major multi-scale issues and multi-dimensional hurdles that must be overcome to develop reliable metrics. The analysis also provides a comparative assessment of the quantity and quality of alternative, available data sets and agroecosystem model-based outputs for crop yield across the Canadian Prairies (provinces of Alberta, Saskatchewan and Manitoba). To the best of our knowledge, no study related to the spatial scaling of crop yield from field to rural municipality (RM), to ecodistrict (ECD), to census agricultural regions (CAR) has been previously reported in the literature. This is despite the importance of further geospatial importance of further geospatial analysis to elucidate the spatial scaling/ translation function in the statistical multiscaling function, and greater confidence in estimates of scale dimension and exponent for crop yield. In the future, the use of such leading metrics may broaden the adoption and consistent implementation of new sustainable management protocols and practices under a precautionary, adaptive management approach.

Methods

Yield data sources

Yield data of spring wheat (all cultivars included) at CAR, RM and ECD scales across the provinces of Alberta, Saskatchewan and Manitoba were used. The data spanned different periods according to the spatial unit: from 1976 to 2013, 1992 to 2010, and 1976 to 2012 for CAR, ECD and RM scales, respectively. For the analysis, spatial units with crop land were considered. The total number of CARs, ECDs, and RMs for each of the Prairie provinces is given in Table 1. Only yield data at RM scale for Saskatchewan is considered. They were retrieved from the Government of Saskatchewan website (<http://www.agriculture.gov.sk.ca/Statistics-Crops>). Yield data at CAR scale come from the Field Crop Reporting Series (FCRS) of the Agriculture Division, Statistics Canada (Statistics Canada 2014). The reported yield is obtained by dividing the total production by the harvested area in a given statistic unit. The methodology and error control of the crop survey are explained in the online documents accompanying the FCRS

Table 1 Total number of spatial units with crop land extent across the Canadian Prairie provinces

Province	Census agricultural region	Ecodistrict	Rural municipality
Alberta	8	103	-
Manitoba	12	47	-
Saskatchewan	20	108	295
All Prairie provinces	40	233	-

Only the number of rural municipalities (RM) for Saskatchewan is provided as yield data is available at the RM scale for this province.

(Statistics Canada 2014). Data at RM scale for the province of Saskatchewan are collected in the same manner. Yields at ECD scale were derived from the November Crop Survey files of the FCRS and geospatial characteristics. Indeed, for each year over the 1992–2010 period, the census subdivision level data were interpolated at the ECD level based on the crop land extent map (derived from the Land Cover for Agricultural Regions of Canada (<http://data.gc.ca/data/en/dataset/f5ded3b0-a5b4-4599-95d6-d853a825792b-2999>)) and the ECD boundary map (Ecological Stratification Working Group 1996). Field-scale data at three selected sites [i.e., Swift Current, Melfort (province of Saskatchewan) and Lacombe (province of Alberta)] from Agriculture and Agri-Food Canada (AAFC) experimental plots were considered in this study. No long-term field-scale yield data were available for the province of Manitoba.

Assessing the scaling behaviour of crop yield

The use of scaling factors that take into account the degree of spatial heterogeneity linked with field or regional yield data may be especially important when integrating yield data in different areas, i.e., from field to regional scales, or inter-comparing crop models across different agricultural areas (with their input data obtained at different number and location of validation sites).

We assume that crop yield (i.e. a leading sustainability metric) follows a power-law distribution given by (Solé and Bascompte 2006):

$$Y(s) = kA(s)^{-\eta}, \quad \eta > 1, \quad s > s_{min} \quad (1)$$

where $Y(s)$ refers to the crop yield at a given spatial unit s ; $A(s)$ is the area of s ; k is a constant; η is the scaling exponent (as defined in standard fractal dimensional analysis); and s_{min} is a minimum threshold (i.e. non-zero) value for area, s , determined from observational data. From a given spatial scale s to another one αs (α being the shift scale), we have $Y(s)$ that changes to $Y(\alpha s)$. $Y(s)$ is scale-invariant, whereby,

$$Y(s) = \alpha^{-\eta} Y(\alpha s), \quad \eta > 1 \quad (2)$$

where $\alpha^{-\eta}$ is a spatial scaling (or scale translation) factor.

There are two possible representations corresponding to whether $\alpha > 1$ or $\alpha < 1$. In the analysis of scale dependence and invariance in hydrology, values of $\alpha < 1$ are typically obtained (Sposito 1998). Such values for α might also be generally expected in the

scaling of crop yield given the strong dependence of soil water and hydrological processes, especially in semi-arid and arid regions.

Power laws, with the attribute of scale invariance, appear widely in physics, biology, earth and planetary sciences, economics and finance, computer science, demography and social sciences (e.g., Newman 2005; Brinsmead et al. 2015a, b). In our analysis we utilize a log-log plot to identify power-law scaling behaviour, though Pareto Q-Q quartile plots and mean residual life plots can also be used. The signature of a power law is a linear relationship (i.e. with the scaling exponent η given by the absolute slope of the straight line) in a log-log plot of $Y(s)$ versus $A(s)$. Power laws with exponents less than unity cannot be normalized as its value diverges and do not normally occur in nature (Newman 2005). As detailed by Newman (2005), after computing the normalization constant, the mean value of a power-law distributed variable (i.e., crop yield) is given by:

$$\langle Y(s) \rangle = \frac{k}{2-\eta} [A(s)^{-\eta+2}]_{s_{min}}^{\infty} \tag{3}$$

which becomes infinite if $\eta \leq 2$. Power-laws with such low exponent values have no finite mean, and the mean would diverge if one had an infinite number of samples. In such cases, the mean is not a “well-defined” quantity or metric because it can vary enormously from one measurement to the next, especially at different scales s . If, however, $\eta \geq 2$, then the mean is well-defined and is given by:

$$\langle Y(s) \rangle = \frac{\eta-1}{\eta-2} s_{min}, \eta > 2. \tag{4}$$

For a power-law distribution, the second distribution moment, or root-mean-square (RMS) is given by (Newman 2005):

$$\langle Y^2 \rangle = \frac{\eta-1}{\eta-3} s_{min}^2, \eta > 3. \tag{5}$$

If $\eta \geq 3$, then the variance is well-defined and finite. Under scale invariance conditions (i.e., power-law distributed yield), the mean is proportional to the scale linearly, but its variance is proportional to the scale quadratically. In summary, for crop yield to be a well-defined metric and finite (i.e. enabling simulation-distribution re-sampling that is scale *invariant*), implies that the scaling exponents of its mean across spatial scale s must satisfy $\eta > 2$, and its variance must satisfy $\eta > 3$. Otherwise, it is scale-dependent with a scaling exponent that must be determined from cross-scale observational data, and used to spatially weight variance between different subregions when upscaling or downscaling is undertaken. Thus, there is a high risk that the variance is underestimated when applying simple neighbourhood averaging approaches, along with a substantial bias in the mean, depending on the particular sampling method employed in measuring yield, and the resampling of its observed distribution.

The scaling variance δV is defined as the difference between the variance of crop yield, $Y(s)$ and $Y(\alpha s)$, related by a given finite spatial scaling or translation factor α (Eq. 6).

$$\delta V = \sigma_{Y(s)}^2 - \sigma_{Y(\alpha s)}^2 \tag{6}$$

where

$$\sigma_{Y(s)}^2 = \frac{1}{n} \sum_{i=1}^n (Y(s)_i - \bar{Y}(s))^2, \sigma_{Y(\alpha s)}^2 = \frac{1}{n} \sum_{i=1}^n (Y(\alpha s)_i - \bar{Y}(\alpha s))^2 \tag{7}$$

Where n is the total number of years in our case. The estimated difference in net-variance, δV , reduces to:

$$\delta V = (1 - \alpha^{2\eta}) \bullet \frac{1}{n} \sum_{i=1}^n (Y(s)_i - \bar{Y}(s))^2 \Rightarrow \delta V = (1 - \alpha^{2\eta}) \bullet \sigma_{Y(s)}^2, \alpha > 0.$$

When there is no scaling ($\alpha = 0$), there is no correction to variance required. However, when upscaling or downscaling the leading metric ($\alpha > 0$), a correction factor that increases or decreases the variance, respectively, is required. Spatial scale-dependent exponents necessitate scaling yield by identifying a single or leading set of climate, genetic/phenotypic and agronomic management variables that best (i.e., statistically) explain the observed variance, taking all field, subregional, regional data into consideration in this computation.

Statistical analysis

We generated a cumulative distribution function (CDF) plot of the yield (pooled across all spatial units) for each of the selected site to distinguishing between power-law and log-normal scaling distributions based on heavy-tail behaviour. We referred to the Kolmogorov-Smirnov (KS) statistic and the fitness significance (referred to as p-value; Clauset et al. 2009) to assess the goodness-of-fit between theoretical and empirical distributions. KS is the maximum distance between the CDFs of the observed data and the fitted distribution (Press et al. 1992). The p-value is defined to be the fraction of the synthetic distances between the distribution of the empirical data and the hypothesized model that are larger than the empirical distance (Clauset et al. 2009). The hypothesis of goodness-of-fit with respect to the theoretical model is rejected if p-value < 0.1 (Clauset et al. 2009). Smaller KS scores denote better fit of the theoretical distribution to the empirical one. The methodology used for computing the KS statistic and the p-value is that proposed by Clauset et al. (2009), i.e., maximum likelihood estimation method for gauging the parameters and Monte Carlo procedure for generating synthetic data sets (i.e., 2500 data sets) and then compute the statistics. The R-package “powerLaw” (Gillespie 2014) was used to perform these calculations.

The scaling behaviour was computed through a log-log plot. In this scaling analysis, we considered the annual harvested areas. Only official harvested areas at CAR scale are available from the data source used (Statistics Canada 2014). Therefore, the annual harvested area in a given ECD or RM was calculated as a percentage of this unit area within the corresponding CAR multiplied by the ratio annual harvested area/unit area in this CAR. Whereas the harvested area at field scale was set at 50 ha for the three selected sites. For each of the selected sites, we determined the spatially-dependent

(i.e. site specific) scaling factors for the mean and variance of crop yield. All the analyses were performed using R (R Core Team 2014).

In addition, an example of the performance of the APSIM model (Agricultural Production Systems Simulator; Keating et al. 2003) in a long-term cropping systems trial in Canadian conditions is provided. Although they are a simplification or abstraction of a real (managed) complex dynamic system, agroecosystem models are used to describe the relationships within the Soil-Plant-Atmosphere continuum. The objective of this example is to highlight the complexity in the set-up of agroecosystems at finer scales and the expecting ones when applying them at larger scales (with yield data at these larger scales used for validation). APSIM was calibrated based on experimental data retrieved from previous studies (Wang et al. 2002, 2003; Gervais et al. 2010; Mkhabela and Bullock 2012) involving the three selected sites (Table 2).

Results and discussion

Yield variability

Spring wheat yield varies inter-annually and spatially across the Canadian Prairies (Figure 2). A great variability in field-scale data was observed. Such variability is relatively high due to different management practices (e.g., soil tillage, use efficiency of water and fertilizer), and genetic (crop cultivars, technological improvement). Whereas at large scales the outliers in crop yield data may originate from misreporting, extreme weather conditions, management errors of farmers, or methodological changes in data collection (Finger 2010). However, general similar patterns are evident across the three spatial scales, namely when relating to extreme climate events such as drought (i.e. abnormal dry to exceptional drought conditions across the Great Plains/Midwestern of North America in 2002).

The use of historical yield data within a sustainability assessment framework implies taking into account existing trends in those data. Generally, a technological trend (i.e., better genetics, improved management practices and farm machinery, etc.) is evaluated. Environmental factors such as seasonal climate indices could also be explored. Understanding the historical and future potential spatial impact of multiple teleconnections on crops may provide some insights on yield inter-annual variability and considerably improve spatially-explicit forecasting of future crop potential production and decision-making involving changes in cropland distribution under changing climate variability. Findings on the spatial impact of long-term climate teleconnections (such as El Niño Southern Oscillation, ENSO) shows that it enacts a strong influence on crop yield and its inter-annual variability (Iizumi et al. 2013; Iizumi et al. 2014; Valdez-Cepeda et al. 2007). In the scaling behaviour analysis, no data detrending was performed. We assumed the existence of the same trend throughout the spatial scales for a given site in our analysis. Future works will include trend analyses at each spatial units, i.e. decompose the whole net trend into various trends (technology, climate-related, etc.).

Spatially-dependent scaling in crop yield

The CDF plots of spring wheat yield for each of the selected site are shown in Figure 3. For Lacombe and Swift Current the spring wheat yield pooled across all the spatial units (field, RM, ECD, and CAR) follow a power-law distribution, compared with a

Table 2 Selected sites and data period considered in the analysis

Site	Longitude (dd ¹)	Latitude (dd)	CAR-ID	ECD-ID	RM-ID	Soil type	Climate data period	Yield data period				
								Official statistics			Field scale	Simulated yield ²
							CAR scale	ECD scale	RM scale			
Melfort	-104.6008	52.817	4780	705	428	Black Chernozem	1961-2012	1976-2013	1992-2010	1976-2012	1946-2009 ³	1961-2012
Swift Current	-107.7330	50.2670	4732	825	137	Brown Chernozem	1940-2012	1976-2013	1992-2010	1976-2012	1949-2009 ⁴	1940-2012
Lacombe	-113.7500	52.4500	4850	737	n.a ⁵	Black Chernozem	1961-2012	1976-2013	1992-2010	1976-2012	1946-2009 ⁶	1961-2012

¹degree decimals.

²Yield simulated using the APSIM model.

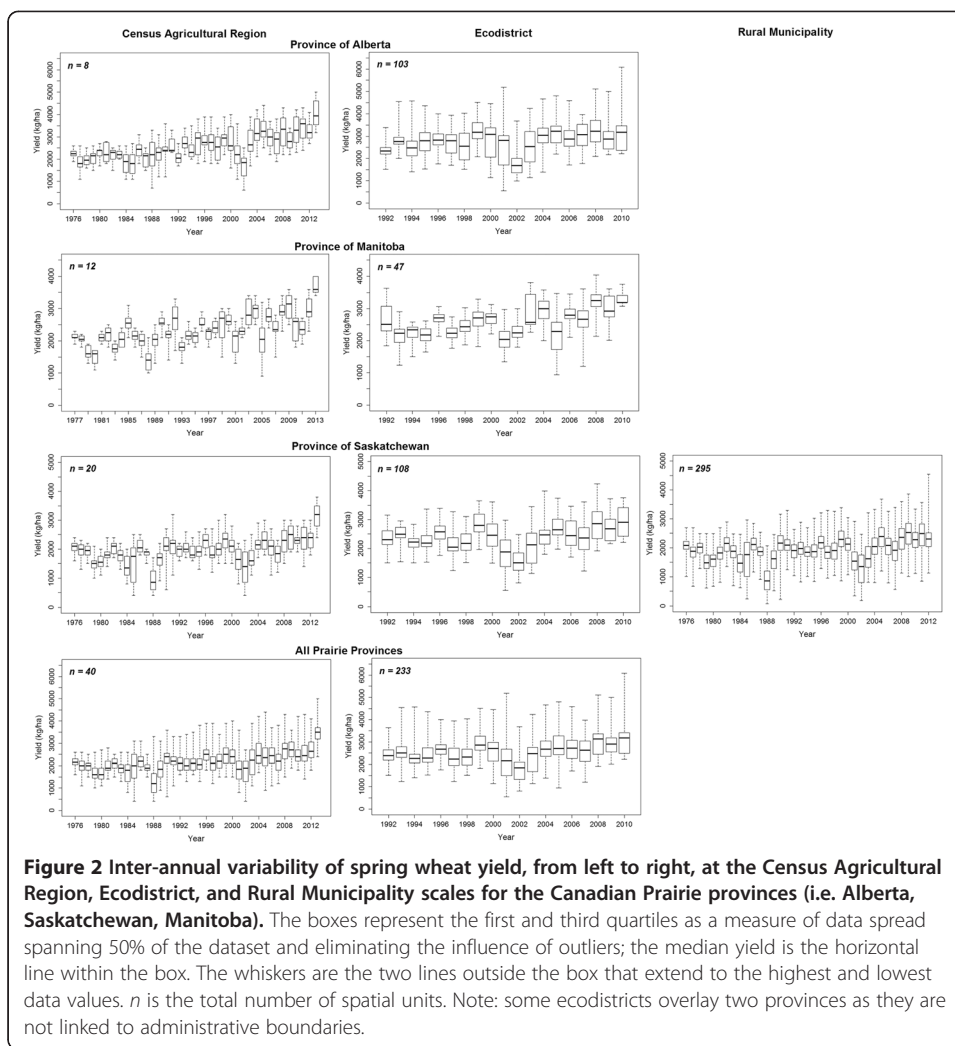
³no data available in 1974–1993.

⁴no data available in 1959, 2006–2007.

⁵not applicable.

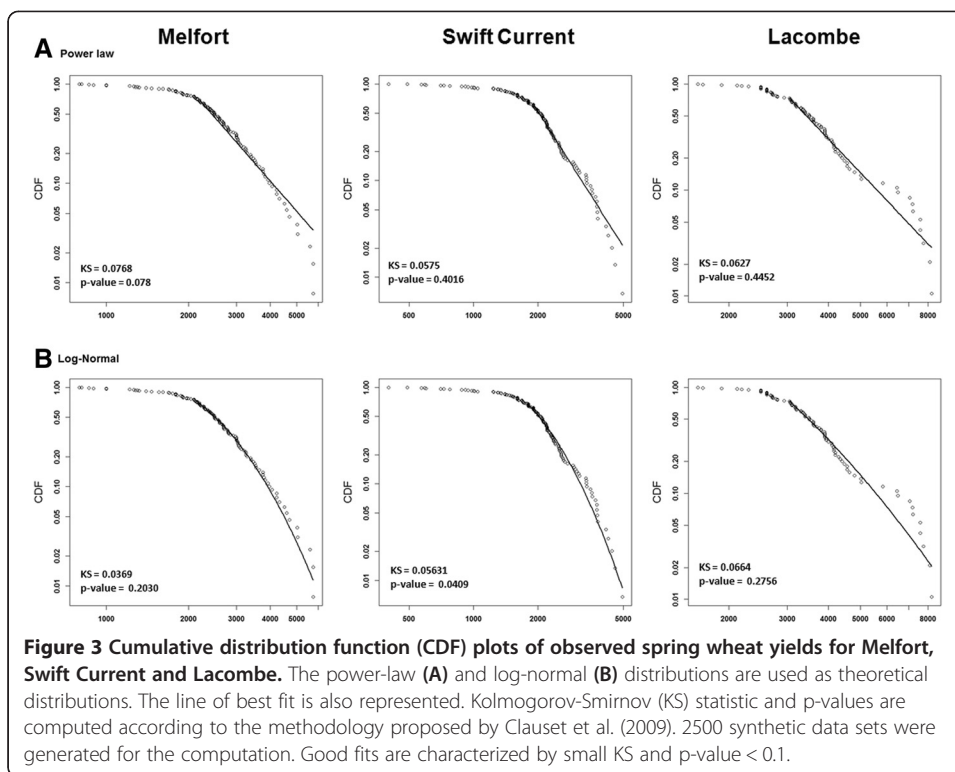
⁶no data available in 1953, 1962, 1971–1993, 2001, 2006–2007.

CAR, ECD, and RM refer to census agricultural region, ecodistrict, and rural municipality, respectively. Climate data include the daily minimum and maximum air temperatures, rainfall and solar radiation (solar radiation data were simulated using the SolarCalc model, Spokas and Forcella (2006)). Field-scale yield data come from AAFC (Agriculture and Agri-Food Canada) experimental plots.



log-normal distribution: lower KS values and p-value greater than 0.1. Whereas for Melfort, the log-normal distribution is the best fit, even though the power-law distribution is also acceptable (KS = 0.0768, p-value ≈ 0.1; Figure 3). The analyses of the distribution assumption of crop yield at each spatial scale for each selected site (not shown) indicated that yield data were power-law distributed in most cases at field scale and that the log-normal distribution outperformed at larger scales (ECD and CAR). However, such results (at each spatial scale) should be dealt with caution since small number of observations is involved (<50). Indeed, the power-law function may appear to be a good fit even if the data are drawn from a non-power-law distribution, in addition to the difficulty to rule out alternative fits to such data (Clauset et al. 2009). Nevertheless, the general trend observed with three sites at four spatial scales in our case study could be replicated with more data. The tail behaviour, as well as the difference of scaling behaviours between the sites in Saskatchewan and the site in Alberta (Figure 3), might be linked to the climate as consequences of its spatial variability.

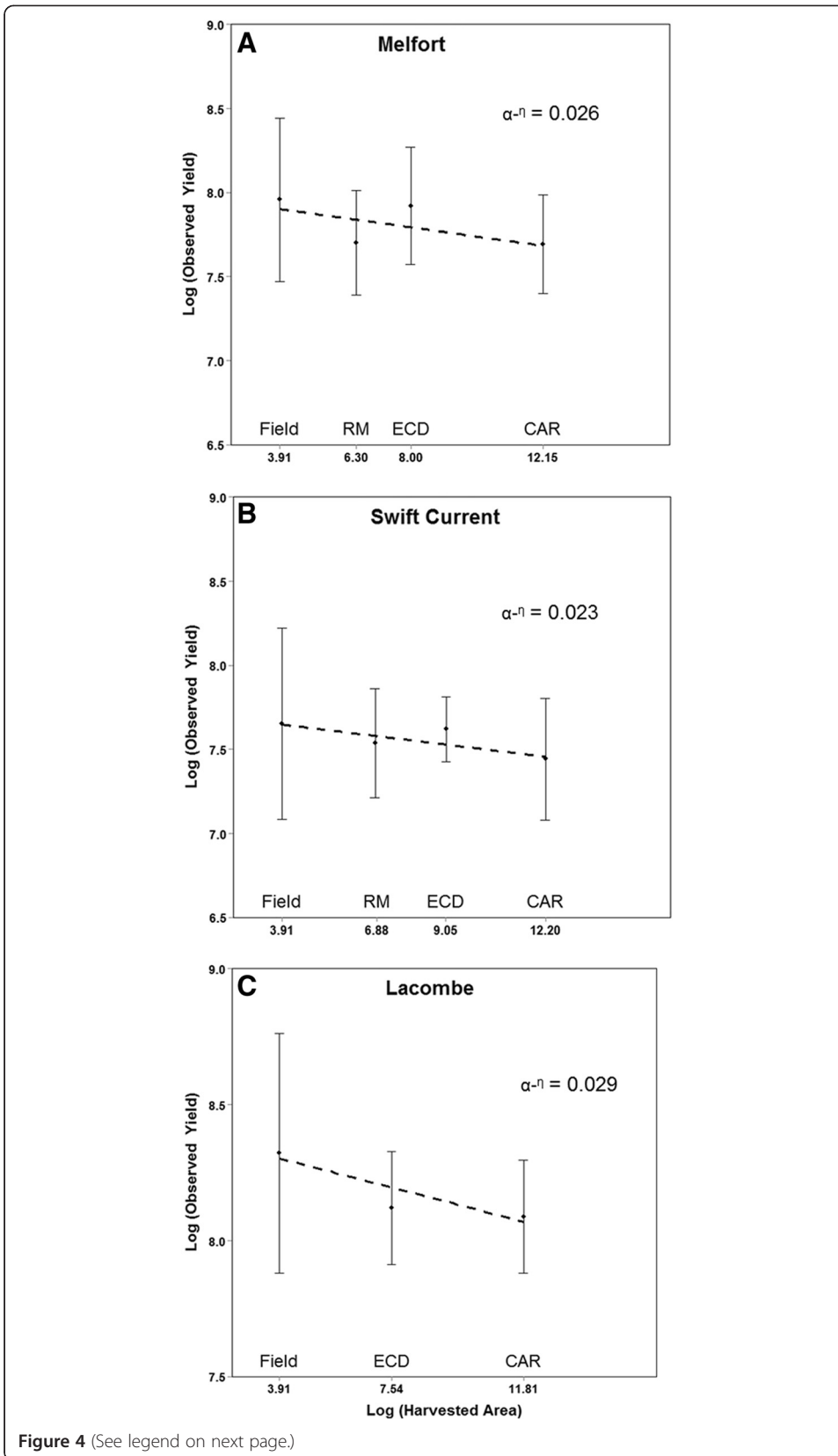
While there are several methods available for estimating the exponent of power-law distributions from empirical data, the maximum likelihood estimation outperforms other methods in both accuracy and precision (White et al. 2008). Generally, one can



expect either a spatially-dependent trend that is homogeneous (i.e. linear) or heterogeneous (e.g. exponential, power-law, etc.) scaling behaviour when comparing the data. An evidence of a spatially-dependent scaling at the selected sites is shown in Figures 4 and 5. Scaling exponents (η) for the mean of crop yield for all the sites (Figure 4) were 2.6%, 2.3% and 2.9%, respectively for Melfort, Swift Current, and Lacombe ($P = 0.013, 0.035, 0.002$, respectively). These values are similar to that obtained in de Wit et al. (2005) with simulated yield data (ratio of ~ 5 for 10–50 km shift). P measures whether an observed result can be attributed to chance. But it cannot answer whether a null hypothesis is correct. The odds depends on how strong the result was, and how plausible the hypothesis is (Nuzzo 2014). Given the scaling factors measure the degree of spatial heterogeneity at each of the selected sites, this highlights the importance of neighbourhood influences on yield data obtained at a specific location (near-field variance), as well as the proportion of explained variance in regional-scale crop yield predictions attributable to far-field variance when sampling across many fields in a given region.

Regarding the variance of crop yield across spatial scales (Figure 5), linear trends were found for Melfort and Swift Current: factors $\alpha^{-\eta} = 2.28$ ($P = 0.021$) and 2.14 ($P = 0.026$), respectively. A quadratic trend was the better fit for the variance in yield for Lacombe, suggesting therefore a heterogeneous scaling behaviour ($\alpha^{-\eta} = 2.48$).

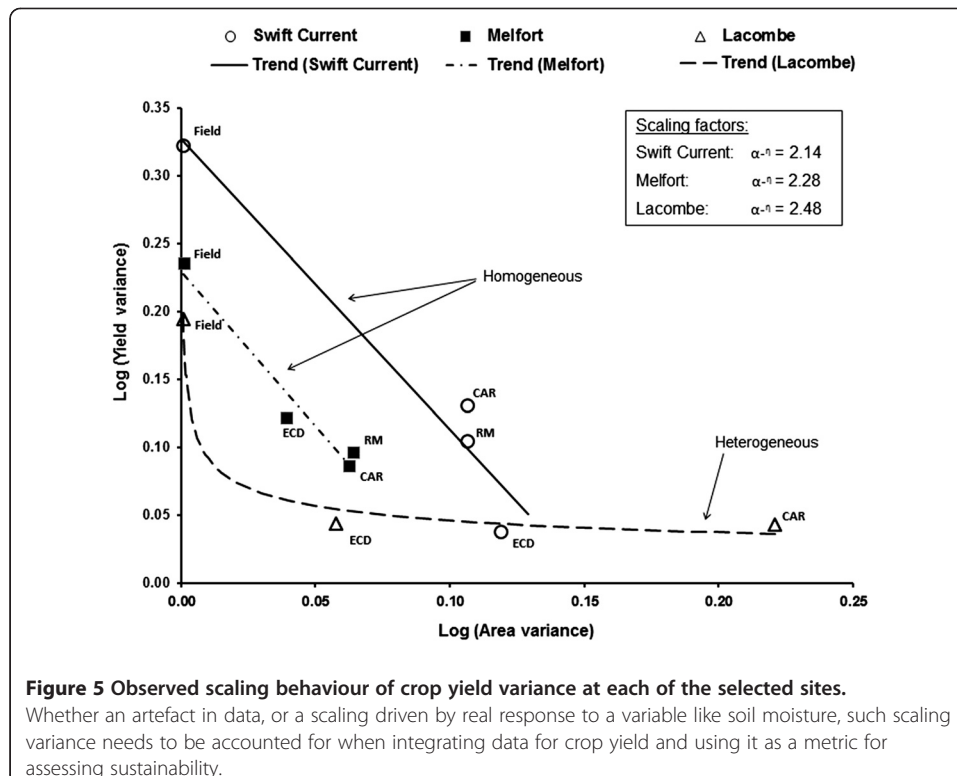
Our findings, based on the estimated scaling exponents in the distribution of the mean and variance of crop yield, indicate that the spring wheat yield is not a well-defined metric at the selected sites. This is because its distribution at any given scale is highly dependent on different variables, conditions, sampling method and other attributes associated with an individual data set. As we demonstrate, employing decision analytics to quantify scaling behaviour in the crop yield distribution, helps to



(See figure on previous page.)

Figure 4 Scaling behaviour of crop yield at each of the selected sites: Melfort (A), Swift Current (B) and Lacombe (C). *RM, ECD, CAR* refer to rural municipality, ecodistrict, and census agricultural region, respectively. The logarithm values of yield data are plotted against the logarithm of the average harvested areas. The scaling factor $\alpha^{-\eta}$ corresponds to the absolute slope of the straight line. For a given site, yield data at RM, ECD and CAR is that of the corresponding spatial unit in which it falls. Data were not detrended. We assumed that the same trend exist throughout the spatial scales for a given site. All the slopes were significant at the confidence level 0.05.

better define it and enable its use as a reliable sustainability metric. Here, specifying the scaling exponent for a given observed/application scale, alongside the cross-scale translation factor (as two additional statistically-based indicators on yield variability) helps to better define the crop yield metric. We find that the spring wheat yield is *not* scale invariant, but scale-dependent whereby a scaling factor $\alpha^{-\eta}$, where it is either constant or non-constant being mediated across spatial scale according to a function of auxiliary soil-crop-water-atmospheric variables (i.e., $\eta = \eta(\dots)$). In either case, when upscaling or downscaling, a scaling correction factor needs to be determined from a scaling analysis across different spatial scales. There is, therefore, a risk when spatially aggregating yield data, whereby variance can be artificially distorted. At the census subdivision (RM and CAR levels) a field survey protocol is typically established to minimize various sources in errors, often by aggregating yield data without adequately accounting for cross-scale spatial variability and extreme values of yield distribution. Such methods frequently ignore specific differences in spatial scaling behaviour and this leads to bias and additional variance when upscaling and/or downscaling data. The use of scaling factors that take into account the degree of spatial heterogeneity linked with field or regional

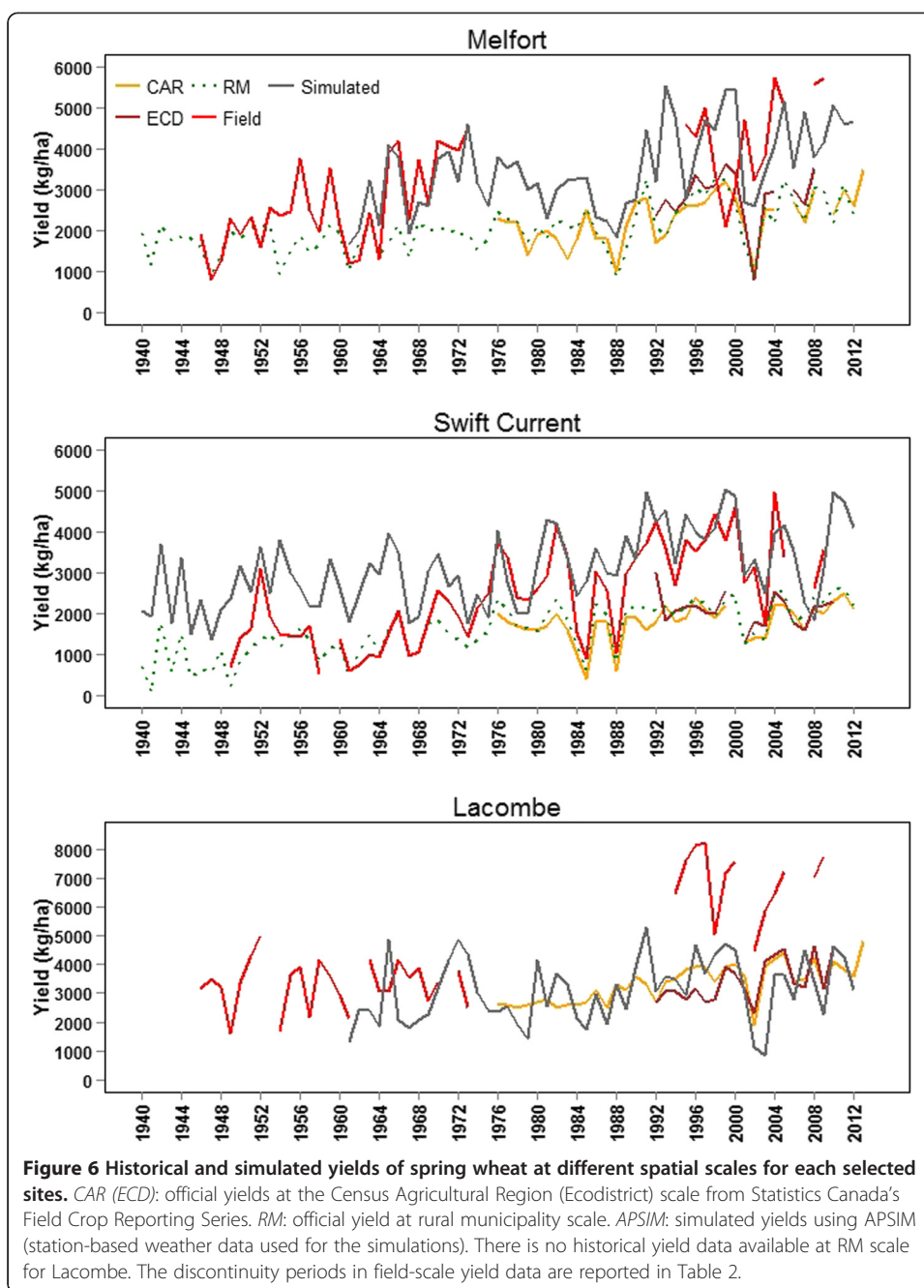


yield data may be especially important when inter-comparing crop models across different agricultural areas and using input data obtained at a different number and location of validation sites and when integrated yield data in different areas from the field to regional scale.

Agroecosystems versus data uncertainties

When supplied with sufficient data, 27 models simulating wheat yield under both normal and climate variability conditions were able to reproduce experimental data well, but forecasting out 50–100 years was more uncertain because of a wide range of possible climate conditions and differences in modelled crop response to increasing temperature and CO₂ (Asseng et al. 2013). In our analysis, though similar patterns are noticed between observed yields at field-scale and simulated yields (Figure 6), the latter are over-estimated in most years at Melfort and Swift Current (1 to 2 times higher than the reported yields), and under-estimated at Lacombe. Nevertheless, the trend in reported yields is captured by the model, suggesting therefore that with sufficient calibration APSIM could give satisfactory outputs for wheat yield forecasting purposes. Using a crop model at regional scale remains challenging because of several aspects, including the number of parameters to be simultaneously calibrated, the availability of data and the uncertainties related to such data when changing scales (e.g. soil properties, sowing dates, irrigation dates).

Reported historical data at regional and national scales are by far the reliable data sets used for modelling purposes (Nelson et al. 2014; Newlands et al. 2014). However, uncertainties (i.e. coefficient of variation) linked to such data are often missing as part of additional information. This kind of information should be readily available to help the scientific community in focusing their research efforts on the scales or units with high variability. Additional biases may also be introduced when integrating different data and models with different assumptions into agricultural sustainability decision-making frameworks. This leads to the incorrect estimation and prediction of crop yield, yield gaps and yield risk. Capturing the distortion or bias in yield when scaling data could be achieved by propagating the uncertainty in the entire distribution of yield (e.g., through Bayesian statistical approaches) and by mapping yield risks based on variance explained by leading explanatory variables within spatially aggregated regions or time-intervals. Typically within sustainability assessment frameworks, additional bias and variance (i.e., uncertainty) introduced when data or model outputs are downscaled or upscaled to align with required spatial and temporal resolution and coverage for a given application, are not considered or assessed sufficiently. However, promising methods do exist and could be applied for this purpose. An example is the probabilistic methodology for adaptively selecting sustainability indicators relevant to a given problem and scale based on a quantitative assessment of the uncertainty of available data and model-based outputs has been devised (Newlands et al. 2013). This approach is able to integrate stakeholder/expert knowledge, empirical and process-based model algorithms using remote-sensing and national agri-environmental datasets, and has undergone preliminary validation testing to assess its ability to predict potential risks and impacts over time and space, in relation to land suitability and nitrogen loading of water for two primary agricultural regions of Canadian (western Prairie provinces and southern Ontario).



Conclusions

Our study emphasizes the spatial scaling behaviour of spring wheat yield across the Canadian Prairie provinces. The results reveal large differences between the mean and variance of crop yield from field to regional scales and different spatially-dependent (i.e., site specific) scaling behaviours. Spring wheat yields at the selected sites are significantly distributed differently across the different spatial scales ($P < 0.05$) and can exhibit a wide range of tail/extremes behaviour. Furthermore, our null hypothesis of invariant scaling in crop yield based on data from three sites across four scales is rejected in favour of the alternative hypothesis of spatial dependence. The clear differences of

scaling behaviour in the distribution of the mean and variance of crop yield could be replicated with more data and extended to other economically important crops. Our analysis also highlights the importance of neighbourhood influences on yield data obtained at a specific location (near-field variance), as well as the proportion of explained variance in regional-scale crop model predictions attributable to far-field variance when sampling across many fields in a given region. Scaling exponents and translation factors can be used to correct either field- or regional-scale yield data in consistent way. In order to take into account the scaling behaviour, reliable protocols for upscaling/downscaling need to be applied, which allow more flexibility on data interpretation and their use in research studies. This should include long time series crop trials data at different locations for efficiently assessing such scaling. Field-scale data are thus urgently needed to enable the estimation of scaling factors for crop yield. Even for a national or sub-national purposes large number of validation sites will be of great importance for the finer-scale data and for choosing the suitable and relevant larger scale to be used. Soil moisture also plays a leading role in influencing crop yield at different scales and supports integrated remote sensing approaches for monitoring and assessing sustainability of open and highly dynamic systems (Phillips et al. 2014). Furthermore, extremes of yield distributions (tail behaviour) need to be taken into account (through fitting a generalized extreme value distribution) when using yield as a metric for sustainability assessment, as well as in upscaling/downscaling techniques and crop modelling and crop yield forecasting.

In Canada, the Federal Sustainable Development Act requires the development and maintenance of systems and procedures to monitor progress on the Federal Sustainable Development Strategy (FSDS (<https://www.ec.gc.ca/dd-sd/default.asp?lang=En&n=9277C8B9-1>)) implementation. To enable such an integration, there is an urgent need to further develop and better assess how agricultural sector-based indicators that rely on historical and scenario geospatial (i.e., spatial-based) data and model outputs could be best integrated into the FSDS framework. A set of agri-environmental indicators and geospatial data on soils, climate and land surface features with statistics on land use and crop and livestock management practices are available through governmental agencies (for example the National Agri-Environmental Health Analysis and Reporting Program of AAFC). The datasets provide valuable, location-specific information on the overall environmental risks and conditions in agriculture across Canada and how they change over time (i.e., every 5 years). However, there is a critical need for long-term time series data from crop trials (i.e., information on cultivars, biomass, yield, crop management, etc.) be made widely available so to enable a more comprehensive analysis to: (i) assess the spatial dependency when upscaling/downscaling data, and (ii) calibrate/validate a suitable agroecosystem to be used in studies involving crop yield forecasting and more generally the impact of the climate change on agriculture. Furthermore, a common data sharing platform for historical climate, yield and other agricultural data that is quality-controlled and continually updated for all end-users is needed. This enables a consistent set of validated scripts for data manipulation and assimilation, modelling and decision analytics.

The case study in this paper shows that there are various hurdles in developing sustainability frameworks, related to the available data and their uncertainties, the spatial scales involved and the performance of process-based crop models used. Biases

introduced when integrating different data and/or crop model outputs with different scaling assumptions, may lead to incorrect estimation and prediction of crop yield, yield gaps and yield risk. While sophisticated, statistical techniques do exist that can correct for distortion of yield variance and bias when aggregating data and for propagating uncertainty when integrating data and model-based output, typically, such methods are not readily available for practitioners. The distortion is therefore either simply ignored, with even simple corrections or remedies that could be minimally undertaken, neglected. We recommend addressing three core aspects in guiding the integration and use of sustainability metrics, namely: complexity (number and type of metrics), applicability (perspective, assumptions, scale) and reliability (known impacts, unknown adaptation, response uncertainty and risk). If these inherent features that are inextricably linked to how coupled environmental, economic and social systems fundamentally behave and interact, are ignored, even a precautionary, adaptive approach could be considered too risky. However, if one can encapsulate these features within a sustainability assessment framework, then they can be better understood, mimicked and controlled. To help in convergence of diverse, interdisciplinary knowledge, participatory decision-making and monitoring approaches are needed, in addition to core metric and sustainability decision analytics.

Abbreviations

AAFC: Agriculture and Agri-Food Canada; APSIM: Agricultural Production Systems Simulator; CAR: Census agricultural region; ECD: Ecodistrict; ENSO: El Niño Southern Oscillation; FCRS: Field Crop Reporting Series; FSDS: Federal Sustainable Development Strategy; RM: Rural municipality.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

NKN designed and funded the research. LK performed the analyses. Both the authors wrote and edited the paper.

Acknowledgement

We thank the Growing Forward II Program and Sustainability Metrics Project of Agriculture and Agri-Food Canada (AAFC) and the National Science and Engineering Council of Canada (NSERC)'s Visiting Fellows in Government Laboratories Program for providing funding support. We also thank Harvey Hill (AAFC) for the project coordination and funding support.

Received: 15 August 2014 Accepted: 17 February 2015

Published online: 20 March 2015

References

- Asseng, S, Ewert, F, Rosenzweig, C, Jones, JW, Hatfield, JL, Ruane, AC, et al. (2013). Uncertainty in simulating wheat yields under climate change. [Letter]. *Nature Climate Change*, 3(9), 827–832. doi:10.1038/nclimate1916.
- Brinsmead, TS, Herr, A, & O'Connell, DA. (2015a). Quantifying spatial dependencies, trade-offs and uncertainty in bioenergy costs: An Australian case study (1) – least cost production scale. *Biofuels, Bioproducts and Biorefining*, 9(1), 21–34. doi:10.1002/bbb.1502.
- Brinsmead, TS, Herr, A, & O'Connell, DA. (2015b). Quantifying spatial dependencies, trade-offs and uncertainty in bioenergy costs: an Australian case study (2) – National supply curves. *Biofuels, Bioproducts and Biorefining*, 9(1), 35–50. doi:10.1002/bbb.1503.
- Clauset, A, Shalizi, CR, & Newman, MEJ. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*, 51(4), 661–703. doi:10.1137/070710111.
- Core Team, R. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- De Grauwe, A. (2009). *Without capacity, there is no development*. Paris, France: United Nations Education, Scientific and Cultural Organization (UNESCO) and International Institute for Educational Planning (IIEP).
- de Vries, BJM. (2013). *Sustainability Science*. New York, NY, USA: Cambridge University Press.
- de Wit, AJW, Boogaard, HL, & Diepen, CAV. (2005). Spatial resolution of precipitation and radiation: the effect on regional crop yield forecasts. *Agricultural and Forest Meteorology*, 135(1–4), 156–168. doi:10.1016/j.agrformet.2005.11.012.
- Ecological Stratification Working Group. (1996). *A National Ecological Framework for Canada. Report and national map at 1:7 500 000 scale*. Ottawa/Hull, Canada: Agriculture and Agri-Food Canada, Research Branch, Centre for Land and Biological Resources Research, and Environment Canada, State of the Environment Directorate, Ecozone Analysis Branch.
- Finger, R. (2010). Revisiting the Evaluation of Robust Regression Techniques for Crop Yield Data Detrending. *American Journal of Agricultural Economics*, 92(1), 205–211. doi:10.1093/ajae/aap021.

- Gervais, M, Bullock, P, Mkhabela, M, Finlay, G, & Raddatz, R. (2010). Improvements to the accuracy of modelled soil water content from the Second Generation Prairie Agrometeorological Model. *Canadian Journal of Soil Science*, 90(3), 523–526. doi:10.4141/CJSS09078.
- Gillespie, CS. (2014). Fitting heavy tailed distributions: the {powerLaw} package. *R package version*, 20, 5. <http://data.gc.ca/data/en/dataset/f5ded3b0-a5b4-4599-95d6-d853a825792b>
<http://www.agriculture.gov.sk.ca/Statistics-Crops>
<https://iris.thegiin.org/>
<https://www.ec.gc.ca/dd-sd/default.asp?lang=En&n=9277C8B9-1>
- Iizumi, T, Luo, J-J, Challinor, AJ, Sakurai, G, Yokozawa, M, Sakuma, H, et al. (2014). Impacts of El Niño Southern Oscillation on the global yields of major crops. [Article]. *Nat Commun*, 5. doi:10.1038/ncomms4712.
- Iizumi, T, Sakuma, H, Yokozawa, M, Luo, J-J, Challinor, AJ, Brown, ME, et al. (2013). Prediction of seasonal climate-induced variations in global food production. [Letter]. *Nature Clim. Change*, 3(10), 904–908. doi:10.1038/nclimate1945.
- INFASA. (2006). From common principles to common practice. In FJ Hani, L Pinter, & HR Herren (Eds.), *Proceedings of First Symposium of the International Forum on Assessing Sustainability in Agriculture (INFASA)* (pp. 1–248).
- Keating, BA, Carberry, PS, Hammer, GL, Probert, ME, Robertson, MJ, Holzworth, D, et al. (2003). An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy*, 18(3–4), 267–288. doi:10.1016/s1161-0301(02)00108-9.
- Mkhabela, MS, & Bullock, PR. (2012). Performance of the FAO AquaCrop model for wheat grain yield and soil moisture simulation in Western Canada. *Agricultural Water Management*, 110, 16–24. doi:10.1016/j.agwat.2012.03.009.
- Morse, S, McNamara, N, Acholo, M, & Okwoli, B. (2001). Sustainability indicators: the problem of integration. *Sustainable Development*, 9(1), 1–15. doi:10.1002/sd.148.
- Nelson, GC, Valin, H, Sands, RD, Havlik, P, Ahammad, H, Deryng, D, et al. (2014). Climate change effects on agriculture: Economic responses to biophysical shocks. *Proceedings of the National Academy of Sciences*, 111(9), 3274–3279. doi:10.1073/pnas.1222465110.
- Newlands, NK, Zamar, DS, Clark, GO, Zhang, Y-C, & McConkey, B. (2013). An integrated assessment model for exploring potential impacts of global change scenarios on the Canadian agricultural system. In *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS)* (pp. 1–10). Maui, Hawaii, USA: IEEE Computer Society Press.
- Newlands, NK, Zamar, DS, Kouadio, LA, Zhang, Y, Chipanshi, A, Potgieter, A, et al. (2014). An integrated, probabilistic model for improved seasonal forecasting of agricultural crop yield under environmental uncertainty. *Frontiers in Environmental Science*, 2. doi:10.3389/fenvs.2014.00017.
- Newman, MEJ. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5), 323–351. doi:10.1080/00107510500052444.
- Nuzzo, R. (2014). Scientific method: Statistical errors. *Nature*, 506, 150–152. doi:10.1038/506150a.
- Phillips, AJ, Newlands, NK, Liang, SHL, & Ellert, BH. (2014). Integrated sensing of soil moisture at the field-scale: Measuring, modeling and sharing for improved agricultural decision support. *Computers and Electronics in Agriculture*, 107, 73–88. doi:10.1016/j.compag.2014.02.011.
- Popp, M, Rudstrom, M, & Manning, P. (2004). Spatial yield risk issues: comparing yield risk across region, crop and aggregation method. In *Proceedings of the 2004 Annual Meeting of the American Agricultural Economics Association* (pp. 1–18). Denver, Colorado, USA: American Agricultural Economics Association.
- Press, WH, Teukolsky, SA, Vetterling, WT, & Flannery, BP. (1992). *Numerical recipes in C: The art of scientific computing* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Solé, RV, & Bascompte, J. (2006). *Self-organization in complex ecosystems (Monographs in population biology)*. New Jersey, USA: Princeton University Press.
- Spokas, K, & Forcella, F. (2006). Estimating hourly incoming solar radiation from limited meteorological data. *Weed Science*, 54(1), 182–189. doi:10.1614/ws-05-098r.1.
- Sposito, G, ed. (1998). *Scale Dependence and Scale Invariance in Hydrology*. Cambridge, UK, Cambridge University Press.
- Statistics Canada (2014). Definitions, data sources and methods of Field Crop Reporting Series, Record number: 3401. URL: <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3401>. (Accessed 15 March 2014): Statistics Canada, Agriculture Statistics Division.
- Tubiello, FN, & Rosenzweig, C. (2008). Developing climate change impact metrics for agriculture. *The Integrated Assessment Journal*, 8(1), 165–184.
- Valdez-Cepeda, RD, Delgado-Ruiz, O, Magallanes-Quintanar, R, Miramontes-De León, G, García-Hernández, JL, Enciso-Muñoz, A, et al. (2007). Scale-invariance of normalized yearly mean grain yield anomaly series. *Advances in Complex Systems*, 10(03), 395–412. doi:10.1142/S0219525907001161.
- Wang, H, McCaig, TN, DePauw, RM, Clarke, FR, & Clarke, JM. (2002). Physiological characteristics of recent Canada Western Red Spring wheat cultivars: Yield components and dry matter production. *Canadian Journal of Plant Science*, 82(2), 299–306. doi:10.4141/P01-107.
- Wang, H, McCaig, TN, DePauw, RM, Clarke, FR, & Clarke, JM. (2003). Physiological characteristics of recent Canada Western Red Spring wheat cultivars: Components of grain nitrogen yield. *Canadian Journal of Plant Science*, 83(4), 699–707. doi:10.4141/P02-166.
- White, EP, Enquist, BJ, & Green, JL. (2008). On estimating the exponent of power-law frequency distributions. *Ecology*, 89(4), 905–912. doi:10.1890/07-1288.1.