A comparison of two methods to predict the landscape-scale variation of crop yield

F.C. Stevenson a, J.D. Knight b,*, O. Wendroth c, C. van Kessel d, D.R. Nielsen e

a 206A Dunlop St., Saskatoon, Sask., Canada S7N 2B7
b Department of Soil Science, University of Saskatchewan, 51 Campus Drive, Saskatoon, Sask., Canada S7N 5A8
 c ZALF Institute of Soil Landscape Research, D-15374 Müncheberg, Germany
d Department of Agronomy and Range Science, University of California, Davis, CA 95616, USA
e Department of Land, Air, and Water Resources, University of California, Davis, CA 95616, USA

Abstract

Landscape-scale variation is a source of information that increasingly is being taken into consideration in agricultural and environmental studies. Models that encompass and interpret this variation in fields and across contrasting management practices have the potential to improve the landscape management of agroecosystems. Our objective was to compare the results of two approaches, analysis of covariance (ANCOVA) and state-space modeling, to determine the factors affecting grain yield in three crop rotations [pea (Pisum sativum L.)–wheat (Triticum aestivum L.)–barley (Hordeum vulgare L.), canola (Brassica napus L.)–wheat–barley, and wheat–wheat–barley] at two sites in Saskatchewan, Canada. Crop rotations were established in adjacent 30 m x 80 m plots arranged in a randomized complete block with five replicates. Variables that were expected to affect resource availability and pest infestations in wheat (second rotation phase) or barley (third rotation phase) were measured. Each sampling point was classified according to landscape position as either a shoulder or footslope. Landscape position was considered as a cross-classified treatment along with crop rotation, and analyzed using ANCOVA procedures. State-space modeling was conducted on a single transect connecting sampling points across all of the rotations and replicates at each site. ANCOVA frequently indicated that grain yield and other measured variables differed between landscape position across all rotations, or in a specific crop rotation. For example, grain yield, soil water content, soil N availability during the growing season, and the incidence of common root rot were higher in the footslopes than the shoulders in all of the crop rotations at one of the sites. However, the landscape position effect for grain yield was never fully explained by the landscape position effects detected for the other variables (e.g., higher soil water content in the footslopes did not correspond with higher grain yields in footslope positions at both sites). State-space modeling indicated that most of the measured variables contributed to the prediction of landscape-scale variation for grain yield in the pea–wheat rotation; whereas only leaf and root disease incidences explained landscape-scale variation in the wheat–wheat rotation. The selective omission of data indicated that state-space modeling was accounting for the varied importance of the predictors across the landscape; i.e., localized response functions. The major reason that ANCOVA did not explain landscape-scale variation of grain yield across the different crop rotations may be because it was unable to account for highly localized variation. However, there is evidence from other studies that the ANCOVA approach is appropriate when the response functions explaining grain yield do not vary significantly within the study area. This situation is most likely to occur in studies with smaller experimental areas. Future research conducted at scales reflecting ‘real world’ field conditions (i.e., study units representative of producer’s

* Corresponding author. Tel.: +1-306-966-2703; fax: +1-306-966-6881.
E-mail address: knight@sask.usask.ca (J.D. Knight).

0167-1987/01/$ – see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S 0167-1987(00)00166-5
fields) should consider the use of state-space modeling or alternative statistical techniques that are designed to address and predict the complex and dynamic nature of landscape-scale processes. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Analysis of covariance; Crop rotation; Grain yield; Landscape position; Landscape-scale; Management practices; State-space modeling

1. Introduction

Once considered a nuisance or something to be minimized or controlled, landscape-scale variation is becoming an integral source of information as farming practices move toward improved landscape stewardship, like variable rate management or precision farming. Developing technologies and environmental issues have forced scientists to evaluate the significance of this background variation that is a combination of landscape-scale variation and measurement error. Furthermore, landscape-scale variation can be dependent on management practices. Stevenson and van Kessel (1996) found that grain yield was about 400 kg ha\(^{-1}\) lower in high-catchment footslopes (depressions) compared with low-catchment footslopes and shoulders in the second year of a pea–wheat and wheat–wheat rotation, but that the cause of the yield reduction was different for the different rotations. The landscape position effect was related to greater soil resource availability in the pea–wheat rotation, and higher grassy weed infestation in the wheat–wheat rotation. Beckie and Brandt (1997) found that the yield advantage of growing wheat on pea stubble rather than flax stubble ranged from 10 to 24% in the lower slope positions but was negligible in the mid- to upper-slope positions.

Effective approaches are needed to identify and understand processes causing landscape-scale variation. Traditional applications of analysis of variance (ANOVA) (e.g., Steel and Torrie, 1980) consider landscape-scale variation as ‘noise’ associated with an average or trend. Recent studies have used ANOVA in conjunction with a variable that classifies different areas of the landscape (e.g., landform complex), in factorial combination with the treatment(s), to detect and explain occurrences of landscape-scale variation (Beckie and Brandt, 1997; Stevenson and van Kessel, 1996). However, Nielsen et al. (1998) demonstrated that classical statistical approaches (e.g., treatments and ANOVA) incorrectly interpret the cause of landscape-scale patterns. Their major concern was that classical experimental designs and data analysis techniques cannot deal with localized variation (e.g., weed patches) and their associated response functions, a phenomenon that typically occurs across a relatively large area (>1 ha) of a field.

State-space modeling is a multivariate, autoregressive technique. Nielsen et al. (1998) and Wendroth et al. (1992, 1998) used state-space modeling to quantify localized variation. Their findings indicate that state-space models provide insight into the spatial patterns of crop and soil variables. Furthermore, Wendroth et al. (1992) showed that 40–50% of the variation in biological N\(_2\) fixation by alfalfa (Medicago sativa L.) across a transect was explained by regression analysis, whereas approximately 90% of the variation was explained by a state-space model.

Crop and land management studies conducted at the landscape-scale could provide an understanding of how anthropogenic activities influence productivity across a landscape. Identifying these factors, as well as identifying and eliminating those factors not taking part in the landscape-scale processes will allow optimal management of soil resources and crop yield. The objective of this study is to compare the results of analysis of covariance (ANCOVA) and state-space modeling and to discuss their ability to identify factors explaining the landscape-scale variation of crop productivity among different management (crop rotation) practices.

2. Materials and methods

2.1. Study sites

Two research sites were established in the northern grainbelt of Saskatchewan, Canada. Sites were established near St. Louis (106°45'W, 52°54'N) and Birch Hills (105°2'W, 53°3'N) in 1996. The terrain at these sites varied from hummocky to undulating. Soils at
Table 1
Soil properties (median) in 1996 at St. Louis and Birch Hills, Saskatchewan, Canada. Values are medians (±interquartile range) for 76 shoulder and 74 footslope positions at St. Louis and 75 footslope and 75 shoulder positions at Birch Hills.

<table>
<thead>
<tr>
<th>Site/landscape position</th>
<th>Bulk density&lt;sup&gt;a&lt;/sup&gt; (g cm&lt;sup&gt;-3&lt;/sup&gt;)</th>
<th>Soil water, 30–60 cm (g kg&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Soil inorganic N, 30–60 cm (kg ha&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Electrical conductivity (μS)</th>
<th>pH</th>
<th>Total C (g kg&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Total N (g kg&lt;sup&gt;-1&lt;/sup&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>St. Louis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoulders</td>
<td>1.49</td>
<td>270 ± 5</td>
<td>4.0 ± 1.7</td>
<td>16.6 ± 6.7</td>
<td>22.4 ± 19.5</td>
<td>7.67 ± 0.42</td>
<td>30.6 ± 9.9</td>
</tr>
<tr>
<td>Footslopes</td>
<td>281 ± 4</td>
<td>4.4 ± 2.7</td>
<td>15.3 ± 8.3</td>
<td>19.0 ± 15.8</td>
<td>7.58 ± 0.39</td>
<td>8.06 ± 0.42</td>
<td>31.1 ± 11.6</td>
</tr>
<tr>
<td>Birch Hills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoulders</td>
<td>1.49</td>
<td>249 ± 4</td>
<td>2.2 ± 0.9</td>
<td>14.9 ± 8.4</td>
<td>20.7 ± 11.5</td>
<td>7.58 ± 0.72</td>
<td>26.6 ± 17.4</td>
</tr>
<tr>
<td>Footslopes</td>
<td>274 ± 7</td>
<td>2.8 ± 1.3</td>
<td>13.3 ± 5.4</td>
<td>13.8 ± 9.4</td>
<td>7.30 ± 0.58</td>
<td>7.79 ± 0.77</td>
<td>37.8 ± 19.0</td>
</tr>
</tbody>
</table>

<sup>a</sup> Bulk density estimates were not different between landscape positions.
both sites are mainly Udic Haploborolls developed on a silty-textured lacustrine parent material, with occurrences of Haplaqueptic Haploborolls in the lower slope positions and Udorthentic Haploborolls in the upper slope positions. Basic soil characteristics for each site are provided in Table 1. At the St. Louis site, soils were sampled from 76 shoulder and 74 footslope positions. At the Birch Hills site, soils were sampled from each of 75 footslope and shoulder positions. A portion of each air-dried soil sample was screened using a 2 mm sieve and pH and EC measured using a soil to solution ratio of 1:2. Total C and total N were determined on finely ground soil samples using a LECO-CNS analyzer. Bulk density was determined using the core method (Blake and Hartge, 1986). Cropping history at the sites included mainly wheat, barley, and canola in the years prior to the study, with barley grown as the most recent crop.

2.2. Experimental design

Pea–wheat–barley, canola–wheat–barley, and wheat–wheat–barley rotations were established in 30 m x 80 m plots encompassing a 4 ha area at each site (Fig. 1). Plots were arranged in a randomized complete block design with five replicates. A systematic grid with 10 m spacing was superimposed across all plots at both sites. A topographic survey was conducted at each grid point and the elevation and horizontal coordinates determined. This information was used to calculate a digital elevation model that in turn was used to generate slope variables (Pennock

Fig. 1. Landscape position/sampling point maps at St. Louis and Birch Hills, Saskatchewan, Canada. The bottom two maps indicate those sampling points that were synthesized into transects for the state-space modeling approach.
et al., 1987, 1994). Slope variables were used to identify shoulder, backslope, footslope, and level landscape positions (landform complexes) across each site at a 10 m × 10 m resolution (Fig. 1). Level landform complexes were recategorized into upper and lower level landscape positions based on elevations above and below the mean elevation for each site. Upper level and backslope positions were recategorized to shoulders, and lower levels were recategorized to footslopes because of their similar productivity potentials and to simplify subsequent analyses.

2.3. Site preparation

2.3.1. First rotation phase

Tillage operations were performed prior to sowing in 1996. Sites were cultivated twice (St. Louis) or once (Birch Hills) with a medium-duty cultivator, followed by harrow packing. Both sites were seeded with a 12.2 m hoe air-drill on 29 May 1996. ‘Highlight’ pea, inoculated with a commercial rhizobium inoculant, ‘45A71,’ B. napus (L.) canola, and ‘Katepwa’ hard red spring wheat were sown in the first rotation phase. Seeding rates were 200 kg ha⁻¹ for pea, 5 kg ha⁻¹ for canola, and 100 kg ha⁻¹ for wheat. During seeding, urea was side-banded relative to the seed row at a rate of 55 kg N ha⁻¹ in the canola plots and 50 kg N ha⁻¹ in the wheat plots. Five kilograms of N per hectare and 11 kg P ha⁻¹ were applied as mono-ammonium phosphate in the seed row during all seeding operations at all sites. Grass and broadleaf weeds in the pea and canola plots were controlled with postemergent applications of imazethapyr/imazamox (2-[4,5-dihydro-4-methyl-4-(1-methylethyl)-5-oxo-1H-imidazol-2-yl]-5-ethyl-3-pyridinecarboxylic acid). Imazethapyr was tank-mixed with clethodim ((E,E)-(±)-2-[1-[[3-chloro-2-propenyl]oxy]jimino]propyl)-5-[2-(ethylthio)propyl]-3-hydroxy-2-cyclohexen-1-one) at St. Louis for more complete control of wild oat (Avena fatua L.) and volunteer barley. A postemergent application of tralkoxydim (2-(cyclohexen-1-one), 2-[1-(ethoxyimino)propyl]-3-hydroxy-5-(2,4,6-trimethylphenyl)-(9CI)) tank-mixed with bromoxynil (3,5-dibromo-4-hydroxybenzonitrile) + MCPA [(4-chloro-2-methylphenoxy)-acetic acid] 1:1 was used to control grassy and broadleaf weeds in wheat plots at both sites. Herbicides were applied at recommended stages and rates with a high-clearance or a tractor-drawn field sprayer.

2.3.2. Second rotation phase (wheat)

Preseeding tillage at both sites in 1997 included one pass with a tandem disc just prior to sowing. The sites were seeded on 20 May 1997, with a 10 m hoe air-drill. The seeding rate for wheat was 100 kg ha⁻¹ at all sites. Urea was side-banded relative to the seed row at a rate of 55 kg N ha⁻¹. Five kilograms of N per hectare and 11 kg P ha⁻¹ were applied as mono-ammonium phosphate in the seed row during seeding at all sites. Clodinafop-propargyl (2-propynyl-(R)-2-[4-(5-chloro-3-fluoro-2-pyridyl)oxy]-phenoxy)propionate) was applied for postemergent control wild oat.Dicamba (3,6-dichloro-2-methoxybenzoic acid) + mecoprop [(±)-2-(4-chloro-2-methylphenoxy)propionic acid] + MCPA at St. Louis and thifensulfuron (3-[[4-(methoxy-6-methyl-1,3,5-triazin-2-yl)amino]carbonyl]jimino]sulfonfyl]-2-thiophencarboxylic acid) + tribenuron (2-[[4-(methoxy-6-methyl-1,3,5-triazin-2-yl)methylamino]carbonyl]jimino]sulfonfyl]benzoic acid) tank-mixed with MCPA at Birch Hills were applied for postemergent control of broadleaf weeds. Herbicides were applied at recommended stages and rates with a high-clearance or tractor-drawn field sprayer.

2.3.3. Third rotation phase (barley)

The two sites were cultivated in the fall prior to the growing season of 1998. The St. Louis site was seeded to ‘Stander’ 6-row barley at a rate of 90 kg ha⁻¹ with a 12 m air seeder on 11 May 1998. The Birch Hills site was seeded to ‘Stein’ 2-row barley at a rate of 100 kg ha⁻¹ with a 17 m air seeder on 7 May 1998. A fertilizer blend of 60 kg N ha⁻¹, 11 kg P ha⁻¹, 3 kg S ha⁻¹, and 5 kg K ha⁻¹ was applied with the seed at St. Louis. The same fertilizer blend only containing 50 kg N ha⁻¹ plus 0.6 kg Cu ha⁻¹ was applied with the seed at Birch Hills. Harrow packing immediately followed the seeding operation. 2,4-D ((2,4-dichlorophenoxy)acetic acid) was applied at St. Louis, and fluroxypyr ((4-amino-3,5-dichloro-6-fluoro-2-pyridinyl)oxy)acetic acid) + clopyralid (3,6-dichloro-2-pyridinecarboxylic acid) + MCPA was applied at Birch Hills, with a tractor-drawn field sprayer to provide postemergent control of broadleaf weeds. Herbicides were applied at recommended stages and rates. Wild oat herbicides were not sprayed at either site because of the dry spring and the sparse wild oat infestation at the time of spraying.
2.4. $^{15}$N application

Atom% $^{15}$N excess measures the size of the plant-available N pool integrated over time. The $^{15}$N:$^{14}$N ratio of the plant tissue varies inversely with the size of the unlabeled inorganic N pool available for plant uptake during the growing season (Senaratne and Hardarson, 1988). It is assumed that if the atom% $^{15}$N excess of wheat plants grown on pea or canola stubble does not differ from the atom% $^{15}$N excess of wheat plants grown on wheat stubble, then there is no extra N contribution to a subsequent wheat crop from the pea and canola.

An approximately 8 l volume of water containing 0.5 kg N ha$^{-1}$ double-labeled 10 atom.% $^{15}$N NH$_4$NO$_3$ was uniformly applied to a 1 m$^2$ microplot at all sampling points. The $^{12}$N solution was applied just after the wheat had emerged in the second rotation phase.

2.5. Data collection and calculations

Soil water and inorganic N content, atom% $^{15}$N excess, common leaf diseases of wheat, common root rot, densities of predominant weed species, and wheat grain yield were measured at 150 sampling points in the second rotation phase (Fig. 1 and Table 2). Weed densities and barley grain yield were measured at 90 sampling points in the third rotation phase (Table 2).

One soil core was taken from the 0–30 cm increment of the soil profile at each sampling point just prior to sowing crops in the second phase of the rotations. A 20 g subsample of soil was oven dried at 105°C to determine soil water content. Another 20 g subsample was extracted with 200 ml of 2 M KCl and soil inorganic N (NO$_3^-$ and NH$_4^+$ combined) determined using a Technicon AutoAnalyzer II System (Labtronics, Tarrytown, NY). Soil inorganic N was calculated on a dry-weight basis and adjusted for the average bulk density at each site.

After postemergent herbicide applications, individual weed species were counted in four 0.25 m$^2$ quadrats at each sampling point in the second and third rotation phases at the sites. Wild oat and Canada thistle [Cirsium arvense (L.) Scop.] were identified as the most abundant species at both sites and were considered for further analysis. Common root rot (Cochliobolus sativus Ito and Kurib.) incidence also was assessed between anthesis and maturity. Approximately 20 plants were excavated in a diagonal across each sampling point. Common root rot lesions on the subcrown internode were rated using a 0–4 scale (0 = 0%, 1 = 1–25%, 2 = 26–50%, 3 = 51–75%, and 4 = 76–100%). A mean of the 20 subsamples was calculated to obtain a single rating for each sample. Leaf disease [tan spot: Pyrenophora triticirepens (Died.) Drechs., and septoria leaf blotch: Septoria avenae Frank f. sp. triticea, Johns., Septoria tritici Rob. in Desm., and Septoria nodorum (Berk.) Berk.] lesions on the flag and upper leaves were rated near anthesis using a 0–11 scale (McFadden, 1991).

A 1 m$^2$ area was harvested at each sampling point in the second and third phases of the rotations at a time when producers would normally begin harvest operations. Samples were dried at 40°C in a forced air drier, threshed, and weighed to determine grain yields.

The aboveground portion of three or four wheat plants within the $^{15}$N-enriched microplots was
harvested at the same time yield samples were collected in the second rotation phase. The samples were dried and threshed by hand. The residue and grain were ground separately in a cyclone mill with a 0.4 mm screen. Residue samples were ground further in a rotating ball-bearing mill. The percentage of N and atom% $^{15}$N content of 1.00 ± 0.10 mg samples of ground plant tissue were determined using a continuous flow, isotope ratio mass spectrometer interfaced with a RoboPrep Sample Converter (Europa Scientific, Crewe, UK). The working standard was $^{15}$N-enriched pea residue (atom% $^{15}$N excess = 0.2322; S.D. = 0.0019) and pea grain (atom% $^{15}$N excess = 0.1455; S.D. = 0.0029). A representative background atom% $^{15}$N was subtracted from each measurement to give atom% $^{15}$N excess. The results for $^{15}$N-labeled grain are not presented because the atom% $^{15}$N excess for residue and grain were correlated ($P < 0.05$).

2.6. Statistical analysis

2.6.1. Analysis of variance

The influence of crop rotation and landscape position on grain yield was assessed using ANOVA. Data from the second and third rotation phases were analyzed separately using the PROC MIXED procedure with block as a random effect, and crop rotation and landscape position as fixed effects (Littel et al., 1996). Contrasts were used to make specific comparisons among the crop rotations and related interactions between crop rotation and landscape position. Weed density data were transformed (natural logarithm) prior to all analyses to reduce the influence of non-normal distributions. Treatment effects were declared significant at $P = 0.05$. A separate analysis was performed for grain yield with the same ANOVA model. Results from this analysis were considered significant when: (i) the covariable(s) was significant at $P = 0.05$; and (ii) a previously significant treatment effect for the univariate ANOVA was no longer significant at $P = 0.05$ for the ANCOVA. A significant result implied that the landscape position effect for the covariables explained the variation associated with a particular treatment effect on grain yield.

2.6.2. State-space modeling

Spatial association between crop yield and underlying variables across the crop rotations at each site was assessed with state-space modeling. State-space modeling is a multivariate autoregressive technique adopted from applied time series analysis (Shumway, 1988) and can be used to quantify the spatial coincidence of a set of variables in agricultural systems (Morkoc et al., 1985; Wendroth et al., 1992, 1999).

The state-space model consists of two major equations. The first equation

$$Z_i = Z \Phi_{i-1} + \omega_i$$

where the transition matrix $\Phi$ (a vector or set of variables) defines the state of a system at a given location $Z_i$, by the state of the system at a previous location $Z_{i-1}$ with the error (unpredicted) variation defined by $\omega_i$. The transition matrix consists of autoregression coefficients (transition matrix coefficients) for each of the variables in the state-space model. The magnitude of the coefficients reflects the importance of each variable in their ability to define the state of the system. The state of a system $Y_i$ reflects the true state of the system only indirectly. Therefore, the state equation $Z_i$ is embedded in an observation equation:

$$Y_i = M_i Z_i + v_i$$

where the true state is related to the observation through an observation matrix $M_i$ with an error $v_i$. Both error vectors $\omega_i$ and $v_i$ are assumed to be zero, uncorrelated, and independent of each other. The initial state vector $Z_0$ is assumed to have a mean vector $\mu$ and $p \times p$ covariance matrix $\Sigma$ (Shumway and Stoffer, 1982). The transition matrix coefficients and state covariance matrix estimates are optimized with a Kalman filtering algorithm/iteration procedure (Kalman, 1960). When a new observation becomes available, the prediction is updated depending on the variance terms in the state and observation equations.
The iterative updating continues until a convergence criterion of 0.005 is reached. In summary, the state equation makes a prediction for a given observation based on the state at the previous observation and a given set of coefficients (state equation), compares the prediction and measured observation (observation equation), and updates each prediction (Kalman filter).

State-space modeling integrates the influence of local effects because the meaning and impact of the transition matrix coefficients for the state vector can change across space or time. Other autoregression techniques, geostatistical analysis, and multiple regression use a single set of relationships among the variables to explain variation across a data series and require that the mean:variance ratio remain constant across space. Variation across fields typically tends to be localized, with weeds being one of the better examples of a patchy phenomenon. Therefore, state-space modeling, compared with other more commonly used analyses, can accurately and precisely predict landscape-scale variation. Other researchers have observed that state-space modeling can be an effective research tool to explain landscape-scale variation in agricultural systems (Morkoc et al., 1985; Wendroth et al., 1992, 1999).

Data were stretched into a transect perpendicular to the direction of the plots (Fig. 1) prior to the analyses. The distance between sampling points of adjacent plots was considered to be 10 m to allow for equally spaced sampling points along the transect. To simplify the analysis, only 90 sampling points from the total 150 sampling points were used for data collection in the second rotation phase and hence, the total length of the transect was 900 m. Also, only the most abundant weed species were considered for state-space modeling; wild oat at St. Louis, and Canada thistle at Birch Hills. Data could not be collected because of flooding at some of the sampling points in the second rotation phase, especially at Birch Hills. Missing data were estimated as the average of adjacent sampling points to ensure optimal state-space estimations. The data were normalized prior to the analysis in order to remove differences in scale among the variables as follows:

\[ X_i = \frac{x_i - (\overline{x} - 2\sigma)}{4\sigma} \]

\( X_i \) is the transformed value of observation \( x_i \), \( \overline{x} \) the mean of the observations, and \( \sigma \) the population standard deviation. The following output and summary were obtained from each state-space analysis. Transition matrix coefficients are equivalent to regression coefficients and indicate how important the different covariables are for predicting grain yield. Fiducial limits of uncertainty were calculated from the S.E. for each prediction. Measured data points outside the fiducial limits indicate those instances where the model did not provide a satisfactory prediction. Coefficients, predictions, and state covariance matrix estimates were monitored when grain yield data were omitted for the pea–wheat or wheat–wheat crop rotations, and every fourth (25% of data), and every second (50% of data) sampling point. Transition matrix coefficients and/or covariance matrix estimates that were affected by data omission indicate those instances where the updating resulting from the Kalman filter was important. Transition matrix coefficients and/or covariance matrix estimates that were affected by data omission from two different crop rotations indicated those instances where management practices affected landscape-scale processes influencing grain yield; i.e., landscape-scale variation was specific to a particular crop rotation.

3. Results

3.1. Differences between landscape positions

Grain yield was higher in the footslopes of all crop rotations at St. Louis in the second rotation phases (Fig. 2). The landscape position effect corresponded with higher soil water contents prior to sowing, a higher incidence of common root rot, and lower atom% \(^{15}\)N excess in the footslopes compared with the shoulders of the pea–wheat and wheat–wheat rotations. ANCOVA did not confirm that the greater soil resource availability in the footslopes was responsible for an increase in grain yield (Table 3). The ANCOVA did not include root rot because it was not logical to explain a yield increase with higher disease incidence. Furthermore, because common root rot ratings overall were very low, the landscape position effect on common root rot was probably not agronomically important (Bailey and Duczek, 1996). The
The landscape position effect for grain yield in the third rotation phase at St. Louis tended to be greatest in the pea–wheat–barley rotation (\( P = 0.087 \) for the landscape position by crop rotation interaction) and occurred despite a significant increase in wild oat density in the footslope positions (Fig. 3).

Grain yield was lower in the footslopes of the wheat–wheat rotation at Birch Hills (Fig. 2). The landscape position effect corresponded with higher soil water content prior to sowing and higher wild oat density in the footslopes compared with the shoulders. ANCOVA confirmed that the higher wild oat density in the footslopes was responsible for part of the lower grain yields in the footslopes (Table 3). Soil water content was not included in the ANCOVA because it was not logical to explain a yield decrease in those areas with greater water availability, considering the relatively dry conditions at the Birch Hills site. It is possible that increased soil water content in the footslopes may have contributed to the greater wild oat density, which in turn was related to the lower grain yield in the footslopes. There was inconclusive statistical evidence (\( P = 0.12 \)) that grain yield was lower (15%) in the footslopes of the wheat–wheat–barley rotation but not in the other crop rotations in the third rotation phase (Fig. 3). Higher wild oat density also occurred in the footslopes of the wheat–wheat–barley rotation but not in the other crop rotations.

Covariance estimates indicated that most of the variation across the sites was associated with the residual component of the experimental design (Table 4). The covariance estimate for blocks was small, and the size of residual covariance estimates for the data analyzed as a completely randomized rather than randomized complete block design were similar (results not shown), indicating that blocking did not account for a major part of the variation (Steel and Torrie, 1980). Furthermore, the magnitude of the covariance estimates for the exponential geostatistical function showed that there was little spatial structure within the residual covariance. These diagnostic analyses indicate that the majority of information remained unexplained by the experimental design and the associated ANCOVA.

### 3.2. State-space modeling

Variation along the transects at the two sites changed considerably over short distances (e.g.,
Table 3
ANOVA for wheat grain yield using the response of other relevant variables for data collected during the second rotation phase (1997) at St. Louis and Birch Hills, Saskatchewan, Canada

<table>
<thead>
<tr>
<th>Covariables/source</th>
<th>St. Louis</th>
<th>Birch Hills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariables</td>
<td>ANOVA</td>
<td>ANOVA with</td>
</tr>
<tr>
<td></td>
<td>(P-value)</td>
<td>covariables</td>
</tr>
<tr>
<td>Soil water</td>
<td>0.122</td>
<td></td>
</tr>
<tr>
<td>Soil inorganic N</td>
<td>0.503</td>
<td></td>
</tr>
<tr>
<td>Atom% 15N excess</td>
<td>0.092</td>
<td></td>
</tr>
<tr>
<td>Common root rot</td>
<td>0.035</td>
<td></td>
</tr>
<tr>
<td>Wild oat</td>
<td></td>
<td>0.168</td>
</tr>
<tr>
<td>Source</td>
<td>Landscape position (L)</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>L × pea–wheat and canola–wheat vs. wheat–wheat</td>
<td>0.285</td>
</tr>
<tr>
<td></td>
<td>L × pea–wheat vs. canola–wheat</td>
<td>0.349</td>
</tr>
</tbody>
</table>

* Only those variables that were identified as significant were included as covariables.

Fig. 3. Average landscape effect in three crop rotations for the third rotation phase (barley) at St. Louis and Birch Hills, Saskatchewan, Canada. Bars represent S.E. for each mean.
Table 4
Covariance estimates from different components of the experimental design for grain yield in two rotation phases in 1997 and 1998 at St. Louis and Birch Hills, Saskatchewan, Canada

<table>
<thead>
<tr>
<th>Rotation phase/site</th>
<th>Blocks</th>
<th>Spatial</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Louis</td>
<td>0</td>
<td>6</td>
<td>161125</td>
</tr>
<tr>
<td>Birch Hills</td>
<td>12724</td>
<td>7</td>
<td>132465</td>
</tr>
<tr>
<td>Third</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Louis</td>
<td>5831</td>
<td>0</td>
<td>286623</td>
</tr>
<tr>
<td>Birch Hills</td>
<td>0</td>
<td>10</td>
<td>188581</td>
</tr>
</tbody>
</table>

>1000 kg ha\(^{-1}\) change in grain yield between sampling points 50 m apart (Figs. 4 and 5). It also was apparent that relationships between grain yield, as the primary variable of interest, and the covariables were complex and localized. There were no large-scale (global) trends for responses in the crop rotations at St. Louis and Birch Hills. Generally, responses across the transects for weed density were the most variable, responses for grain yield, soil inorganic N prior to sowing, atom% \(^{15}\)N excess, and diseases were moderately variable, and responses for elevation and soil water content prior to sowing were the least variable.

Locally high/low responses for wheat grain yield occurred along the transects but there were no consistent set of associated local response functions. For example, at St. Louis at 225 m, reduced atom% \(^{15}\)N excess and increased soil water content corresponded with higher grain yield (Fig. 4). Locally, a low incidence of leaf diseases corresponded with high grain yields at 575 m. At Birch Hills, low incidence of common root rot coincided with higher grain yield at 700 m, and a high density of Canada thistle coincided with lower grain yield at 675 m (Fig. 5).

An analysis of the complete data set at St. Louis indicated that atom% \(^{15}\)N excess, leaf diseases, and wild oat density were particularly important predictors of the landscape-scale variation for wheat grain yield (Fig. 6). Deleting every other data point reduced the frequency at which yield predictions were updated by the Kalman filter. The importance of updating on grain yield predictability should be reflected through larger (either positive or negative) transition matrix coefficients and state covariance matrix estimates for the covariables used to explain grain yield when portions of the data are selectively omitted. Omitting data collected from the pea–wheat rotation put greater dependence on all covariables as predictors of landscape-scale variation for grain yield (Fig. 6). State covariance matrix estimates decreased for leaf diseases, and increased for atom% \(^{15}\)N excess and wild oat density, when data from the pea–wheat rotation were omitted (Table 5). Omitting the data collected from the wheat–wheat rotation decreased the dependence on all covariables except wild oat density (Fig. 6). State covariance matrix estimates, in particular, increased for leaf diseases, and decreased for common root rot and wild oat density when data from the wheat–wheat rotation were omitted (Table 5). Therefore, the factors explaining landscape-scale variation of grain yield contrasted between the various rotations.

Table 5
State covariance matrix estimates for state-space models for data collected during the second rotation phase (1997) at St. Louis and Birch Hills, Saskatchewan, Canada

<table>
<thead>
<tr>
<th>Site/data omission</th>
<th>Soil water</th>
<th>Atom% (^{15})N excess</th>
<th>Leaf diseases</th>
<th>Common root rot</th>
<th>Weeds(^a)</th>
<th>Grain yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>St. Louis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>-0.0019</td>
<td>-0.0071</td>
<td>&lt;0.0001</td>
<td>-0.0040</td>
<td>0.0066</td>
<td>0.0187</td>
</tr>
<tr>
<td>Pea–wheat omitted</td>
<td>-0.0013</td>
<td>-0.0123</td>
<td>&lt;0.0001</td>
<td>-0.0015</td>
<td>0.0122</td>
<td>0.0184</td>
</tr>
<tr>
<td>Wheat–wheat omitted</td>
<td>0.0015</td>
<td>-0.0099</td>
<td>0.0075</td>
<td>-0.0001</td>
<td>-0.0012</td>
<td>0.0224</td>
</tr>
<tr>
<td>25% omitted</td>
<td>0.0007</td>
<td>-0.0070</td>
<td>-0.0055</td>
<td>-0.0025</td>
<td>0.0055</td>
<td>0.0124</td>
</tr>
<tr>
<td>50% omitted</td>
<td>0.0043</td>
<td>-0.0027</td>
<td>-0.0121</td>
<td>-0.0036</td>
<td>-0.0028</td>
<td>0.0238</td>
</tr>
<tr>
<td>Birch Hills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>-0.0071</td>
<td>0.0012</td>
<td>-0.0195</td>
<td>-0.0012</td>
<td>-0.0147</td>
<td>0.0400</td>
</tr>
<tr>
<td>Pea–wheat omitted</td>
<td>0.0040</td>
<td>-0.0024</td>
<td>-0.0240</td>
<td>0.0024</td>
<td>-0.0076</td>
<td>0.0196</td>
</tr>
<tr>
<td>Wheat–wheat omitted</td>
<td>0.0012</td>
<td>0.0034</td>
<td>-0.0192</td>
<td>0.0029</td>
<td>-0.0112</td>
<td>0.0125</td>
</tr>
<tr>
<td>25% omitted</td>
<td>-0.0070</td>
<td>0.0026</td>
<td>-0.0180</td>
<td>-0.0040</td>
<td>-0.0162</td>
<td>0.0308</td>
</tr>
<tr>
<td>50% omitted</td>
<td>-0.0061</td>
<td>0.0002</td>
<td>-0.0164</td>
<td>-0.0109</td>
<td>-0.0161</td>
<td>0.0310</td>
</tr>
</tbody>
</table>

\(^a\) The predominant weed species at St. Louis was wild oat and at Birch Hills was Canada thistle.
Fig. 4. Responses at 90 sampling points during the second rotation phase (wheat) of three rotations for the transect at St. Louis, Saskatchewan, Canada. Responses for grain yield are presented with each of the other variables as a dashed line with triangle symbols. Sampling points classified as shoulders are designated with a ‘1’ along the bottom of the chart and footslopes are designated with a ‘3’ along the top of the chart.
Fig. 5. Responses at 90 sampling points during the second rotation phase (wheat) of three rotations for the transect at Birch Hills, Saskatchewan, Canada. Responses for grain yield are presented with each of the other variables as a dashed line with triangle symbols. Sampling points classified as shoulders are designated with a ‘1’ along the bottom of the chart and footslopes are designated with a ‘3’ along the top of the chart.
Fig. 6. State-space models for 90 sampling points during the second rotation phase (wheat) of three rotations for the transect at St. Louis, Saskatchewan, Canada. Models were generated with different types of data omission. The solid circles represent normalized grain yield that was measured, empty circles represent those sampling points that were selectively omitted, and the thin lines represent the upper and lower 95% fiducial limits. Transition matrix coefficients for all variables in the model are presented immediately above each chart.
Fig. 7. State-space models for 90 sampling points during the second rotation phase (wheat) of three rotations for the transect at Birch Hills, Saskatchewan, Canada. Models were generated with different types of data omission. The solid circles represent normalized grain yield that was measured, empty circles represent those sampling points that were selectively omitted, and the thin lines represent the upper and lower 95% fiducial limits. Transition matrix coefficients for all variables in the model are presented immediately above each chart.
crop rotations. Omitting data from every fourth or second sampling point tended to decrease the size of
the transition matrix coefficients (Fig. 6). Selective
data omissions increased the size of state covariance
matrix estimates for leaf diseases and had varied but
relatively smaller effects on estimates for the other
variables (Table 5). The presence of points outside the
fiducial limits, following the omission of every fourth or
second sampling point, also suggests that reducing
the frequency of updating resulted in less reliable
predictions of grain yield at St. Louis.

An analysis of the complete data set from Birch
Hills showed that atom% 15N excess, and to a lesser
extent soil water content and common root rot, were
important predictors of the landscape-scale variation
for wheat grain yield (Fig. 7). Omitting data collected
from the pea–wheat rotation reduced the size of the
transition matrix coefficient for common root rot and
increased the size of the coefficient for the other
covariables, especially soil water content (Fig. 7).
Omitting data collected from the wheat–wheat rota-
tion increased the dependence on soil water content,
Canada thistle density, and especially common root rot
(Fig. 7). The state covariance matrix estimates
decreased for soil water content and increased for
atom% 15N excess and common root rot when data
from either crop rotation was omitted (Table 5). Omit-
ting data from every fourth or second sampling point
had minimal effects on the size of the transition matrix
coefficients (Fig. 7). However, these data omissions
affected the state covariance matrix estimates for
atom% 15N excess and common root rot (Table 5).
The greater reliance and increased covariance esti-
mates for certain covariables following selective data
omission suggests that reducing the frequency of
updating for grain yield predictions also was important
at the Birch Hills site.

4. Discussion

4.1. Comparison of methods

ANOVA frequently detected significant landscape
position effects and landscape position by crop rota-
tion interactions for grain yield and the other variables.
The presence of significant landscape position effects,
however, was not accompanied by explanations of the
landscape position effects for grain yield in the dif-
ferent crop rotations. There are a number of reasons
why the ANCOVA approach should have successfully
explained the landscape-scale variation for grain yield.
Different landscape positions encompass attributes of
the landscape that exemplify changes in crop produc-
tivity across landscapes (Fiez et al., 1994; Halvorson
and Doll, 1991; Stevenson et al., 1995; Stevenson and
van Kessel, 1996). Only small changes in soil quality
(e.g., total soil C and N) and productivity potential
occurred between the shoulders and footslopes at both
sites (Table 1). Despite this similarity, there was
enough topographic relief associated with the different
landscape positions to influence the availability of soil
resources, microclimatic conditions (e.g., crop canopy
humidity), and ultimately crop production potential
(see Fig. 2). Greater success also was expected with
the ANCOVA approach because we measured the
most obvious, and probably the most important, vari-
ables determining cereal crop productivity in the
different crop rotations and across the landscape. It
is unlikely that variables that were not measured (e.g.,
soil P availability, micronutrients, etc.) would have
significantly improved our success in explaining the
landscape position effects for grain yield.

A statistically significant explanation, albeit partial,
of the landscape position effect for grain yield by a
relevant covariable(s) occurred in only one of the four
possible instances, despite the frequent detection of
significant landscape position effects. There are a
number of reasons why ANCOVA did not successfully
explain landscape-scale variation for grain yield
among the different crop rotations. The unbalanced
number of sampling points for each landscape position
(Table 2) within the crop rotation and block combina-
tions of the experimental design may have slightly
reduced the power of the analysis to explain landscape
effects for grain yield. Reduced statistical power
would occur because the S.E. of the means would
be larger for crop rotation by landscape position
combinations with fewer observations. Larger S.E.
would reduce the probability of detecting means that
differed significantly relative to those means with
more observations. It is important to consider that
this unbalance is an attribute of the landscape and
should be appropriately considered when choosing a
research approach to understand why variation occurs
across a landscape.
Localized variation and the continuously changing importance of the different explanatory variables for grain yield (response functions) across the transects are more likely reasons why ANCOVA was not able to consistently explain the landscape-scale variation. For example, greater weed infestations and lower grain yields occurred in the footslopes of only the wheat–wheat rotation at Birch Hills. However, the raw data indicated that this yield reduction probably occurred only at a small proportion of the footslope sampling points (i.e., 550–700 m, see Fig. 5). Therefore, the ANCOVA results were based on a small amount of information, despite the fact that the ANCOVA assumptions were not seriously violated.

Complementary to the problem associated with localized responses is the incidence of more than one set of response functions explaining grain yield across a landscape. The covariance associated with each response function across the landscape would not be properly accounted for with techniques such as ANCOVA that use a single response function to explain all the variation. For example, a particular patch of common root rot may have reduced grain yield in one or more of the crop rotations. An area with higher soil water and N content may have increased grain yield in an area 100 m away. A poor explanation of the unique response functions explaining grain yield across the landscape ultimately means that the residual covariance becomes inflated. Error inflation with the ANCOVA would explain why this approach was unable to explain the landscape-scale variation for grain yield even though landscape position effects on yield frequently were detected with univariate ANOVA.

Randomization and blocking are two components of an experimental design that are supposed to compensate for the occurrence of unexplainable variation and prevent the confounding of variation between adjacent blocks and plots. Randomization of treatments among the experimental units within blocks may not reduce the possibility of spurious variation, particularly variation associated with patchy pest infestations. Moreover, blocking and accounting for spatial structure in the residual covariance did not reduce the amount of unexplained variation for the ANCOVA. Other reports that blocking may not effectively explain spatial variation when variation within blocks is obvious also have been made (Gotway Crawford et al., 1997; Lin et al., 1993).

The size of the study area may influence how well the ANCOVA approach explains landscape-scale variation. A previous study conducted in a 2 ha area clearly demonstrated the factors controlling grain yield in two crop rotations (Stevenson and van Kessel, 1996). It is thought that an ANOVA approach was successful for these researchers because the factors explaining landscape-scale variation across the study site were relatively consistent. It seems logical that smaller study areas will tend to have fewer explanatory response functions. However, it is important to remember that scaling the size of the experimental area to accommodate statistical convenience, rather than reflecting a ‘real world’ situation may not further our understanding of landscape-scale variation.

In addition to the technical problems outlined, the basic concept of ANOVA may not be valid for agronomic research that is conducted at the landscape-scale. The ANOVA approach was developed to remove or control ‘spurious’ sources of variation; i.e., landscape-scale variation. Furthermore, a landscape position effect is deemed to be important only when it is detected at a chosen level of significance. In reality, landscape-scale variation is always present and its intensity may differ between fields. Technical and philosophical problems with the ANOVA approach indicate that a new paradigm is necessary to improve our understanding of landscape-scale processes as related to agronomic questions.

State-space modeling is a technique that can identify landscape-scale processes and generate reliable predictions in the presence of highly localized responses and different functional subunits. The model necessary to explain the landscape-scale variation of grain yield in the pea–wheat rotation was more complex than the model necessary to explain the landscape-scale variation for grain yield in the wheat–wheat rotation at both sites. Of particular importance was the atom% $^{15}$N excess values (indicators of total plant available N during the growing season) that explained the landscape-scale variation of grain yield in the pea–wheat rotation. It appears that environmental conditions at both sites and the low C:N ratio of pea residue increased the net N mineralization or reduced N immobilization in certain areas of the landscape. In contrast, leaf and root diseases mainly explained the landscape-scale variation of grain yield in the wheat–wheat rotation. The temporal persistence
of wheat disease patches may be responsible for the dominant effect of wheat diseases on grain yield in the wheat–wheat rotation. It is thought that the areas of the field under greater physiological stress (e.g., nutrient deficiency, drought, etc.) were more conducive to disease development leading to lower crop yields. Leaf and root diseases have been shown by others to be most prevalent when crop plants are stressed (Bailey et al., 1989).

State-space modeling also provided accurate interpretations of landscape-scale variation, especially at St. Louis. All of the observations were within the fiducial limits for the model at St. Louis, thus indicating a reliable prediction of landscape-scale variation on grain yield. About 30% of the observations were outside the fiducial limits for the various models estimated for Birch Hills. Unmeasured factors may have improved predictions of the landscape-scale variation for wheat grain yield at this site. State-space modeling does have some technical and practical advantages over other approaches as an agronomic tool to understand landscape-scale variation. Most agricultural scientists would agree that landscape-scale variation always is present. The ANOVA approach first has to detect if landscape-scale variation is present before it can provide further explanation. State-space modeling does not have this restriction, which means that it always can offer an explanation of landscape-scale variation. The updating process (Kalman filter) deals with highly localized variation better than other approaches that consider single response functions and averages. Our results from the data omission phase of the analyses indicated that localized updating by the Kalman filter improved that predictability of grain yield across the landscape. This is why the $r^2$ for state-space models is generally higher than for ANOVA and regression models (Wendroth et al., 1992). Other researchers also have found that state-space modeling was an effective technique to understand landscape-scale variation in a variety of agricultural systems (Morkoc et al., 1985; Nielsen et al., 1998; Wendroth et al., 1992, 1998, 1999). More importantly, our comparison of ANCOVA and state-space modeling has shown that alternative analytical tools and experimental designs may allow a better diagnosis of variable productivity across a landscape and garner a broader understanding of landscape-scale variation.

A couple of realities must be recognized to effectively use state-space modeling to explain landscape-scale variation. The state-space approach assigns coefficients to assess the importance of each covariable as a predictor of some primary variable, such as grain yield. Unlike ANOVA, it does not assign a $P$-value to their importance. Also, state-space modeling does not directly assess the statistical significance of treatments and their interaction with a factor(s) describing landscape-scale variation like ANOVA. Neither of these issues is a problem per se, but will require a new philosophy when dealing with the explanation of landscape-scale variation. In particular, agronomists interested in state-space modeling will have to shift emphasis away from statistical significance and traditional experimental design, and put more emphasis on innovative ways to assess landscape-scale variation among different management practices.

5. Summary and conclusions

Effective and economic implementation of variable rate applications of agricultural inputs (Nielsen et al., 1998; Sadler et al., 1998) and environmental protection strategies require an in-depth understanding of the processes that lead to variable yields across a landscape. This understanding at the process level also must include the potential impact and benefit of different management practices imposed across a field. Landscape-scale variation is a complex set of interrelated processes that vary continuously across a field and can be dependent on treatments or management practices. The challenge facing scientists to explain landscape-scale variation in agricultural fields cannot be solved completely with the current agronomic research approaches available that utilize treatments and ANOVA, geostatistics, etc. (Nielsen et al., 1998). We determined that the classical randomized complete block design in combination with an ANCOVA did not account for localized variation (e.g., weed patches) and the complex interrelationships that varied among spatial subunits across the large (4 ha) experimental sites. Furthermore, the overwhelming source of variation at this scale probably was caused by landscape variation, rather than variation associated with superimposed treatments. If this phenomenon is present in all large field-scale experiments, this would impose a
challenge to researchers who conduct agronomic research at the landscape-scale. The challenge could be surmountable, however, if the research approach used to understand landscape-scale variation is better adapted to address the processes underpinning agro ecosystems, like state-space modeling. On the other hand, state-space modeling does not deal with treatment structures and does not assess the importance of individual covariables in the model in the same way as ANOVA. We anticipate that the adoption of alternative analytical tools may further our understanding of landscape-scale variation.

Acknowledgements

The Saskatchewan Agricultural Development Fund and the Saskatchewan Pulse Crop Development Board provided funding for the study. The expertise and assistance of Dr. D. Pennock, M. Forster, G. Parry, M. Dick, B. van Melle, and B. Anderson are greatly appreciated.

References


McFadden, W., 1991. Econometric analysis of economic behavior when the dependent variable is ordered. J. Econ. 39, 381–403.


