Economics of Management Zone Delineation in Cotton Precision Agriculture

MargaritaVelandia, Roderick M. Rejesus, and Eduardo Segarra
Department of Agricultural and Applied Economics
Texas Tech University
Lubbock, TX 79409-2132
USA

Kevin Bronson
Texas Agricultural Experiment Station
Texas A&M University
Lubbock, TX 79401
USA

Contributed paper prepared for presentation at the International Association of Agricultural Economists (IAAE) Conference, Gold Coast, Australia
August 12-18, 2006

Economics of Management Zone Delineation in Cotton Precision Agriculture

Abstract:

This paper develops a management zone delineation procedure based on a spatial clustering approach and evaluates its economic impact for the case of Texas cotton production. The results show that there is potential economic value in using a spatial approach to management zone delineation, but its value critically depends on the cost of collecting soil test information needed to delineate those zones.

JEL Classification: Q12

Introduction

Optimally configuring management zones for better management of farm inputs is one of the most important issues in precision farming and variable rate application. Management zones are geographical areas that can be treated as homogenous, so that input application and decision-making can be treated separately for each zone. This approach may then lead to more efficient management of the farm. The objectives of this paper are: (1) to develop a univariate management zone delineation procedure based on a specific spatial clustering approach called ESDA (Exploratory Spatial Data Analysis), and (2) to evaluate the potential economic impact of this management zone delineation procedure for the case of cotton production in the Texas High Plains. Moreover, this paper implements spatial econometric techniques and shows its importance in economically evaluating management zone delineation procedures.

Empirical Methodology

Data and the ESDA Approach to Management Zone Delineation

The data used to establish management zones is based on a 2002 agronomic cotton experiment designed to study nitrogen (N) use for cotton production in the Southern High Plains of Texas. The experiment is a randomized complete block design with three replicates and each replicate was within a center pivot irrigation span. The data was originally collected as point data (135 data points). But we spatially interpolated the data into 443 grids (16m x 16m) in order to obtain a balanced design for analysis. The spatial structure of the yield data used in the analysis is presented in Figure 1.

As mentioned in the introductory section, we use a spatial clustering approach called ESDA as the main procedure for establishing management zones. ESDA can be defined as a method that combines different techniques to visualize spatial distributions, identify patterns of different locations, and identify patterns of association between these locations (Anselin, 1998). This method is based on the concept of spatial autocorrelation, which is the relationship between spatial units, and makes use of the concept of distance between locations. Hence, this approach takes the spatial structure of the data into account when delineating management zones, which is an improvement to simple clustering algorithms used in the past.

The step-by-step procedure for the ESDA approach to management zone delineation can be described as follows: (1) Define the 'neighborhood' structure of each grid (see Bivand (1998) for procedures to define the neighborhood structure); (2) Establish a 'weight matrix', that defines the neighbor structure (see, Bivand (1998)); (3) Test for the presence of spatial autocorrelation using the Moran's I statistic (see, Anselin (1998)); (4) Graphically visualize the spatial correlation structure (if step (3) indicates there is spatial autocorrelation) with a Moran Scatterplot; and (5) Establish the management zones. Since we have a grid-based data structure, we used a "rook" structure (four neighbors to each cell, north, south, east and west) to define the neighborhood in our management zone delineation procedure (Anselin, Bongiovanni, and Lowenberg-Deboer , 2004).[1]

Using soil nitrate as the variable of interest, the computed global Moran's I statistic, based on the "rook" neighborhood structure, is 14.38 and this has a p-value of <0.001. This indicates that there is spatial autocorrelation in the data. Based on this result, a Moran scatterplot is created and management zones based on this scatterplot is then determined (Figure 2). There are three management zones established based on our procedure: management zone 1 (MZ1) represents high nitrate areas (i.e. grids with high nitrate levels have "neighbors" with high nitrate levels), management zone 2 (MZ2) represents low nitrate areas (i.e. grids with low nitrate levels have "neighbors" with low nitrate levels), and management zone 3 (MZ3) represents the area with a mix of high and low nitrate levels.

Economic Model and Estimation Procedures

The economic model to assess the impact of the management zone delineation procedure is based on a mathematical programming model for spatial profit (or net return) maximization (See, among others, Lowenberg- Deboer and Boehlje, 1996; Bongiovanni and Lowenberg-Deboer, 1998; Anselin, Bongiovanni, and Lowenberg-Deboer, 2004; Bullock, Lowenberg-DeBoer, and Swinton, 2002). In this framework, we compute the expected net returns from: (1) a uniform N rate application based on an agronomic optimum (URA), (2) a uniform N rate application based on an economic optimum (URE), and (3) a variable rate N application based on the economic optimum for each of the management zones established through our spatial procedure above (VRN). Hence, our economic analysis evaluates the economic impact of our

[1] There are other contiguity-based neighborhood structures like the "queen" (eight neighbors to each cell) or the "bishop" (four neighbors with common vertex) structure. We also used these structures for defining management zones and found very similar results to the rook structure. The management zone delineation results for the alternative neighborhood structures are not reported here, but are available from the authors upon request.

management zone delineation procedure relative to the uniform N rate application based on the agronomically recommended rate and the economically optimum rate calculated from the model. In addition, we also compare the expected net returns from VRN to another scenario where N is variably applied based on landscape position (VRL). Landscape positions considered are high elevation, medium elevation, or low elevation.

For the uniform N application, we first use the agronomically recommended N rate (52 lbs/acre) and then calculate the corresponding net returns based on the parameters of the spatial profit maximization model (described below). An economically optimal uniform N rate application is computed by maximizing the spatial profit function with respect to N (described in equation (2) below). We then compare the net return figures for the uniform rate cases (URA and URE) to the net returns figures for both of the variable rate N application scenarios (VRN and VRL). These net return calculations utilize the spatial optimization model below, where the main component is a spatial cotton yield response function (for each management unit). We use the quadratic specification for the yield response function:

(1) $$Yield_{ij} = \alpha_i + \beta_i N_{ij} + \gamma_i N_{ij}^2 + \varepsilon_{ij}$$

where $Yield_{ij}$ is the cotton yield, N_{ij} is the N rate, *i* indexes management zone, and *j* is the location (in this case, the grids) within each management zone. We firs estimate (1) using Ordinary Least Squares (OLS) and test for the presence of spatial autocorrelation in the residuals. If it is present, then appropriate spatial econometric techniques are implemented to account for the spatial autocorrelation in the residuals. In our case, the spatial error model is used to account for spatial autocorrelation in the residuals (See Anselin, 1988).

Once the parameters of the cotton yield response function are estimated, these estimates are used to formulate an optimization model to maximize profit for a representative farm. In particular, we maximize net returns over N cost (i.e. fixed costs are not considered) using the yield response parameters estimated and available data on prices/costs.

The net return above N cost for the farm is defined as the weighted sum of the net returns in each management zone (for the case of variable rate application), where the weights are the proportion of the area in the management zone. For the case of finding the economically optimum uniform N rate application, this weight is set to one and there is no management zone delineation. More formally, the mathematical programming model can be expressed as:

(2) $$\underset{N}{Max}\ E[\pi] = \sum_{i=1}^{m} (A\omega_i E[P_c(\alpha_i + \beta_i N_i + \gamma_i N_i^2) - r_N N_i]$$

where: E = Expectation operator, π = Total net returns over N fertilizer and fixed cost ($), A = Total land area (22,000 acres), ω_i = Proportion of total land area allocated to management unit i (i.e. for the management zones based on the spatial approach, zone 1= 37%, zone 2= 48%, zone 3= 15%), i = Management unit (either the whole field or the management zones), m = Total number of management units (m = 1 for uniform rate application and m = 3 for variable rate based on the management zones delineated using the spatial approach), P_c = Price of cotton ($0.47 per lb, see Bronson et. al, 2005), N_i = Quantity of N applied in management unit i (in lbs/acre), and r_N = Price of N fertilizer applied ($0.21/lb, see Bronson et. al, 2005)

Results and Discussion

Response Function Estimation Results

The results of both the OLS and spatial error estimation procedures are presented in Table 1.[2] All the coefficients follow our a priori expectations and are all statistically significant (at the 10% level). These results suggest that there are differences in the yield response for each

[2] Note that the yield response function estimated in Table 1 is based on the management zones delineated using our spatial approach. Although not reported here, we also estimate the yield response function when the management zones are based on landscape position, in order to calculate the net returns for the VRL scenario.

management zone. Further, the magnitudes of the coefficients and standard errors are different in the spatial error model as compared to the traditional OLS. This suggests that economic inferences from these two models would be different and that incorrect decisions could be made when only traditional OLS techniques, rather than spatial econometric methods, are used in the yield response estimation. Additionally, when the spatial error structure is modeled, the fit of the model improves as shown by the increase of the log likelihood and a decrease in Akaike Information Criteria (AIC). The improvement of the model was also to be expected because of the highly significant spatial error (lambda) coefficient.

Mathematical Programming Results: Yield, Nitrogen, and Profitability

Based on the estimated response function(s) and the optimization model described above, we estimated the yield, the N application levels, and the net returns over N cost for each of the different application techniques considered: URA, URE, VRN, and VRL. Each of these application scenarios was examined by using a yield response function estimated both by OLS and by using the spatial error model (SEM) estimated through a maximum likelihood technique (ML). This allows us to see the potential magnitude of inference or recommendation errors that could be committed when spatial autocorrelation is not properly accounted for in the yield response estimation.

A comparison of the returns for the different N rate application techniques is presented in Table 2. The OLS technique tends to overestimate the benefits from variable rate application relative to the uniform rate based on the agronomic recommendations (VRN-URA), and OLS tends to underestimate the benefits from variable rate application relative to the uniform rate based on the economic optimization model (VRN-URE). Note that with the use of the spatial error model, the variable rate application of N based on the management zones delineated still tend to have a higher net return relative to the uniform rate based on the agronomic optimum (VRN-URA), albeit smaller than if OLS was used. The spatial error model for the variable rate application of N based on the management zones delineated also tend to have a higher net return relative to the uniform rate base on the economic optimum (VRN-URE), albeit higher than if OLS was used. Another notable comparison is the higher net return of VRN relative to VRL, once we correct the model for spatial autocorrelation. This shows that our spatial approach to management zone delineation has added value (when used to variably apply N) relative to a management zone delineation technique based solely on landscape position.

The average N levels for the different application techniques are presented in Table 3. Our results show that, on average, the variable rate system using the delineated management zones based on the spatial approach tend to have higher yields than the uniform rate application techniques (Table 3). The VRN scenario also generated a higher average yield than the VRL scenario. With regards to N application levels, the variable rate scenario (VRN) tends to utilize more N (on average) than the URE technique (Table 4). But the variable rate scenario tends to have lower N levels relative to the URA scenario. Note, however, that the variable rate scenarios (VRN) tend to more efficiently utilize N because it applies less N in zones with high soil nitrate levels and more N in zones with low soil nitrate levels.[3] Therefore, even if N application is higher (on average) for the variable rate techniques, the more efficient use of the N fertilizer may possibly reduce nitrate run-off in the soil and, consequently, reduce non-point source pollution.

Note that the results with regards to the net returns are based on an approach that does not take into account fixed costs. This is reasonable since we only undertake a short run analysis that utilize one year of data. However, there is truly a fixed cost associated with the soil test nitrate

[3] In the interest of space, the exact figures for the applied N in each management zone are not explicitly reported here, but are available from the authors upon request.

information and this needs to be taken into consideration when evaluating the variable rate approach based on the delineated management zones. For the experiment considered in this analysis, the estimated cost for the nitrate soil analysis is $9.60/acre (Bronson et. al, 2005). If we consider this cost, then VRN may not be more profitable than URA and URE. A breakeven analysis, where the breakeven fee is simply calculated as the difference between net returns under VRN and net return under URA, shows that for VRN to be more profitable than URA and URE, the cost of the soil analysis needs to be less than $2.21/acre.

Sensitivity Analysis

The two important components that underlie the results presented above are the choice of neighborhood structure and the yield response estimation technique. The rook neighborhood structure is used as the basis for the spatial weights matrix in the delineation of the management zones and in modeling the error structure of the SEM yield response function. Standard OLS techniques and a ML approach to estimating the yield response function are the estimation techniques used to produce the economic results above. In order to check for the sensitivity of the economic results, we also examine the economic effect of using an alternative neighborhood structure (e.g. a queen structure) and/or alternative estimation techniques (Table 4). In general, we find that regardless of neighborhood structure or estimation technique VRN still tend to have higher net returns relative to the uniform rate approaches (URA and URE).

Conclusions

Based on a spatial clustering approach that utilizes a spatial autocorrelation statistic, we are able to develop a procedure for delineating management zones using precision agriculture data from cotton production in the Texas high plains. The results of the optimization model suggest that applying variable N rates based on the management zones delineated (using the spatial approach developed), would result in higher yields and higher net returns above N cost relative to the traditional uniform rate application and relative to a variable rate application based on landscape position. However, the economic advantage of variable rate N application using the delineated zones critically depends on the fixed cost of collecting the soil test information necessary for spatial clustering. If the current cost of collecting this information stays at it current levels, then a variable rate approach using the spatial clustering algorithm may not be economically feasible relative to uniform rate approaches. The results of our analysis also reinforce the observation in past studies that incorrectly estimating yield response functions without correcting for spatial dependence may lead to misleading inferences about the economic impact of variable rate technologies.

References:

Anselin, L. Spatial Econometrics: Methods and Models. Dordrecht, Netherlands: Kluwer Academic, 1988.

Anselin L., Bongiovanni R. and Lowenberg-DeBoer J. “A spatial econometric approach to the economics of site-specific nitrogen management in corn production.” American Journal of Agricultural Economics. 86 (August 2004): 675-687

Bivand, R. “A review of spatial statistical techniques for location studies.” Online. Available at http://www.nhh.no/geo/gib/gib1998/gib98-3/lund.html. 1998. Retrieved October 2003.

Bongiovanni, R. and J. Lowenberg-DeBoer. “Economics of Variable Rate Lime in Indiana,” pp. 1653-1666. Proceedings of the 4th International Conference on Precision Agriculture. St. Paul, MN, July 19-22, 1998.

Bronson, K.F., J. D. Booker, J.P. Bordovsky, J. W. Keeling, T.A. Wheeler, R.K. Boman, M.N. Parajulee, E. Segarra, M. Velandia-Parra, and R.L. Nichols. “Site-Specific Irrigation and Nitrogen Management for Cotton Production in the Southern High Plains,” Working paper. April, 2005.

Bullock, D. S., Lowenberg-Deboer, J. Swinton, S. “Adding value to spatially managed inputs by understanding site-specific yield response.” Agricultural Economics 27(November 2002): 233- 245.

Lowenberg-DeBoer, J. and M. Boehlje. “Revolution, Evaluation or Deadend:Economic Perspectives on Precision Agriculture.” In: Robert, P.; Rust, H. and R. Larson, eds. Proceedings of the 3rd International Conference on Precision Agriculture. Minneapolis, MN, June 23-26, 1996.

Table 1. Parameter estimates of the cotton yield response function for the management zones delineated using the spatial approach

Variables	OLS (Ordinary Least Squares) COEFF ($lbs\ ac^{-1}$)	OLS P-value	SEM (Spatial Error Model) COEFF ($lbs\ ac^{-1}$)	SEM P-value
Constant	827.10	0.0000	916.26	0.0000
N	7.38	0.0000	2.71	0.0006
N^2	-0.10	0.0000	-0.03	0.0021
MZ1	806.09	0.0000	916.23	0.0000
MZ2	814.52	0.0000	893.64	0.1071
MZ3	860.70	0.0000	955.70	0.0369
N x MZ1	10.93	0.0000	2.69	0.0000
N x MZ2	6.59	0.0000	2.22	0.0074
N x MZ3	4.61	0.0002	2.88	0.2214
N^2 x MZ1	-0.19	0.0000	-0.06	0.0000
N^2 x MZ2	-0.06	0.0616	-0.01	0.0885
N^2 x MZ3	-0.055	0.0745	-0.034	0.0975
Lambda	NA	NA	0.64	0.0000

Measures of fit	OLS	SEM
Log Likelihood	-2675.32	-2536.82
AIC	5368.64	5091.65

Diagnostic tests	d.f.	Value	Value	P-value
Lagrange multiplier(error)	1	NA	147.71	0.0000
Robust LM(error)	1	NA	126.68	0.0000
Lagrange multiplier (lag)	1	NA	28.10	0.0000
Robust LM (lag)	1	NA	7.07	0.0781

Table 2. Net returns under different application methods and estimation procedures

	OLS	SEM	Difference (OLS-SEM)
	--- Net Returns ($ acre^{-1}) ---		
Uniform rate, agronomic optimum (URA)	431.09	447.83	-16.73
Uniform rate, economic optimum (URE)	444.47	448.30	-3.83
Variable rate, spatial approach (VRN)	444.76	450.04	-5.28
Variable rate, landscape position (VRL)	445.46	447.45	-1.99
Differences across application techniques			
URE vs. URA (URE – URA)	13.38	0.47	12.9
VRN vs. URA (VRN – URA)	13.67	2.21	11.45
VRN vs. URE (VRN – URE)	0.29	1.74	-1.45
VRL vs. URA (VRL – URA)	14.37	-0.38	14.74
VRL vs. URE (VRL – URE)	0.99	-0.85	1.84
VRN vs. VRL (VRN – VRL)	-0.7	2.59	-3.29

Table 3. Nitrogen levels under different application methods and estimation procedures

	OLS	SEM	Difference (OLS-SEM)
	--- N level (lbs acre^{-1}) ---		
Uniform rate, agronomic optimum (URA)	52.00	52.00	0.00
Uniform rate, economic optimum (URE)	34.21	33.24	0.97
Variable rate, spatial approach (VRN)	34.71	42.66	-7.95
Variable rate, landscape position (VRL)	27.91	28.73	-0.82
Differences across application techniques			
URE vs. URA (URE – URA)	-17.79	-18.76	0.97
VRN vs. URA (VRN – URA)	-17.29	-9.34	-7.95
VRN vs. URE (VRN – URE)	0.5	9.42	-8.92
VRL vs. URA (VRL – URA)	-24.09	-23.27	-0.82
VRL vs. URE (VRL – URE)	-6.3	-4.51	-1.79
VRN vs. VRL (VRN – VRL)	6.8	13.93	-7.13

Table 4. Sensitivity of the differences in net returns under alternative neighborhood structure and estimation method assumptions

Neighborhood structure[1] Estimation Method[2]	Difference in net returns ($ acre^{-1}) across application techniques[3]					
	URE-URA	VRN-URA	VRN-URE	VRL-URA	VRL-URE	VRN-VRL
Rook Structure						
OLS	13.38	13.67	0.29	14.37	0.99	-0.7
SEM (ML)	0.47	2.21	1.74	-0.38	-0.85	2.59
SEM (GM-Two step)	6.25	9.70	3.45	5.61	-0.64	4.09
SEM (GM-Iterated)	5.49	6.76	1.26	4.03	-1.46	2.72
SEM (GM-GHET)	5.26	6.92	1.65	4.10	-1.17	2.82
Queen Structure						
OLS	16.65	20.50	3.85	17.35	0.70	3.15
SEM (ML)	4.66	6.76	2.10	4.25	-0.41	2.50
SEM (GM-Two step)	11.43	14.19	2.76	12.16	0.73	2.03
SEM (GM-Iterated)	11.43	13.28	1.86	11.45	0.03	1.83
SEM (GM-GHET)	10.87	11.61	0.74	9.80	-1.06	1.81

Note: (1) The neighborhood structures considered are rook and queen. Note that these structures are assumed both in the delineation of the management zones for the spatial approach and in specifying the error structure in the SEM model.

(2) The alternative estimation methods considered (aside from the traditional OLS and SEM (ML)) are: SEM using two stage general method of moments (GM-Two step), SEM using iterated general method of moments (GM-Iterated), and SEM using general method of moments that corrects for groupwise heteroskedasticity (GM-GHET).

(3) Application techniques are: uniform rate based on agronomic optimum (URA), uniform rate based on economic optimum (URE), variable rate based on the spatial approach (VRN), and variable rate based on landscape position (VRL).

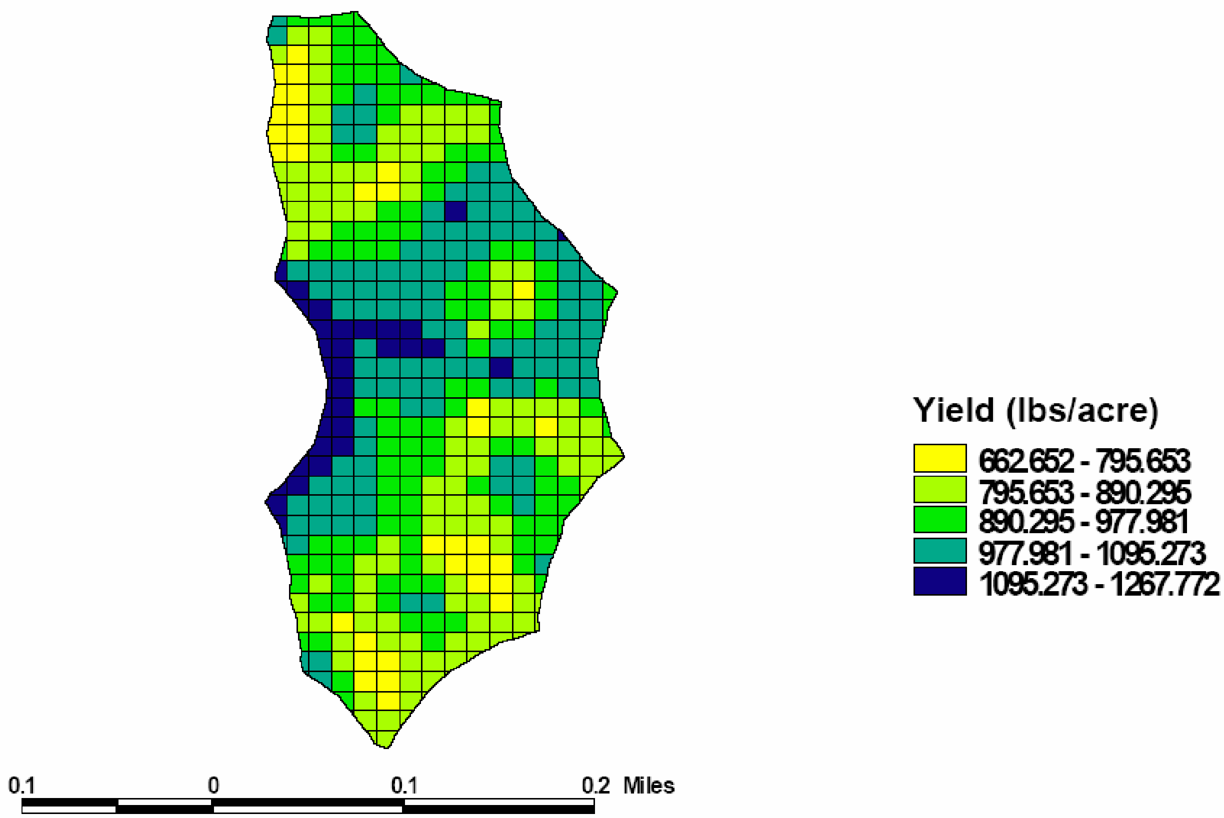

Figure 1. Digitized Grids for Cotton Yield (lbs/acre)

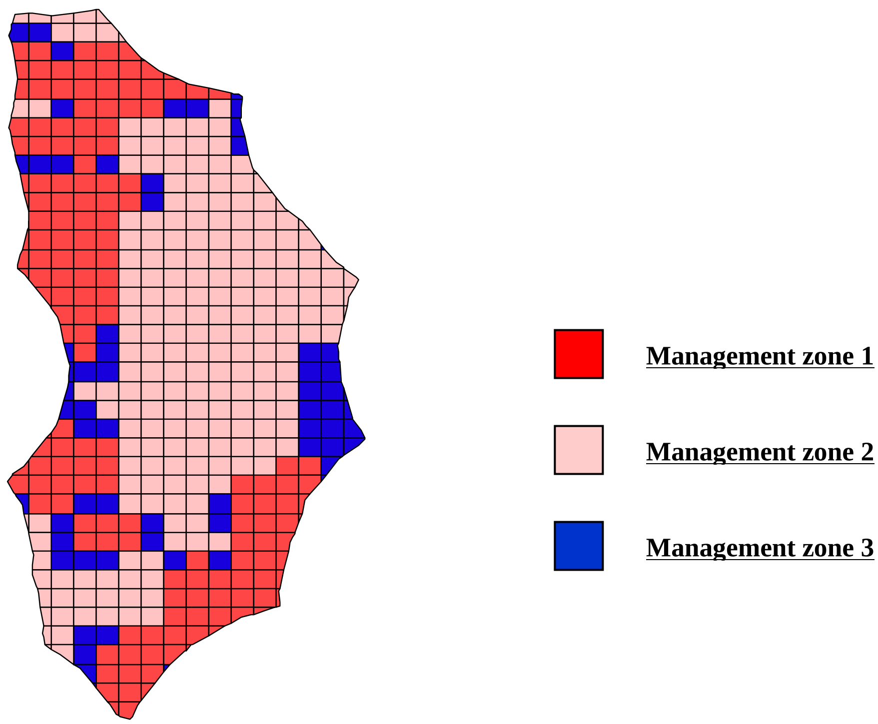

Figure 2. Delineated Management Zones from the ESDA Procedure