

Optimización bajo incertidumbre en Agricultura.

Víctor M. Albornoz S.

Departamento de Industrias, Campus Santiago Vitacura.
Universidad Técnica Federico Santa María.

Santiago de Chile, 26 de Mayo de 2022

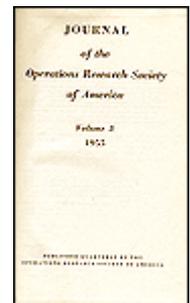
Agenda

1. Introducción.
2. Problema de zonas de manejo agrícola.
3. Optimización bajo incertidumbre.
4. Métodos de descomposición.
5. Palabras finales.

1. Introducción.

Numerosos problemas en agricultura han sido abordados mediante el uso de modelos de optimización.

A partir del trabajo seminal de Thornthwaite (1953) su aplicación ha incluido la planificación y rotación de cultivos, planificación de cosechas, gestión de cadenas de suministro y riego en tiempo real, entre muchos otros.



En los diferentes niveles de decisión surgen problemas complejos que involucran muchas variables interdependientes y un menor o mayor grado de certidumbre en los antecedentes del problema.

La publicación de numerosos trabajos en apoyo a la toma de decisiones en cadenas de suministro agrícola bajo incertidumbre puede revisarse en Borodin et al. (2016).



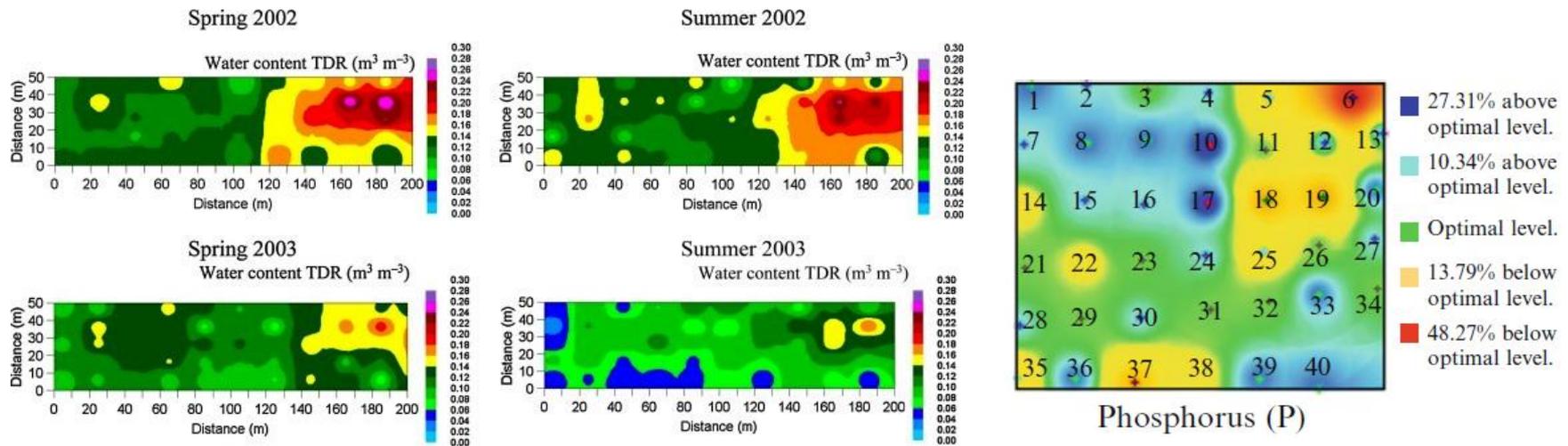
Posteriormente, Behzadi et al. (2018) elaboran una revisión bibliográfica con trabajos en manejo del riesgo en cadenas de suministro agrícola.

En esta presentación mostraremos algunos problemas de la **Agricultura 4.0** y como estos pueden ser abordados igualmente desde la perspectiva de la I.O.



La variabilidad espacial de las propiedades del suelo y los vegetales representan uno de los aspectos más importantes que determina la productividad y calidad de la cosecha obtenida.

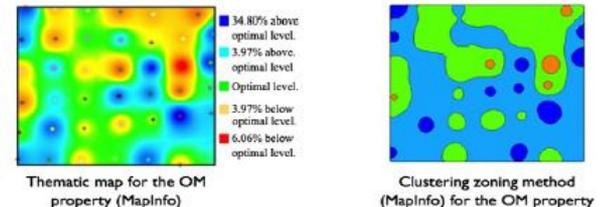
Las siguientes figuras muestran tal variabilidad:



La Agricultura de Precisión propone un manejo eficiente del campo de una manera sitio-específica para responder a dicha variabilidad.

Desde hace 25 años que se emplean métodos de clustering basados en muestras del suelo.

Más recientemente, propusimos dividir un terreno en zonas de manejo rectangulares a partir de un modelo BIP.



La optimización estocástica y apropiadas técnicas algorítmicas de resolución pueden contribuir a resolver numerosos problemas en agricultura y plantean interesantes desafíos de investigación.

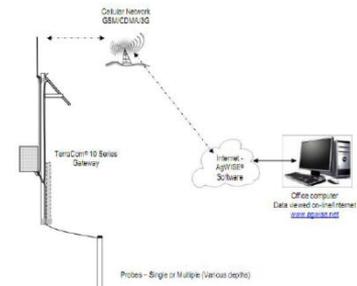
En lo que sigue de esta charla:

- modelo de zonas de manejo
- esquema jerárquico en planificación de cultivos.
- empleo de un método de descomposición.

2. Problema de zonas de manejo agrícola.

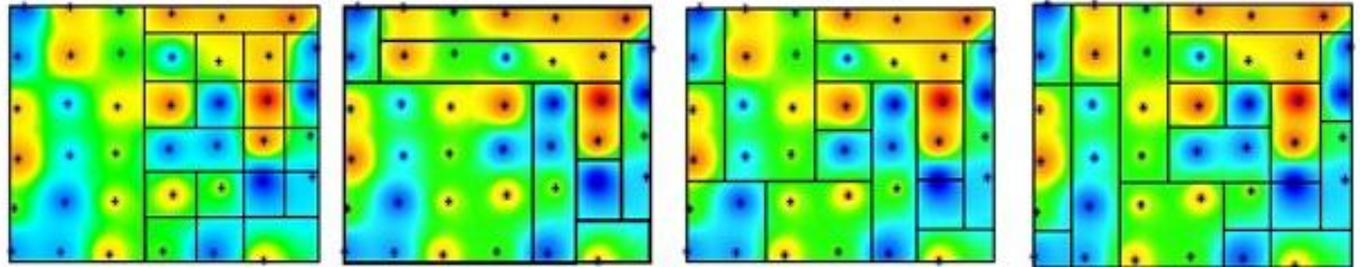
El problema consiste en dividir un terreno en zonas de manejo a partir de datos del suelo o de los vegetales.

En Agricultura 4.0 los datos pueden conseguirse con drones, sensores o robots y la ayuda de IoT y sistemas GIS, disponiendo de grandes volúmenes de datos en apoyo a la toma de decisiones.



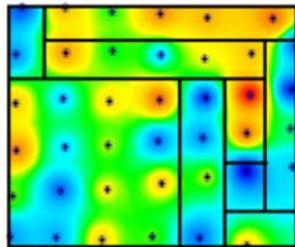
La visualización de los datos, para una propiedad dada, se conoce como un mapa temático.

La siguiente figura muestra el mapa temático de un terreno y su división en zonas de manejo rectangulares en relación al problema que nos ocupa.

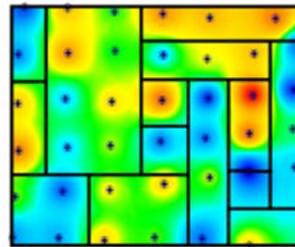


Una medida de la efectividad de un método de zonificación es el concepto de **varianza relativa**.

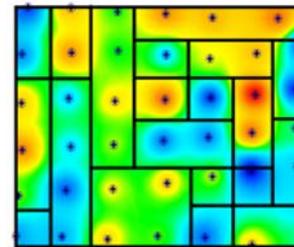
Una alta varianza relativa ($0 \leq RV \leq 1$) implica una alta homogeneidad interna en cada una de las zonas en las cuales se divide un terreno.



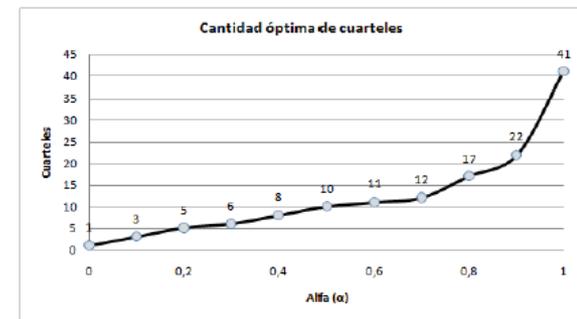
RV=0.5



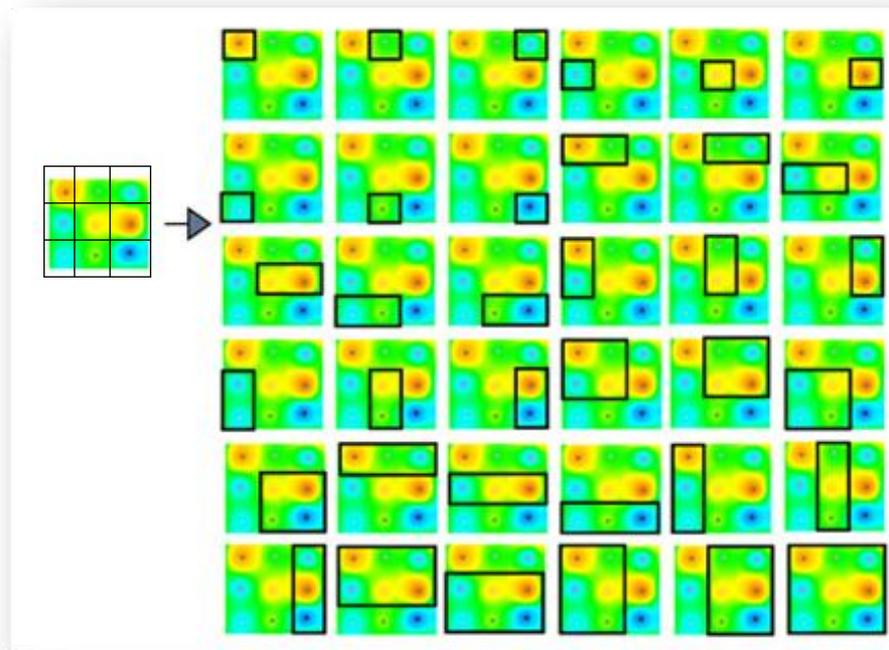
RV=0.7



RV=0.9



En Cid-García et al. (2013) las zonas de manejo son elegidas considerando un conjunto de potenciales zonas de manejo (rectangulares), a partir de una división inicial del terreno en celdas, en cada una de las cuales se dispone de un punto de muestra de la propiedad considerada.



Asumimos dado entonces un conjunto S de **puntos muestrales** asociados a cada celda que conforma el terreno.

Generamos un conjunto de **potenciales zonas de manejo rectangulares** M a partir de un cierto tamaño mínimo y máximo de celdas dado.

Los datos previos permiten definir un parámetro $a_{sz}=1$ **si la zona z incluye el punto muestral s y**
 $a_{sz}=0$ **sino**

El modelo de programación entera binaria propuesto emplea la siguiente notación:

Parámetros.

σ_z^2 = varianza de la propiedad en cada zona $z \in M$.

n_z = número de muestras en la potencial zona $z \in M$.

σ_T^2 = varianza de la propiedad en todo el terreno.

α = nivel de varianza relativa impuesta a la partición.

Variables de decisión.

$q_z = 1$ si la potencial zona de manejo $z \in M$ pertenece a la *partición* y 0 sino,

Modelo de Programación Entera Binaria (BIP):

$$\begin{aligned} \min \quad & \sum_{z \in Z} q_z \\ \text{s.t.} \quad & \sum_{z \in Z} a_{sz} q_z = 1 \quad s \in S \\ & 1 - \frac{\sum_{z \in Z} (n_z - 1) \sigma_z^2 q_z}{\sigma_I^2 [N - \sum_{z \in Z} q_z]} \geq \alpha \\ & q_z \in \{0, 1\} \quad z \in Z \end{aligned}$$

Al considerar instancias de diferentes tamaños, los resultados en la siguiente tabla muestran lo alcanzado con Cplex 12.4 en un Intel core i5 de 2.5 GHz con 8 GB de RAM, incluido el tiempo de generación del conjunto de zonas potenciales:

Instance	N	K	Time[s]	Optimal Value
1	42	588	0.48	10
2	100	3,025	11.52	22
3	150	6,600	106.14	24
4	225	14,400	1,468.62	34
5	300	25,200	8,761.71	47
6	400	44,100	47,654.8	58
7	500	68,250	-	-
8	600	97,650	-	-
9	750	151,125	-	-
10	900	216,225	-	-

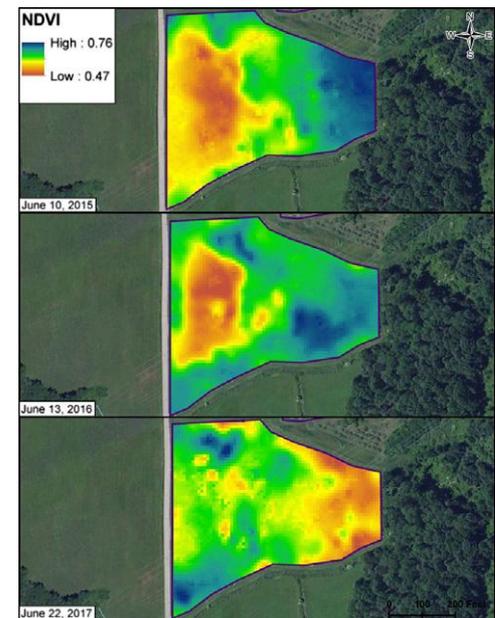
La zonificación pudo ser aplicada en un conjunto de instancias en el contexto de un problema de planificación táctica de cosecha, basado en el NDVI como propiedad de los vegetales, obteniendo una reducción promedio de un 9.5% en los costos totales:

Instancia	S	Z	Zonas Dadas	Costo Cosecha	Z*	Costo Cosecha con Zonificación
1	80	1980	16	8406590	26	8151850
2	100	3025	20	10259560	33	10314700
3	120	4290	24	11658140	36	11973200
4	140	5575	28	14902000	38	13411600
5	160	7480	32	15826000	38	14765900
6	180	9405	36	14745200	42	13060100
7	200	11550	40	15734700	45	14879600
8	220	13915	44	18774400	46	15209600
9	240	16500	48	18574400	48	16134700
10	260	19305	52	20674500	47	17298900

3. Optimización bajo incertidumbre.

La variabilidad temporal de la propiedad empleada representa un aspecto importante al momento de determinar la productividad y calidad de la cosecha obtenida.

Lo anterior lleva a considerar que el indicador pueda ser incorporado en el modelo a través de distintos escenarios, dando origen a un modelo bajo incertidumbre.



Esta y otras situaciones llevan a considerar la formulación y resolución de modelos de optimización que incluyen explícitamente la presencia de **parámetros bajo incertidumbre**.

Bajo el nombre de *Programación Estocástica* se reúne una parte importante de la literatura referente a la incorporación de variables aleatorias en la toma de decisiones óptimas.

Un modelo con recurso de dos-etapas para el problema de zonas de manejo considera la siguiente notación adicional (parámetros):

Ω = conjunto de escenarios para la propiedad empleada.

p_ω = peso de cada escenario $\omega \in \Omega$.

$\sigma_{z\omega}^2$ = varianza de la propiedad en $z \in M$ bajo escenario $\omega \in \Omega$.

σ_ω^2 = varianza de la propiedad del terreno en escenario $\omega \in \Omega$.

M = un costo de penalización.

Variables de decisión de primera etapa:

$q_z = 1$ si la potencial zona de manejo $z \in M$ es parte de la partición y 0 sino,

Variables de decisión de *recurso*:

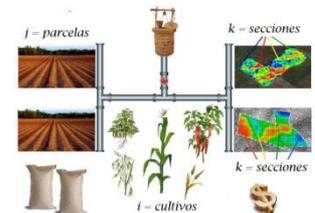
s_ω = variable que penaliza incumplimiento de la varianza relativa de acuerdo al escenario $\omega \in \Omega$

Modelo bajo incertidumbre para la definición de zonas de manejo :

$$\begin{aligned}
 & \text{Min } \sum_{z \in Z} q_z + M \sum_{\omega \in \Omega} p_{\omega} s_{\omega} \\
 & \text{s.t. } \sum_{z \in Z} a_{sz} q_z = 1, \quad s \in S \\
 & 1 - \frac{\sum_{z \in Z} (n_z - 1) \sigma_{z\omega}^2 q_z}{\sigma_{T\omega}^2 (N - \sum_{z \in Z} q_z)} \geq \alpha - s_{\omega}, \quad \omega \in \Omega \\
 & q_z \in \{0, 1\} \quad \text{for all } z \in Z, \quad s_{\omega} \geq 0 \quad \text{for all } \omega \in \Omega
 \end{aligned}$$

En Albornoz et al. (2020) se emplea el modelo descrito en un problema de planificación de rotación de cultivos en las zonas seleccionada, empleando un esquema jerárquico.

El modelo minimiza los costos del plan propuesto para satisfacer las demandas, considerando los rendimientos futuros bajo condiciones de incertidumbre.



Variables de decisión del modelo son:

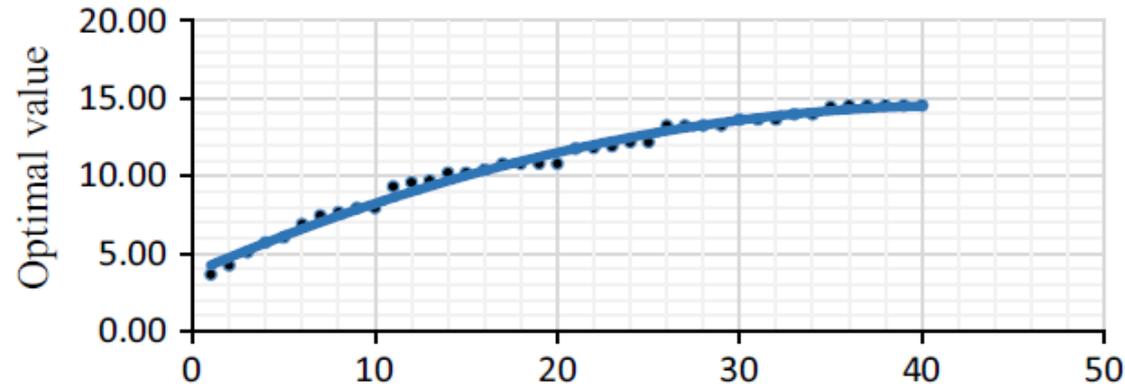
$x_{kz} = 1$ si en la zona z (ya definida) se siembra la rotación k
 0 sino.

$y_{it\omega} =$ demanda insatisfecha del cultivo i en el
 período t bajo el escenario $\omega \in \Omega$

Modelo de selección y rotación de cultivos:

$$\begin{aligned} & \text{Min} \sum_{k \in K} \sum_{z \in Z^*} c_{kz} x_{kz} + \sum_{\omega \in \Omega} p_{\omega} \left(\sum_{i \in I} \sum_{t \in T} f_{it} y_{it\omega} \right) \\ \text{s.t.} \quad & \sum_{k \in K} \sum_{z \in Z^*} r_{it\omega}^{kz} x_{kz} + y_{it\omega} \geq d_{it}, \quad i \in I; \quad t \in T, \omega \in \Omega \\ & \sum_{k \in K} x_{kz} = 1, \quad z \in Z^*, \\ & x_{kz} \in \{0, 1\} \quad k \in K \text{ and } z \in Z^*; \quad y_{it\omega} \geq 0 \quad i \in I; t \in T, \omega \in \Omega. \end{aligned}$$

En el artículo se muestran resultados con datos de 35 escenarios empleando AMPL/CPLEX 12.4



Instance	N	$ Z $	$ Z^* $	Average value of s_ω	Average penalty cost	Optimal value	Time (s)
1	50	825	8	0.130	6.485	14.485	39.000
2	70	1540	8	0.115	5.760	13.760	84.719
3	100	3025	8	0.117	5.827	13.827	114.375
4	120	4290	8	0.077	3.840	11.840	553.907
5	140	5775	8	0.088	4.403	12.403	289.797
6	160	7480	7	0.110	5.522	12.522	776.656
7	180	9405	6	0.091	4.572	10.572	546.969
8	200	11,550	9	0.093	4.663	13.663	652.828
9	220	13,915	8	0.074	3.068	11.680	1093.670

4. Métodos de descomposición.

Existen problemas altamente estructurados y de gran tamaño que contribuyen a la toma de decisiones y cuya resolución plantea desafíos algorítmicos por su complejidad.

$$\begin{bmatrix} A_1 & A_2 & A_3 & \cdots & A_q \\ D_1 & & & & \\ & D_2 & & & \\ & & D_3 & & \\ & & & \cdots & \\ & & & & D_q \end{bmatrix}$$

Los trabajos seminales en algoritmos como los llamados métodos de descomposición son:

Dantzig & Wolfe Decomposition (1960).

Column Generation (Gilmore and Gomory, 1961)

Benders Decomposition (1962).

Para abordar computacionalmente un problema mediante el uso de un Método de Descomposición la idea básica consiste en:

descomponer o simplificar la resolución del problema original resolviendo separadamente un Problema Maestro (Reducido) y un Subproblema, enviando información entre ambos hasta alcanzar la solución óptima o una aproximación a la misma en un número finito de iteraciones.

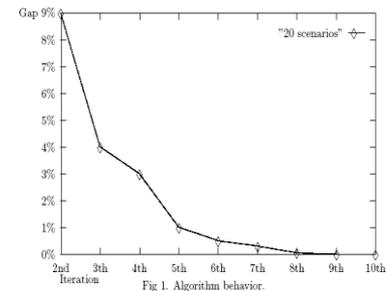
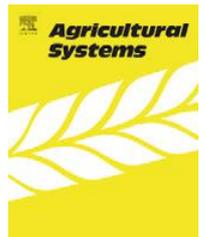


Fig 1. Algorithm behavior.

En problemas de agricultura, Ahumada et al. (2012) aplicaron el Método de Descomposición de Benders, donde:

Problema Maestro. Encuentra un plan de siembra que minimiza los costos de los cultivos para productores de vegetales.

Subproblema. maximiza el ingreso esperado conocida la realización de parámetros inciertos asociados a rendimientos y demandas futuras.



Costa et al. (2014) aplicaron *Generación de Columnas* en cultivos de vegetales, donde:

Problema Maestro. Determina una rotación óptima de cultivos para diferentes productos en un conjunto de terrenos dados de modo de satisfacer las demandas por los mismos.

Subproblema. Provee una nueva rotación de cultivos.



Albornoz y Ñanco (2016) aplicaron *Generación de Columnas* a un problema determinista de agricultura de precisión, donde:

Problema Maestro. Encuentra una partición óptima de un terreno a partir de un conjunto de potenciales zonas de manejo agrícola.

Subproblema. Provee una nueva zona de manejo (rectangular).



Escallón et al. (2020) aplicaron *Generación de Columnas* en la planificación de cosecha, donde:

Problema Maestro. Maximiza los ingresos generados ante distintos patrones posibles de cosecha de las zonas que componen el terreno.

Subproblema. Provee un nuevo patrón de cosecha de las distintas zonas.



Más recientemente, Albornoz y Zamora (2021) aplican Generación de *Columnas* (y filas) a un problema de agricultura orgánica, donde:

Problema Maestro. Encuentra una partición óptima de un terreno en zonas de manejo, asigna cultivos de vegetales e impone restricciones de adyacencia (entre familias botánicas).

Subproblema. Provee una nueva zona de manejo (rectangular).

First period			Second period			Third period		
1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6
7	8	9	7	8	9	7	8	9

En A. y Nanco (2016) la estrategia adoptada fue implementada en AMPL y resuelta con Cplex 12.4 en un IntelCore i5 de 2.5 GHz con 8 GB de RAM, que para su resolución directa obtuvo los siguientes resultados, cuyos tiempos de ejecución incluye la obtención de todas las zonas potenciales:

Instance	N	K	Time[s]	Optimal Value
1	42	588	0.48	10
2	100	3,025	11.52	22
3	150	6,600	106.14	24
4	225	14,400	1,468.62	34
5	300	25,200	8,761.71	47
6	400	44,100	47,654.8	58
7	500	68,250	-	-
8	600	97,650	-	-
9	750	151,125	-	-
10	900	216,225	-	-

La implementación del método muestra en cambio:

Instance	#col	UB	LB	% gap	IS	ISLB	time [s]
1	188	8.89	8.89	4.50E-14	10	9	1.01
2	636	21.08	20.86	1.05	22	21	5.31
3	981	23.77	23.77	9.71E-14	24	24	17.17
4	1,700	33.23	33.23	1.87E-13	34	34	66.23
5	2,111	46.5	45.71	1.69	47	46	146.47
6	3,229	57.52	57.21	0.53	58	58	429.56
7	3,852	76.51	76.51	1.83E-13	77	77	743.093
8	5,160	84.56	84.56	2.93E-13	85	85	1,484.66
9	7,061	106.59	105.21	1.3	107	106	3,395.15
10	8,794	124.33	124.33	6.00E-13	125	125	6,061.11

Posteriormente, Albornoz et al. (2020) resuelven diferentes instancias del modelo de zonificación bajo incertidumbre en un PC Intel Core i5 6200U de 2.3Ghz de 8GB RAM con AMPL/CPLEX 12.7.

Table 1: Small size instances.

#	N	$ Z $	RO	$\sum_{z \in Z} q_z^*$	Time (s)
1	80	1980	30.438	28	28.8
2	100	3025	38.677	33	55.3
3	120	4290	46.437	39	93.1
4	140	5775	51.212	44	143.8
5	160	7480	50.701	46	217.8
6	180	9405	49.458	45	308.9
7	200	11550	48.594	43	411.5
8	220	13915	45.015	40	544.4
9	240	16500	43.749	38	680.6
10	260	19305	44.695	41	1146.4

Table 2: Results for small size instances using the strategy based on CG.

#	$ Z_{RMP} $	RMP	$\sum_{z \in Z} q_z^*$	Time (s)
1	319	30.438	28	1.5
2	473	38.677	33	3.0
3	624	46.437	39	5.6
4	841	51.212	44	11.1
5	910	51.221	45	12.3
6	1064	50.341	46	20.9
7	1412	49.228	43	31.0
8	1721	45.393	37	51.4
9	2074	44.334	38	75.6
10	2246	45.260	40	95.1

Table 4

Medium size instances.

#	$N(n \cdot m)$	$ Z $	RO	$\sum_{z \in Z} q_z^*$	Time (s)
11	320 (32 · 10)	29040	39.708	38	1673.8
12	340 (34 · 10)	32725	38.368	38	1715.2
13	360 (36 · 10)	36630	37.548	37	2501.4
14	380 (38 · 10)	40755	36.000	36	2742.3
15	400 (40 · 10)	45100	35.276	35	3077.5
16	420 (42 · 10)	49665	35.000	35	3438.8
17	440 (44 · 10)	54450	35.001	35	4707.5
18	460 (46 · 10)	59455	34.626	34	4967.3
19	480 (48 · 10)	64680	34.000	34	5260.0
20	500 (50 · 10)	70125	35.000	35	5785.2

Table 5

Results for medium size instances using the strategy based on CG.

#	$ Z_{RMP} $	RMP	$\sum_{z \in Z} q_z$	Time (s)
11	3019	40.259	39	221.7
12	3409	38.459	38	334.6
13	3789	38.000	38	418.5
14	3947	36.379	36	447.6
15	4651	42.407	39	702.0
16	4530	35.091	35	799.8
17	4557	35.577	35	972.2
18	4856	35.000	35	912.5
19	4971	34.299	34	1063.9
20	5529	38.000	38	1272.7

Table 6

column generation results for large size instances.

#	$N(n \cdot m)$	$ Z $	$ Z_{RMP} $	RMP	$\sum q_z$
21	550 (25 · 22)	84700	4962	53.000	53
22	600 (25 · 24)	100650	5902	54.536	54
23	650 (25 · 26)	117975	6207	58.324	58
24	700 (25 · 28)	136675	6479	65.000	65
25	750 (25 · 30)	156750	7189	71.000	71
26	800 (32 · 25)	178200	8411	54.455	54
27	850 (34 · 25)	201025	9464	52.489	52
28	900 (36 · 25)	225225	9244	53.742	53
29	950 (38 · 25)	250800	10461	51.000	51
30	1000 (40 · 25)	277750	11356	50.884	50



Original papers

Delineating robust rectangular management zones based on column generation algorithm

Victor M. Albornoz^a, Linco J. Nanco, José L. Sáez^a Universidad Técnica Federico Santa María, Departamento de Ingeniería, Campus Santiago Vitacura, Av. Santa María 6400, Santiago, Chile

ARTICLE INFO

Keywords:
Precision agriculture
Management zones
Robust optimization
Column generation
OR in agriculture

ABSTRACT

This paper considers a management zone delineation problem in a precision agriculture framework that defines a partition of a field into rectangular and homogeneous management zones according to certain vegetation or soil indexes. Modeling the problem as a robust optimization model accounts for spatial and temporal variability of the index considered, representing this variability according to a finite set of sampled realizations (scenarios). The model assumes that the complete enumeration of all the possible rectangular management zones is known. To deal with this assumption, we propose a strategy based on a column generation algorithm to exploit the structure of the model. To solve the model and implement the strategy, an algebraic modeling language and state of the art MIP software was used. Results from the modeling approach and the computational strategy were obtained by solving different instances of a real case-study. We show that the robust optimization model can be useful in real situations under temporal variability of the index used for delineating management zones. Results also show that the algorithmic strategy is appropriated for solving the instances considered.

5. Palabras finales.

Existen numerosos problemas de toma de decisiones en agricultura que pueden ser abordados con modelos de optimización y la selección de zonas de manejo sitio-específicas.

La naturaleza de los problemas abordados puede llegar a requerir el uso de modelos bajo incertidumbre y estrategias algorítmicas de resolución.

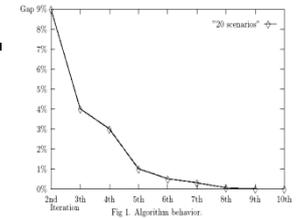
Homogeneous.



Heterogeneous.



El Método de Generación de Columnas provee una técnica numérica de optimización eficiente para resolver problemas de gran tamaño.



En general, los resultados mostrados permiten apreciar la contribución que todas estas metodologías pueden llegar a tener en la agricultura moderna y sustentable de nuestros días.

Referencias:

Cid-García, N., Albornoz, V., Ortega, R. and Ríos-Solís (2013). Rectangular shape management zone delineation using integer linear programming. *Computers & Electronics in Agriculture* 93, 1-9.

Albornoz, V.M., Ñanco, L.J (2016). An empirical design of a column generation algorithm applied to a management zone delineation problem. In *Computational Management Science. State of the Art 2014*. Fonseca, R.J., Weber, G.W. and Telhada, J. (Eds). *Lecture Notes in Economics and Mathematical Systems* 682, 201-208, Springer.

Albornoz, V.M., Ñanco, L.J, Sáez, J.L. (2019). Delineating robust rectangular management zones based on column generation algorithm. *Computers and Electronics in Agriculture* 61, 194-201.

Albornoz, V.M. Véliz, M.I, Ortega, R., Ortíz-Araya, V. “*Integrated versus hierarchical approach for zone delineation and crop planning under uncertainty*”. *Annals of Operations Research* 286(1-2), 617–634, 2020.

Albornoz, V.M., Zamora, G.E. “*Decomposition-based heuristic for the zoning and crop planning problem with adjacency constraints*”. *TOP* 29, 248–265, 2021.

Optimización bajo incertidumbre en Agricultura.

Víctor M. Albornoz S.

Departamento de Industrias, Campus Santiago Vitacura.
Universidad Técnica Federico Santa María.

Santiago de Chile, 26 de Mayo de 2022