A MANAGEMENT OPPORTUNITY INDEX FOR PRECISION AGRICULTURE

Alex. B. McBratney, Brett M. Whelan, James A. Taylor, and Matt. J. Pringle

If the opportunity indices reported here are used in scientific papers (or commercially) please refer to the following. This is good and fair scientific practice.

(1) Pringle M.J., McBratney, A.B., Whelan, B.M. and Taylor, J.A. (in press) A preliminary approach to assessing the opportunity for site-specific crop management in a field, using yield monitor data. *Journal of Agricultural Science, Cambridge*.

(This paper describes the opportunity index O_c in some detail and contains more results on the Fairfield Smith analysis.)

(2) A.B. McBratney, B.M. Whelan, J.A. Taylor, and M.J. Pringle (2000). A management opportunity index for precision agriculture. In (P.C. Robert, R.H. Rust & W.E. Larson, eds). *Proceedings of the 5th International Conference on Precision Agriculture and Other Resource Management*, July 16-19, 2000, Radisson Hotel South, Bloomington, Minnesota, USA.

For the convenience of those interested, a facsimile of

Fairfield Smith, H. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. *Journal of Agricultural Science, Cambridge* **28**: p. 1–23.

has been placed on our website http://www.usyd.edu.au/su/agric/acpa/

under ‘Publications and references’ and then click on ‘Classic PA Papers’. This is one of the finest, and most important early, papers in precision agriculture.

July 14th 2000
Paper presented at the 5th International Conference on Precision Agriculture and Other Resource Management, July 16-19, 2000, Radisson Hotel South, Bloomington, Minnesota, USA

A MANAGEMENT OPPORTUNITY INDEX FOR PRECISION AGRICULTURE

A.B. McBratney, B.M. Whelan, J.A. Taylor, and M.J. Pringle

Australian Centre for Precision Agriculture
The University of Sydney
New South Wales
Australia

Phone *+61 2 9351 3214*
Fax *+61 2 9351 3706*
Email Alex.McBratney@acss.usyd.edu.au
http://www.usyd.edu.au/su/agric/acpa/people/AlexMcB.htm
http://www.usyd.edu.au/su/agric/acpa/

ABSTRACT

Many farmers are employing emerging technologies to characterise the variation in their production systems. The most common of these technologies are real-time yield sensors. Farmers, however, are often left wondering how the subsequent yield maps can be used to justify a change to a Precision Agriculture management philosophy. The Opportunity Index is an attempt to provide a pragmatic solution to this problem.

The Opportunity Index for Site-Specific Crop Management (SSCM) is conditional on three components: i) the magnitude of variation (CV_a) present in a yield map, relative to a given threshold; ii) the spatial structure (S) of yield variation, relative to the minimum area within which variable-rate controllers can reliably operate; and, iii) the economic and environmental benefit (E) of SCCM relative to uniform management.

Methods for assessing the magnitude and spatial structure of variation in yield maps are proposed. These methods are then incorporated into an Opportunity Index to predict the potential for SSCM. Results are encouraging however further research especially into the economic and environmental impact of SSCM is required before the Opportunity Index can be considered complete.

Keywords: Precision Agriculture, Opportunity Index, uniformity trials, yield variation

INTRODUCTION

From a farmer's perspective, a barrier to adoption of SSCM is deciding whether or not a crop displays enough variation, both in terms of magnitude and spatial structure to justify the cost of a shift from traditional (uniform) to site-specific (differential) management. A farmer's database of yield maps should provide the most significant clue towards the opportunity for SSCM. However while methods for quantifying yield variation do exist (discussed below), a way of quantifying the opportunity for SSCM has yet to be defined. This paper presents a first attempt.

Co-efficient of Variation

Perhaps the easiest and most common method currently employed is a 'Co-efficient of Variation' (CV) analysis. The relative magnitude of yield variation *could* be found by comparing CVs to a median value, however, we disagree with the use of a standard CV in this situation. Firstly, the CV is non-spatial and therefore potentially misleading when dealing with different sized areas (as illustrated by Fairfield Smith's work). Larger fields will, on average, have larger CVs for the same crop. Secondly, the CV tells nothing of the difference between autocorrelated yield variation (which is manageable), and uncorrelated ('nugget') variation (which is *not* manageable). The CV is therefore undesirable, and a better method of describing the magnitude of yield variation is needed.

Fairfield Smith's (1938) empirical law of yield heterogeneity

Secondly, the search for management opportunity through crop variation pointed to the work of Fairfield Smith (1938).

Fairfield Smith's empirical law of yield heterogeneity was derived from many uniformity trials. A uniformity trial is simply a field (or part thereof) treated with blanket applications of all agronomic inputs and subsequently harvested in small plots as if there were an experiment over the area (Mercer and Hall, 1911). These uniformity trials served as investigations into the resolution at which to perform effective agronomic management. Extending this concept into modern times, any uniformly managed field which is harvested with yield monitoring technology can be seen as a uniformity trial from which the opportunity for SSCM can be investigated. Fairfield Smith found that as the logarithm of area increased the logarithm of yield variation per unit area decreased linearly. The gradient of this relationship (*b*') was used as a heterogeneity coefficient that applied across all areas: the lower the absolute value of the gradient, the more heterogeneous the crop. It can also be thought of as relating to a fractal dimension (McBratney et al, 1997). It was reasoned that once *b*' was established for a field it could be used as an opportunity function for deciding future experimental plot sizes.

While useful for ranking crop yield variation, Fairfield Smith's methodology has some shortcomings if used as a measure of the potential for SSCM. The aggregation of individual 'plots' in a yield map is cumbersome and inefficient when applied to dense yield data. Furthermore, Fairfield

Smith's empirical law cannot be expected to fully describe the variation present in crop yield because *b*' only relates the rate of change of variation with area. No consideration is given to either the magnitude of yield variation or the economic/environmental impacts. Thus on it own Fairfield Smith's empirical law is not ideal in describing the opportunity for SSCM.

The Variogram

Since the rise of geostatistics within the environmental sciences, the variogram has become another popular method of describing yield variation (*e.g.*, Perrier and Wilding, 1986; Mulla, 1993; Lark *et al.* 1999), mainly because it shows how variation *changes* (usually) through space. Variograms model variance as a function of separation distance between pairs of points. Pairs of points that have a greater lag should generally have a greater semivariance than those that are closer together. The reader is referred to Webster and Oliver (1990) and Isaaks and Srivastava (1989) for detailed explanations of the concept.

There are two primary problems with the use of variograms in explaining yield variation. As with the Fairfield Smith analysis, variograms do not give any indication of the magnitude of variation in relation to the mean or the economic/environmental impact. Furthermore variograms only represent distances not areas; thus they may not necessarily reflect the 'areal' structure of the variation.

AIMS

This paper aims to improve on the shortcomings of Fairfield Smith's method, the CV and the variogram by developing a SSCM Opportunity Index (O_c) based on yield monitor data. We propose that this O_c must account for three parameters of the production system:

- the *magnitude* of yield variation relative to some threshold;
- the area within which yield variation is autocorrelated (i.e. the *spatial structure* of variation) relative to the minimum area within which variable-rate controllers effectively operate; and,
- the *economic and environmental* benefit of SSCM relative to uniform management.

METHODS

Data Preparation

To establish a SSCM Opportunity Index, yield monitor data were gathered for 5 types of crop grown in Australia: wheat, grapes, cotton, lupins, and sorghum. Data from 20 harvests were recorded for 16 fields in the period 1995–1999 (some fields are represented more than once). All crops were managed using the traditional, uniform approach to ground preparation, sowing rates, and fertiliser and pesticide applications.

As much as possible, output from the various yield monitors has been trimmed of doubtful data, *e.g.*, distribution and spatial outliers, and crop headlands. Although dependent on the crop and yield monitor, generally data outside ± 3 standard deviations from the mean yield were regarded as distributional outliers and eliminated. Spatial outliers (arising from the loss of differential correction signal to the GPS) and crop headlands (where the harvester changes direction, leaving an undesirable artifact in the yield map) are more problematic and, as such, usually removed according to one's enthusiasm for the task. This illustrates the need for the development of automated yield correction procedures (*e.g.*, those presented by Blackmore and Moore, 1999; Lars and Antje, 2000)

Definition of Components

It is proposed that the opportunity for SSCM will be a function of i) the magnitude of variation present in a yield map, relative to a certain threshold; ii) the spatial structure of yield variation, relative to the minimum area within which variable-rate controllers can reliably operate; and, iii) the economic and environmental benefit of SCCM relative to uniform management. The following section gives a detailed description of how these components have been quantified.

Magnitude (*M*)

Obviously one of the main constraints to SSCM is the magnitude of the variation within the field. If variation ranges from 2.4 to 2.6 Mg/ha, with a mean of 2.5, there is little opportunity for differential management (unless the crop is of very high value). If however another field with a mean of 2.5 Mg/ha has a variation in yield from 0.5 to 4.5 Mg/ha there would seem to be a strong case for SSCM. A large magnitude of yield variation should allow greater differentiation between input applications, hence greater economic and environmental benefits in comparison to uniform management.

The method for quantifying this magnitude of variation presented here is an 'Areal coefficient of variation' (CV_a). This is a method of standardising the previously non-spatial CV to an area and is based on the double integral of the yield variogram. In this case, because we are only interested in autocorrelated variation, the C_0 parameter was excluded from the integration. The CV_a procedure is outlined here.

Variograms were made of the raw yield of each field and fitted, weighted by m at each lag (McBratney and Webster, 1986), exponential, spherical, double exponential, double spherical and power models. If there was a trend in the variogram (*i.e.*, no obvious sill), the maximum range was constrained to 1000 m, which thereby forced a sill upon the variogram. The fit of the models was assessed using the Akaike Information Criterion (AIC) (after Webster and McBratney, 1989). The parameters from the model with the best fit (lowest AIC) were (numerically) double-integrated (minus the C_0 parameter) to the standardising area (V). This area was selected as 1000ha and considered the upper limit of field size. The numerical definition of the double integration is (after Journel and Huijbregts, 1978: Goovaerts, 1997):

$$\bar{\gamma}(V) = \approx \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\gamma(\mathbf{x}_i - \mathbf{x}_j) - C_0'\right], \tag{1}$$

where: $\bar{\gamma}(V)$ = average autocorrelated yield variation within the block of size V

N = number of points that discretise V;
$\mathbf{x}_i$ = a discrete point in V; and,
$\mathbf{x}_j$ = any other discrete point in V.
C0 = nugget variance of yield variogram

The square root of $\bar{\gamma}(V)$ was then divided by the field's mean yield and multiplied by 100 to obtain the CV$_a$:

$$\mathrm{CV}_a = \left(\frac{\sqrt{\bar{\gamma}(V)}}{\bar{Y}}\right)100, \tag{2}$$

where: $\bar{Y}$ = mean yield

The 50% quantile (q_{50}) or median CV$_a$ of all the fields was found, and used as the quantity against which to compare the magnitude of autocorrelated yield variation. Therefore the magnitude of yield variation can be expressed as:

$$\mathrm{M} = \left(\frac{\mathrm{CV_a}}{q_{50}\left(\mathrm{CV_a}\right)}\right) \tag{3}$$

In Equation 3, the division of CV$_a$ by its median effectively states that the opportunity for SSCM will be increased if a crop is more variable than what is usually observed. Using q_{50}(CV$_a$) as a value against which to make comparisons is the 'best guess', given present knowledge. It is assumed that 50% of the fields in the study will display enough variation for SSCM. With more experimentation into within-field variability, site-specific input response and experience in variable-rate technology, q_{50}(CV$_a$) will be replaced by q_α(CV$_a$), a minimum CV$_a$. This q_α(CV$_a$) will define the magnitude of variation below which uniform treatment is advisable

Spatial structure (*D*)

The second consideration when evaluating the potential for SSCM is the spatial distribution of yield variation. A strong spatial structure is desirable because variable-rate controllers, which physically implement SSCM, operate at maximum efficiency when proposed application patterns are smooth and broad. While trended yield maps may be highly desirable for SSCM, problems do arise when trying to analysis such data. The manifestation of trend in a yield map's variogram will imply that the average autocorrelation area of yield is infinite, which will lead to extremely large (and potentially unrealistic) opportunities. To reduce this effect a trend surface was fitted to the data to calculate the average area within which yield was autocorrelated and the resultant residuals used for analysis. Although dependent on the size of the field, yield monitor data is usually so abundant that a reasonably complex trend model can be afforded; hence a fourth-order model was used here.

$$Y(E,N)=\begin{pmatrix}\text{Int.}+E+N+E^2+N^2+EN+E^3+N^3+E^2N+\\ EN^2+E^4+N^4+E^3N+E^2N^2+EN^3\end{pmatrix}+\varepsilon, \qquad (4)$$

where, E = Easting coordinates of yield (with minimum subtracted to prevent numerical overflow);
N = Northing coordinates of yield (again, with the minimum subtracted);
$Y(E,N)$ = yield as a function of its Eastings and Northings;
Int. = intercept of regression;
ε = error term (residuals).

Empirical variograms were made of the trend surface residuals, and fitted with the four *bounded* theoretical models, exponential, spherical, double exponential and double spherical. The best-fitting model was again found by the AIC. This model of spatial variation was then used to find the 'areal scale' in hectares of the yield residuals (J_a). Russo and Bresler (1981) employed this 'integral scale' concept to determine the spatial dependence of soil hydraulic properties. We have adapted Russo and Bresler's (1981) idea to approximate the average area within which the residuals of a yield trend-surface are autocorrelated:

$$J_a \approx \frac{\left\{2\int_0^{\infty}\left(1-\frac{\gamma(h)}{(C_0+C_1+C_2)}\right)hdh\right\}}{10000}, \qquad (5)$$

where: $\gamma(h)$ = theoretical variogram of yield residuals;
C_0, C_1, and C_2 = parameters of the residual theoretical variogram (if the best fitting model was not a 'double', C_2 was equal to zero);
A divisor of 10000 is used to standardise J_a to a hectare.

Equation 5 converts the best-fitting residual variogram model into an equivalent correlogram. This procedure requires that the variogram have a sill (hence the use of residuals from the trend surface and theoretical models with finite sills).

The proportion of total yield variance explained by the quartic trend-surface (P_t) is calculated. Because a trend-surface is theoretically autocorrelated to an infinite area, a limit must be employed; this was chosen as the area of each field (A). Multiplying P_t by A gives the contribution of the trend surface to the average area within which yield is autocorrelated. Multiplying J_a by (1- P_t) provides the contribution of the residuals. Adding these two terms together produces the average area within which yield is autocorrelated (S)

$$S = (P_t A) + (1 - P_t) J_a. \tag{6}$$

Now let s be an estimate of the minimum area (in hectares) within which variable-rate controllers can reliably operate. It is calculated as:

$$s = \frac{(\beta \nu \tau)}{10000}, \tag{7}$$

where: β = width of application swath (m);
ν = speed of vehicle (m/s);
τ = time required to alter application rate (s).
A divisor of 10000 is used to standardise s to a hectare.

Values of these parameters are given in Table 1 and are based on personal experience with variable-rate applicators. It was necessary to distinguish between grapes and the four other crops used in this study because viticulture operates within much smaller areas than broadacre cropping.

Table 1. Parameter values for the determination of *s*.

Parameter	Grapes	Other crops
β (m)	6	20
ν (m/s)	3	6
τ (s)	3	3

The contribution of the spatial structure of yield variation to the potential for SSCM can therefore be calculated as,

$$D = \frac{S}{s}. \tag{8}$$

Equation 8 effectively states that the opportunity for SSCM will be increased when farm machinery can operate within the average area within which yield is autocorrelated; if this is not the case then SSCM is hardly feasible.

Economic/environmental benefit (*E*)

At present, little is known about the nature of parameter E, and it has therefore been assumed constant (= 1) in this study. Future studies into the opportunity for SSCM will benefit from knowledge of E but it is a topic that requires further research. Some of the factors that E must consider will be short- and long-term economic goals, the on-farm and off-farm environmental impact of management practices, government legislation and a changing consumer preference.

By combining the three parameters and applying a square root function to remove skewdness in the data, (logarithmic transforms were tried but proved too powerful), an interim continuous Opportunity Index is produced:

$$\mathrm{O_c} = \sqrt{\left(\frac{CV_a}{q_{50}(CV_a)}\right)\left(\frac{S}{s}\right)E} = \sqrt{M \bullet D \bullet E} \tag{9}$$

RESULTS AND DISCUSSION

Descriptive details of the 20 yield-monitored fields are given in Table 2. The CV$_a$ and J_a values of the 20 fields are shown in Table 3. Also shown are the proportions of total yield variation that is contributed by the quartic trend-surface (P_t), the average area within which yield is autocorrelated (S), the limitation of variable-rate technology (s), the SSCM Opportunity Index (O$_c$) and finally the Fairfield Smith b' value. Fields have been sorted in order of decreasing O$_c$.

The values of O$_c$ in Table 3 range from 2.8 to 47.2, with the median equal to 19.3. A range of parameter combinations for CV$_a$, q_{50}[CV$_a$] J_a, P_t, and A, with resultant O$_c$, were simulated in an 8^6 factorial arrangement (results not shown) to determine the probable range of O$_c$ values. The median O$_c$ of this factorial trial was 16.6, with 90% of the distribution being less than 95; it takes an extraordinary combination of parameter values to gain an O$_c$ above this.

Scaled yield maps for ten of the studied fields are shown in Figure 1. They are ranked in decreasing order of opportunity. Fields with the largest O$_c$ had significant magnitudes of trend in the yield data. As O$_c$ decreases, so does the contrast between high and low yielding sub-regions of fields, such that C9-11 (a grape field), with the lowest O$_c$ of all the fields, exhibits something akin to white noise.

Temporal fluctuations in O$_c$ size are worth noting. The 1996 and 1998 seasons at West Creek (Fig. 1f-g) displayed less opportunity than in 1997 (Fig. 1e). The differences in these yield maps, which are reflected in their O$_c$, have been attributed to rain – both 1996 and 1998 recorded above normal within-season rain, whereas 1997 experienced little. In drier years crop production is dependent on stored soil moisture thus soil texture/available soil moisture become important yield determining factors. The large triangular feature on the left of Fig. 1e is a red ridge of lighter textured soil running through a field of predominantly heavy clay. In wetter years (1996 and 1998) there was sufficient within season rain to continually replenish soil moisture thus the effect of texture is not as dominant. This field illustrates an important point in determining management zones and opportunity: it is unlikely that a single season's yield maps will characterise the expected variation in most crop fields. Single seasons have been used here just to illustrate the method.

Figure 1 is an interesting case because, while it has the largest magnitude of O$_c$, it presents one of the pit-falls of this method: the effect of natural disasters. The yield of Maidens in 1995 was severely affected by frost. The low mean yield lead to a very large magnitude of CV$_a$ (74.7%) which, when coupled with a trend (which is strongly correlated to topography and frost damage), has given the illusion of a large opportunity for SSCM. Spatial catastrophes – such as frost, waterlogging, and insect damage – may appear disguised as spatial opportunity. Local knowledge is necessary for correct interpretation of the O_c value.

The point-to-point accuracy of yield monitors is also acknowledged as a significant contributor to the validity of the O_c assessment procedure. As with any analysis the outcome is dependent on the quality of the data that is entered into the model. Noisy or poorly calibrated monitors will produce erroneous values.

The last column of Table 3 presents each field's *b*' value The range of *b*' is from 0.44 (Rowlands 5) to 0.89 (C9-11). Interestingly, the three grape fields recorded the three largest *b*' values. Fairfield Smith reported *b*' values for wheat of 0.44–0.72; our values of *b*' for wheat, even though from much larger fields, are very similar and range between 0.44–0.76. When the O_c and *b*' are compared, a moderate negative correlation is found (r = -0.43).

As mentioned in the Methods, $q_{50}(CV_a)$ is currently a 'best guess' of the minimum magnitude in yield variation ($q_\alpha CV_a$) that is needed for SSCM to be viable. Differences in both crop type and production systems will require that $q_\alpha(CV_a)$ be determined for each unique production area. For example wheat growers in Western Australia with yields of 1 Mg/ha may consider 0.4 Mg/ha a significant increase whilst European growers, averaging more than 7 Mg/ha may not. Similarly minimum threshold values for a winegrape crop yielding 25 Mg/ha will differ from the wheat growers in Western Australia. Further research needs to be conducted to determine ($q_\alpha CV_a$) for a range of crops and production systems.

Ultimately, it is envisaged that limits will be set to the O_c, whereby one can decide whether there is a 'high', 'medium' or 'low' opportunity for SSCM. At this stage, any proposed limits are only tentative, however subjective appraisal of the yield maps in Figure 1 suggests that an O_c of 20 represents a threshold above which SSCM is more viable than uniform management.

In the future, as our database of information grows, a number of years of data may be analysed to provide a mean O_c. It is also envisioned that different layers of information (*e.g.* yield estimates from remote sensing (Boydell and McBratney) will be combined to form an 'integrated variation map' that may be analysed for the opportunity of SSCM. Thus opportunity may be judged on the variability of the entire production system and not just yield. This may allow the O_c to shift from a retrospective to a predictive assessment of production variation.

Finally, having established that there *is* an opportunity for a change in field management practices, the next step is to ask what component(s) of management can be changed. This requires a more detailed investigation of site-specific crop variation that can only be achieved through field-scale experimentation (Gotway Crawford *et al.* 1997; Cook *et al.* 1999) and continued crop monitoring.

CONCLUSIONS

Without the backing of information, many farmers may feel reluctant to change their traditional agronomic practices. A database of yield maps contains a vast amount of information that can be utilised to assess the opportunity for Site-Specific Crop Management.

While searching for a method to quantify the opportunity for SSCM, the empirical law devised by Fairfield Smith (1938) was initially applied to yield monitor data, but was found lacking. The SSCM Opportunity Index offers greater possibilities.

The Opportunity Index calculated for 20 fields showed that no particular crop is suited to SSCM over another. There is evidence of temporal instability in the values of O_c for a given field. Perhaps several seasons of crop data are needed before stability in O_c is found. This is a subject that requires more work, but unfortunately, at this stage of SSCM's development, very few fields have more than five years of yield maps.

A tentative proposal is that an O_c greater than 20 suggests a good opportunity for changing one's management practices to SSCM, although further research is needed to justify this recommendation.

Further information on the application and development of the O_c can be found in a paper by Pringle *et al.* (submitted).

ADDENDUM

A Management Zone Opportunity Index

As defined above, O_c is a 'continuous' management index. It does not imply management zones. Since current PA technology may be more effective when applied in a management-zone rather than a continuous context, a management-zone opportunity index (O_{zs}) should be defined.

Briefly, an index based on statistical parsimony (Lark, in press) derived from the Akaike Information Criterion could be used:

$$\mathrm{AIC} = n \bullet \ln(\mathrm{RMS}) + 2p \tag{10}$$

where:
n = the number of observations
RMS = the residual mean square for the model fit
p = the number of parameters in the fitted model.

In this context we can write it as:

$$O_{zs}(z) = 2z - n \bullet \ln(r^2) \tag{11}$$

where:
z = the number of zones or spatial units in the field.
r^2 = the fit of the model

In this situation z is not the number of classes that might be fitted by a (fuzzy) k-means algorithm but rather the number of discrete contiguous spatial zones. These contiguous spatial zones can be formed by including the spatial co-ordinates as well as the yield in the numerical classification. For this paper Eastings and Northings were added as variables to ensure zones were single entities. (N.B. these zones are not the same as treatment classes. While individual zones may be discrete treatment classes, other treatment classes may be composed of two or more zones (that are spatially discrete from each other). An additional 2 zone model determined “by eye” (E_2) was also added (discussed later). For the example we have chosen West Creek 1997 as it has a high O_c and the large feature on the left hand side indicates it is suitable for zone management. For this analysis the headland artifacts shown in Figure 1e were removed manually to the best of our ability.

The goodness of fit of the zone models (in describing yield variation) was determined from the r^2 of a one-way analysis of variance model. The use of n is problematic however, because it is the number of independent observations. For this discussion the method of Bishop *et al.* (in press) has been adapted to raw yield data to determine n using 5000 data points. The number of zones (z), which minimises O_{zs}, can be regarded as optimal.

From Table 4 we can see that O_{zs} becomes asymptotic after 7 zones indicating that there is little benefit in managing additional zones. The amount of variation accounted for in the management zones can be determined by comparing the yield variogram to that of the residuals. In this case a spherical model has been fitted to the yield data and the exponential model to the residual data. There has been no penalty imposed to models with more parameters, thus as the number of zones increases the amount of variance explained by the model increases and the sill decreases.

Figure 2 shows the variograms for the yield and residuals from the ANOVA for 5, 6 and 7 zone models. The area between the sills of the yield and the individual residual variograms can be considered representative of the amount of variation explained by the zonal models. Taking the variance at 1000m, the 5, 6, and 7 zone models explain 51%, 63% and 67% of the variation in yield respectively. These numbers are comparable to the r^2 derived from the ANOVA of the cluster means however they also include a spatial component.

While determining which model best describes variation the calculation of O_{zs} does not take into account issues such as differences between means of zones, gross margins etc. Ideally a better opportunity index would be an economic one (O_{ze}), measured in dollars (per hectare):

$$O_{ze} = \sum_{i=1}^{z}\left[\frac{G_i}{A_i}\right] \qquad (12)$$

where A_i = the area of zone i.
G_i = the gross margin for zone i which is calculated from;

$$G_i = P_i - C_i - F_i \qquad (13)$$

Where: P_i = value of production
C_i = agronomic cost of production
F_i = environmental cost of production (which is still difficult to calculate).

Here the assumption is that the zones are suitable for PA. This time, the optimal z is the number that maximises O_{ze}.

Currently, it may be difficult to obtain all the data to calculate O_{ze} but developing methods to obtain these data should be an aim of further research. In the meantime, we might think of using a compromise between the statistical and economic indices, which is really what our O_c is. If we replace S in Equation 8 by

$$S(z) = r^2\left[\frac{A}{z}\right] + J_a\left[1 - r^2\right] \tag{14}$$

Table 4. Estimates of Opportunity for managing different numbers of zones for West Creek 1997.

z	n	r^2	O_{zs}	O_{zi}
1	112	0	∞	-
2	112	0.002	389.51	2.99
3	112	0.061	189.67	6.70
4	112	0.181	158.87	9.54
5	112	0.484	140.09	13.64
6	112	0.582	124.57	13.63
7	112	0.616	49.86	12.98
8	112	0.660	48.82	12.56
9	112	0.692	49.33	12.12
10	112	0.700	48.84	11.57
E_2	112	0.257	156.17	15.73

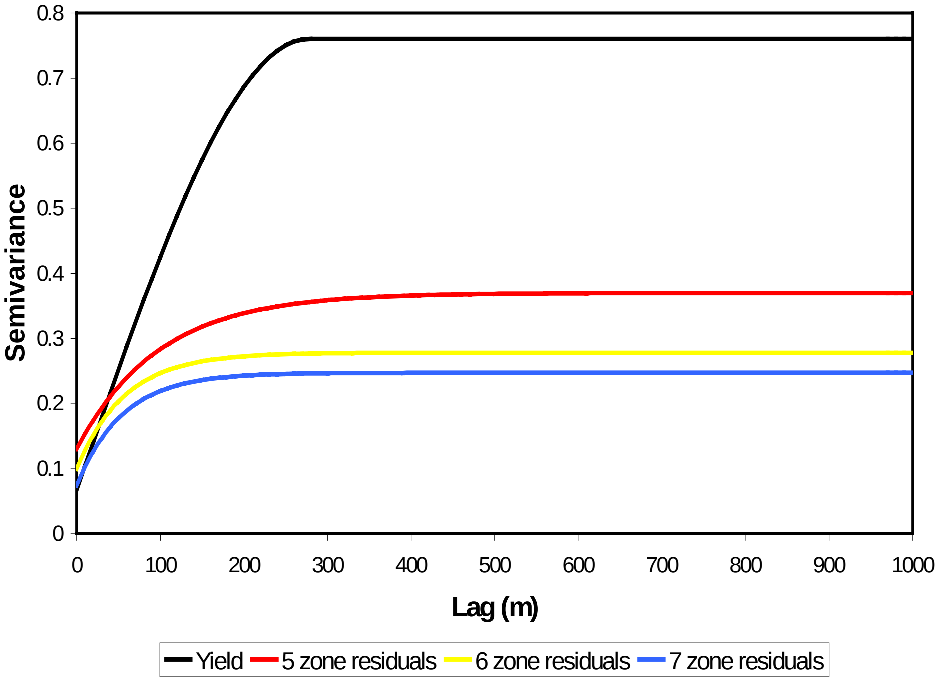

Figure 2: Variograms of Yield and Residuals from the ANOVA between management zones.

it is possible to calculate O_{zi}, which should be maximised. Results of this are presented in Table 4. O_{zi} values indicate that the optimal number of management zones in this field is either 5 or 6 (less than that indicated by the O_{zs} analysis). It should be noted that these values cannot be compared directly to the O_c values as they are situated on a different scale. From Figure 3, diagrams with only 3 and 4 zones are reflecting the heavy weighting of the spatial coordinates in the analysis thus have poor r^2 values when compared with the yield. When compared with the yield map, diagrams with 5, 6 and 7 zones highlight the main management zones in the field. Diagrams with 8 or more zones are starting to identify small areas in the field, which are probably not viable units with current variable-rate technologies.

In this field we would expect two management zones to have a reasonable opportunity due to the large feature on the left-hand side. However the heavy weighting with spatial coordinates in the numerical classification negates this in the 2 zone model. By applying expert knowledge we can segregate this feature into z_1 and specify the rest of the field as z_2 and analyse this model (E_2). The r^2 for the ANOVA between z_1 and z_2 is 0.257 significantly higher than the 2 and even the 3 and 4 zone model derived using yield and spatial coordinates in Table 4. Plotting the variogram of the residuals against the yield shows that this minimal segregation already accounts for 28% of the variation in yield.

The O_{zi} for this model (15.73) is the highest of any of the models due to S_z being weighted to minimise zones. As expected two large discrete contiguous zones provide a good opportunity for PA. Whether it should have a higher O_{zi} than a more complex model that better fits yield variation is a point for further research and discussion. The r^2 O_{zs} and O_{zi} values from E_2 highlight the need for a better algorithm for deriving the discrete contiguous zones.

The authors would like to emphasis that this is only a preliminary model and presented here as an example. Considerable work still needs to be done especially on the development of a zonal algorithm. Further we would like to reiterate that while this analysis has been done on a field for one year, data from several years will be necessary before a true indication of the opportunity will be known.

ACKNOWLEDGMENTS

We acknowledge the Australian Grains Research and Development Corporation (GRDC) and Australian Cotton Research and Development Corporation (CRDC) for supporting this research.

We should like to thank Dr. Simon Cook of CSIRO Land and Water in Perth Agriculture for permission to use the Wyalkatchem data, and Craig and Judy Boydell (at Moree), Doug and Noela Maitland (at Wyalkatchem), Peter Glennie (Norwood), Nick Barton (Telleraga) and Orlando-Wyndham (Cowra) for access to their fields.

Thanks to Broughton Boydell for his helpful discussion.

REFERENCES

Bishop, T.F.A., McBratney, A.B. and Whelan, B.M. (in press). Measuring the quality of digital soil maps using information criteria. Geoderma

Boydell, B. and McBratney, A.B. 1999. Identifying potential within-field management zones from cotton yield estimates. In *Precision Aagricultur '99: Proceedings of the 2nd European Conference on Precision Agriculture* (Ed. J.V. Stafford), Part 1, pp. 331-341. Sheffield Academic Press.

Cook, S.E., Adams, M.L., Corner, R.J. 1999. On-farm experimentation to determine site-specific responses to variable inputs. In *Precision Agriculture: Proceedings of the 4th International Conference* (Eds P.C. Robert, R.H. Rust, and W.E. Larson), Part A, pp. 611–621. ASA/CSSA/SSSA, Madison, Wi .

Fairfield Smith, H. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. Journal of Agricultural Science, Cambridge 28: p. 1–23.

Gotway Crawford, C.A., Bullock, D.G., Pierce, F.J., Stroup, W.W., Hergert, G.W., and Eskridge, K.M. 1997. Experimental design and statistical evaluation techniques for site-specific management. In *The State of Site-Specific Management for Agriculture* (Eds F.J. Pierce and E.J. Sadler), p. 301–335. ASA/CSSA/SSSA, Madison, Wi .

Isaaks, E.H., and Srivastava, R.M. 1989. Applied Geostatistics. Oxford University Press, New York, NY.

Journel A.G. and Huijbregts, C. 1978. Mining Geostatistics. Academic Press, London

Lark, R.M., Bolam, H.C., Mayr, T., Bradley, R.I., Burton, R.G.O., and Dampney, P.M.R. 1999. Analysis of yield maps in support of field investigations of soil variation. In *Precision Agriculture '99* (Ed. J.V. Stafford), p. 151–161. Sheffield Academic Press, Sheffield, UK.

Lark, R.M. (in press). Some tools for parsimonious modeling and interpretation of soil and crop systems. Soil & Tillage Research.

Lars, T. and Antje, G. 2000. An Expert Filter Removing Erroneous Yield Data. *Proceedings of the 5th International Conference on Precision Agriculture and Other Precision Resource Management. Minneapolis 2000* (this conference)

McBratney, A.B., Whelan, B.M. and Shatar, T.M. 1997. Variability and uncertainty in spatial, temporal and spatiotemporal crop-yield and related data. In *Precision Agriculture: Spatial and Temporal Variability of Environmental Quality* (Eds J.V. Lake, G.R. Bock, and J.A. Goode), p.141–160. John Wiley, Chichester, UK..

McBratney, A.B. and Webster, R. 1986. Choosing functions for semi–variograms of soil properties and fitting them to sampling estimates. Journal of Soil Science 37: p. 617–639.

Mercer, W.B., and Hall, A.D. 1911. The experimental error of field trials. Journal of Agricultural Science, Cambridge 4: p. 107–127.

Mulla, D.J. 1993. Mapping and managing spatial pattern in soil fertility and crop yield. In *Soil Specific Crop Management* (Eds P.C. Robert, R.H. Rust, and W.E. Larson), p. 15–26. ASA/CSSA/SSSA, Madison, Wi .

Perrier, E.R., and Wilding, L.P. 1986. An evaluation of computational methods for field uniformity studies. Advances in Agronomy 39: p. 265–312.

Pringle M.J., McBratney, A.B., Whelan, B.M. and Taylor, J.A. (submitted) A preliminary approach to assessing the opportunity for site-specific crop management in a field, using yield monitor data. Journal of Agricultural Science, Cambridge.

Russo, D., and Bresler, E. 1981. Effect of field variability in soil hydraulic properties on solutions of unsaturated water and salt flows. Soil Science Society of America Journal 45: p. 675–681.

Webster, R., and McBratney, A.B. 1989. On the Akaike Information Criterion for choosing models for variograms of soil properties. Journal of Soil Science 40: p. 493–496.

Webster, R., and Oliver, M.A. 1990. Statistical Methods in Soil and Land Resource Survey. Oxford University Press. New York, NY.

Table 2. Summary statistics of the fields used for development of the opportunity index.

Crop	Location	Field	Year	~ Area (ha)	Mean yield (Mg/ha)	Std. dev. yield (Mg/ha)	CV (%)
Wheat	Moree, NSW	B4	1995	8	1.90	0.73	38.4
Wheat	Moree, NSW	East Creek	1997	77	3.86	1.45	37.6
Wheat	Moree, NSW	Maidens	1995	88	0.96	0.78	81.3
Wheat	Moree, NSW	West Creek	1996	80	5.40	0.67	12.4
Wheat	Moree, NSW	West Creek	1997	80	3.69	0.92	24.9
Wheat	Moree, NSW	West Creek	1998	80	5.58	0.95	17.0
Wheat	Wyalkatchem, WA	Home 1	1998	61	1.83	0.47	25.7
Wheat	Wyalkatchem, WA	Home 5	1998	40	1.09	0.38	34.9
Wheat	Wyalkatchem, WA	Rowlands 1	1995	75	1.49	0.52	34.9
Wheat	Wyalkatchem, WA	Shire 4	1997	64	1.09	0.24	22.0
Wheat	Wyalkatchem, WA	Shire 4	1999	64	2.25	0.45	20.0
Grapes	Cowra, NSW	C3-8	1999	14	21.76	7.02	32.3
Grapes	Cowra, NSW	C9-11	1999	2	24.13	4.98	20.6
Grapes	Cowra, NSW	D3-4	1999	6	20.37	6.97	34.2
Cotton	Moree, NSW	Norwood 28	1998	42	2.33	0.63	27.0
Cotton	Moree, NSW	Telleraga 10	1998	97	1.76	0.38	21.6
Lupins	Wyalkatchem, WA	Blackies 6	1998	51	1.08	0.35	32.4
Lupins	Wyalkatchem, WA	Home 8	1997	30	0.54	0.16	29.6
Sorghum	Moree, NSW	East Creek	1996	77	6.90	1.07	15.6
Sorghum	Moree, NSW	W80	1997	42	4.21	1.02	24.2

Table 3. Parameters used in the determination of the opportunity index.

Field	Crop	Year	CV_a (%)	J_a (ha)	P_t	S (ha)	s (ha)	$\mathbf{O_c}$	b'
Maidens	Wheat	1995	74.7	0.074	0.330	29.101	0.036	**47.2**	0.57
Home 5	Wheat	1998	47.9	0.118	0.589	23.627	0.036	**34.1**	0.44
C3-8	Grapes	1999	36.7	0.038	0.293	4.135	0.005	**33.5**	0.88
D3-4	Grapes	1999	43.1	0.003	0.364	2.187	0.005	**26.4**	0.86
Blackies 6	Lupins	1998	26.2	0.301	0.491	25.174	0.036	**26.0**	0.65
Rowlands 1	Wheat	1995	30.4	0.108	0.251	18.927	0.036	**24.3**	0.60
Home 1	Wheat	1998	28.1	0.220	0.314	19.308	0.036	**23.6**	0.55
W80	Sorghum	1997	26.7	0.026	0.473	19.885	0.036	**23.3**	0.67
West Creek	Wheat	1997	21.7	0.300	0.266	21.489	0.036	**21.9**	0.52
Shire 4	Wheat	1997	20.1	0.050	0.287	18.397	0.036	**19.5**	0.70
Telleraga 10	Cotton	1998	18.3	0.112	0.202	19.702	0.036	**19.2**	0.70
East Creek	Wheat	1997	29.9	0.076	0.137	10.622	0.036	**18.0**	0.64
Shire 4	Wheat	1999	17.0	0.094	0.289	18.561	0.036	**18.0**	0.70
West Creek	Wheat	1996	10.6	0.074	0.245	19.632	0.036	**14.6**	0.76
East Creek	Sorghum	1996	29.9	0.076	0.204	15.866	0.036	**13.4**	0.77
West Creek	Wheat	1998	10.3	0.035	0.199	15.918	0.036	**13.0**	0.74
Home 8	Lupins	1997	24.9	0.054	0.183	5.548	0.036	**11.9**	0.78
Norwood 28	Cotton	1998	15.3	0.043	0.141	5.946	0.036	**9.7**	0.77
B4	Wheat	1995	37.2	0.019	0.271	2.178	0.036	**9.1**	0.62
C9-11	Grapes	1999	7.6	<0.001	0.068	0.136	0.005	**2.8**	0.89

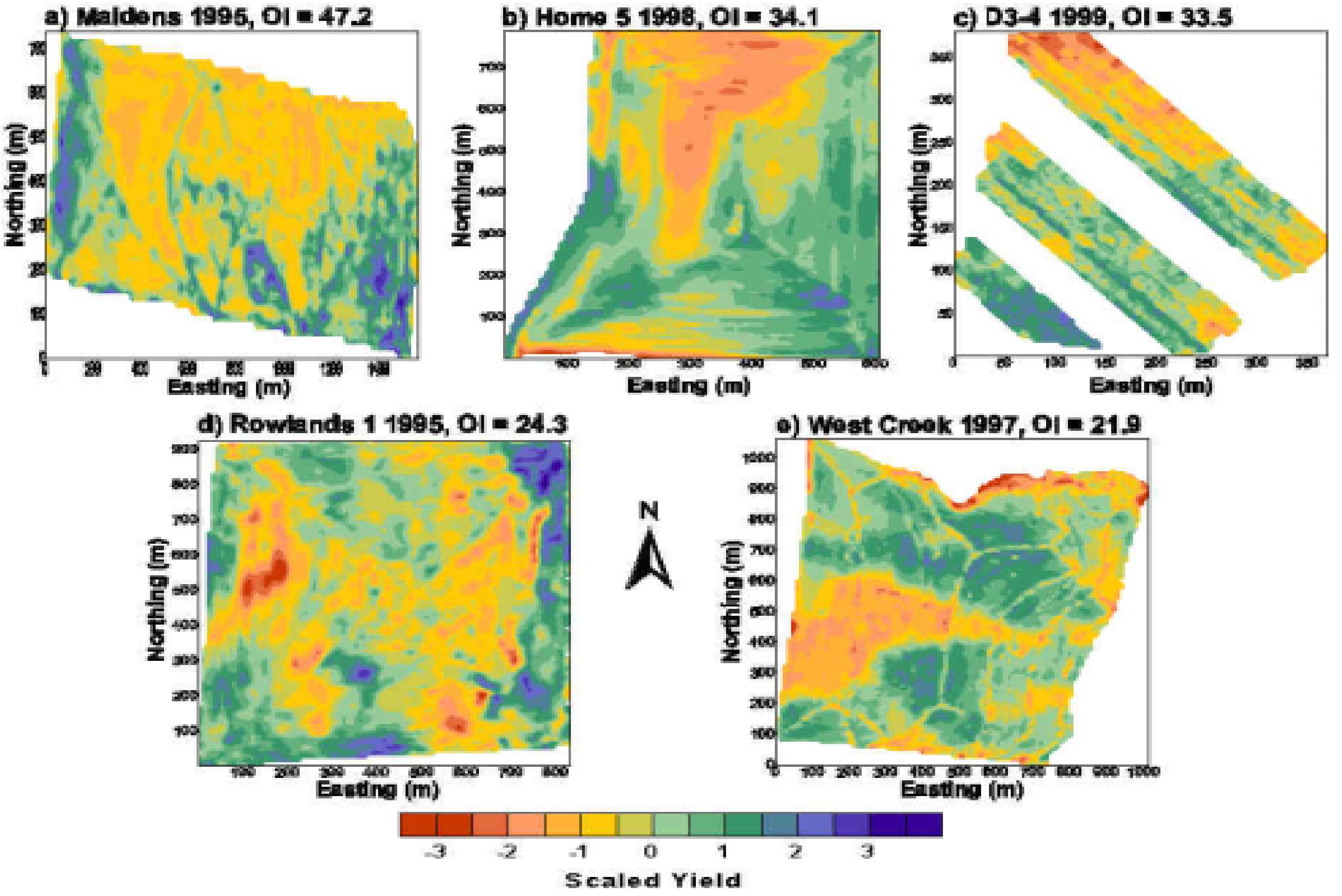

Figure 1(a-e). Scaled yield maps of fields used in the determination of the O_c.

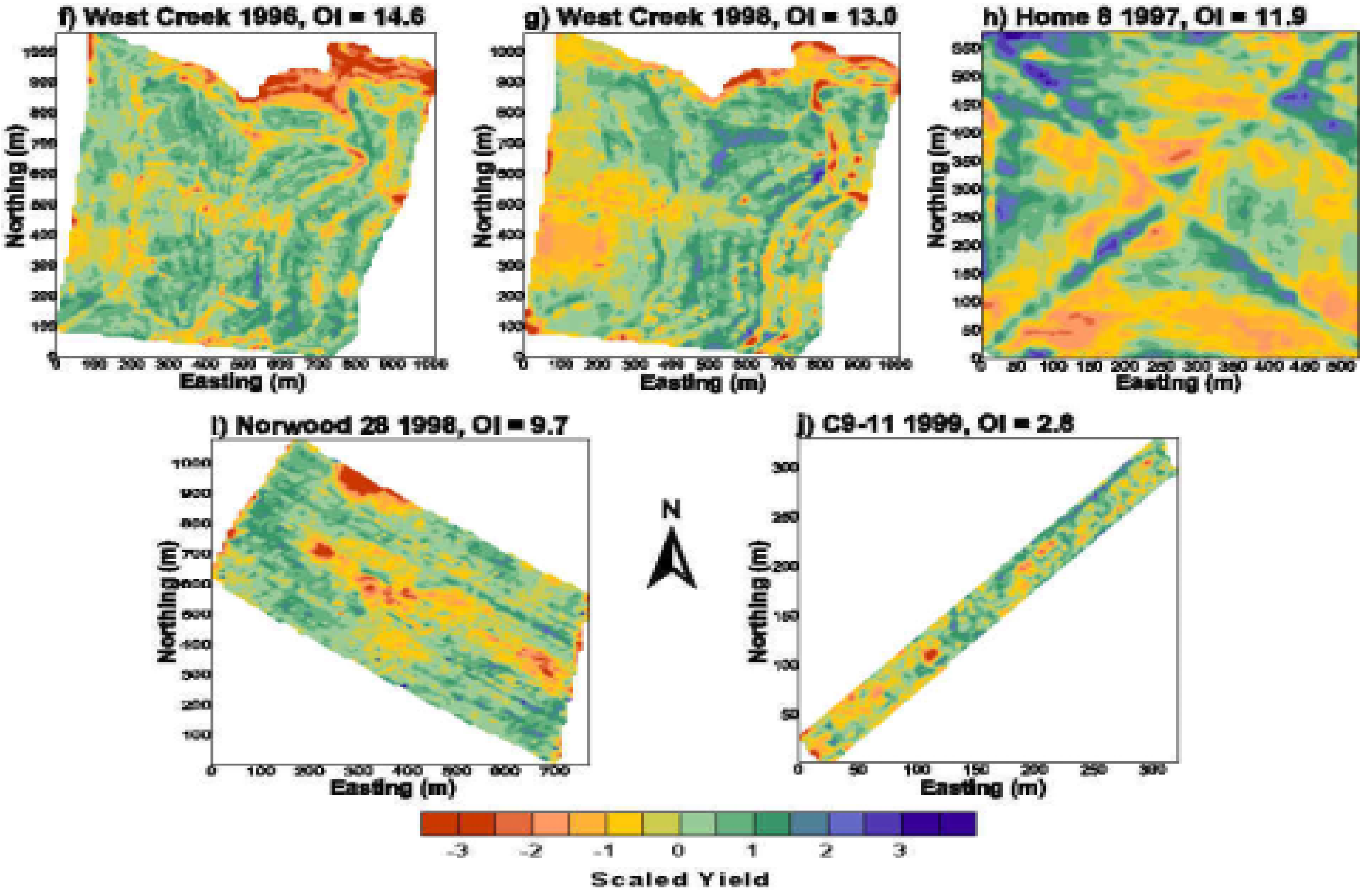

Figure 1(f-j). Scaled yield maps of fields used in the determination of the O_c.

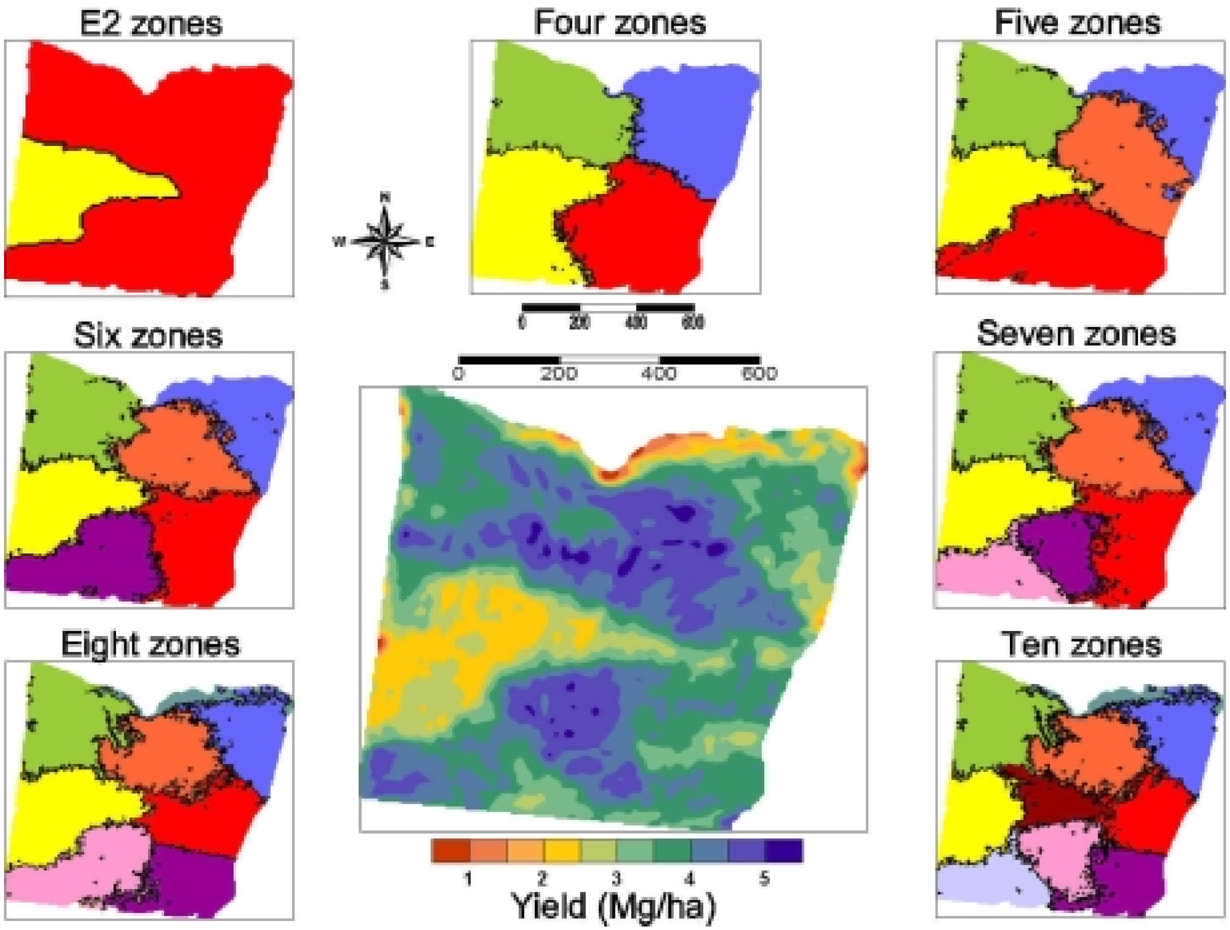

Figure 3. Maps of Yield, Eastings and Northings showing the spatial distribution of potential management zones. 2, 3 and 9 zone models not shown.