

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INDUSTRIAS
SANTIAGO - CHILE



Modelo de Localización de Sensores para la Agricultura de Precisión

TIARE BELÉN TORRES HERRERA

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA
Y AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

PROFESOR GUÍA : SR. Víctor Albornoz S.
PROFESORES CORREFERENTES : SR. Rodrigo Ortega B.
SR. David Godoy
SR. Alejandro Mac Cawley

15 de noviembre de 2023

Este trabajo está dedicado con profundo cariño a mis abuelos. Aunque ya no están físicamente a mi lado, nunca han dejado de acompañarme.

Quiero expresar mis sinceros agradecimientos a todas las personas que han sido parte de este proceso. En primer lugar a mi hermana Javiera, por ser un pilar fundamental en todos los aspectos de mi vida, en especial por su contención cuando más lo he necesitado.

A mis padres su apoyo incondicional, por entregarme los valores que hoy me forman como persona y como profesional, y por enseñarme la importancia de la educación.

A mis queridas compañeras de cuatro patas, Perla y Haru, por su compañía durante días y noches de estudio y por alegrar cada momento.

A mis amigos, me siento agradecida por haber tenido la fortuna de recorrer este camino junto a ustedes, desde los primeros días de nuestra formación hasta hoy, que nos acercamos al final de esta etapa haciendo cada momento memorable.

A mis profesores, por su dedicación y enseñanzas entregadas a lo largo de estos años. En especial al profesor Victor Albornoz, por confiar en mí y guiarme hasta el final de esta etapa.

En resumen, a todas las personas que han parte de mi vida durante mi formación, les agradezco profundamente por su influencia positiva y su apoyo constante en mi camino académico y personal. Su contribución ha sido invaluable y ha dejado una huella perdurable en esta etapa de mi vida.

RESUMEN EJECUTIVO

La agricultura de precisión es una disciplina que ha logrado aumentar la eficiencia, calidad y productividad de los cultivos agrícolas, así como también, disminuir el impacto ambiental de los mismos a través de la aplicación de tecnologías y análisis de datos. Por esta razón, la recopilación de la data es un paso clave en la implementación de estas tecnologías. Para ello, en el mercado existe una amplia variedad de sensores tanto fijos como móviles, cada uno con objetivos específicos y un amplio rango de precios.

Usualmente, los datos se recopilan de manera puntual y se utilizan técnicas de interpolación para poder tener una muestra en términos continuos del índice de interés, es por ello, que en este trabajo se busca determinar dónde localizar un número limitado de sensores de manera de aumentar la representatividad de la muestra obtenida.

A través de dos modelos de optimización lineal y en variables enteras binarias se busca encontrar las ubicaciones óptimas para un número limitado de sensores y además determinar cuántos de ellos es necesario instalar para obtener un muestreo eficiente.

Ambos modelos son comparados en un caso de estudio resuelto a través del Solver Gurobi Optimizer en Julia JuMP.

Palabras clave. Investigación de operaciones en la agricultura, zonas de manejo, modelo de localización p-mediana, programación entera.

ABSTRACT

Due to the benefits of using precision agriculture, many farmers want to install sensors in their fields. However, the high cost that a sensor could achieve limit the number of these to be acquired and to interpolate the information obtained to the whole extension of the field. For this reason it is interesting to determine where to locate a limited number of sensors, increasing the representativeness of the sample and, consequently, the reliability of the interpolation.

Therefore, this paper addresses an optimization–based approach that integrates the management zone problem and a discrete location problem. The input data of the proposed model assumes a set of sample points of a soil or crop property to find an optimal location of a given number of sensors to obtain many others soil or crop properties in the resulting locations, and by interpolating the information obtained, reliable data about new or future properties can be obtained.

The proposed model is solved using the integer programming solver Gurobi Optimizer in Julia JuMP and computational results from a set of instances are presented to show the impact of the adopted methodology.

Keywords. OR in agriculture, management zones, p-median location problem, integer programming.

Índice de Contenidos

1. Introducción	1
1.1. Antecedentes preliminares del problema	1
1.2. Objetivos	3
1.2.1. Objetivos Generales	3
1.2.2. Objetivos Específicos	3
1.3. Alcances y Límitaciones	4
1.4. Organización del trabajo	4
2. Marco Teórico	5
2.1. Zonas de manejo	5
2.2. Problemas de Localización	8
3. Metodología	11
3.1. Generación de zonas potenciales	11
3.2. Esquema Jerárquico	12
3.2.1. Modelo de zonificación	12
3.2.2. Modelo de localización	15
3.3. Modelo Integrado	17
4. Caso de estudio y resultados	20
4.1. Implementación Computacional	20
4.2. Análisis Gráfico	26
4.3. Relajación en el límite superior LS	30
4.4. Costo Computacional	31
5. Conclusiones	33
Bibliografía	35
6. ANEXOS	38
6.1. Algoritmo 1	38

6.2. Implementación esquema jerárquico	41
6.3. Implementación modelo integrado	44
6.4. Datos	46
6.5. Resultados implementación computacional	47



Índice de Tablas



4.1.	Sección de 6x10 puntos muestrales de la data de NDVI ampliada	21
4.2.	Resultados para ambos modelos con diferentes valores de p	23
4.3.	Costo computacional de ambos modelos para distintos tamaños de muestras	31
6.1.	Sección de 40 puntos de la muestra utilizada para implementación computacional	46
6.2.	Resultados de ambos modelos para una instancia de 80 puntos muestrales	47

Índice de Figuras

3.1. Conjunto de zonas potenciales en una muestra de 9 puntos	12
4.1. Valores del NDVI y su interpretación según el estado de la vegetación	21
4.2. Mapa de calor realizado a partir de 121 puntos muestrales del índice NDVI.	22
4.3. Resultados de ambos modelos para una instancia de 42 puntos muestrales	24
4.4. Resultados de ambos modelos para una instancia de 60 puntos muestrales	25
4.5. Resultados de ambos modelos para una instancia de 80 puntos muestrales	25
4.6. Solución gráfica del modelo de zonificación para dos tamaños de muestras	27
4.7. Localización de 3 sensores utilizando el esquema jerárquico para distintos tamaños de muestras	28
4.8. Localización de 3 sensores utilizando el modelo integrado para distintos tamaños de muestras	28
4.9. Relajación del Límite superior	30

1 — Introducción

1.1. Antecedentes preliminares del problema

Hoy en día, la agricultura es una de las industrias más importantes en todo el mundo, por su aporte en el PIB y crecimiento económico de los países, por su rol protagónico en la obtención de materias primas para otras industrias como la forestal, textil, ganadería, entre otras. También en la creación de empleos, donde según cifras del ministerio de Agricultura la fuerza laboral agrícola se ha mantenido invariante representando un 9,2 % de la tasa de ocupación total del país (Domínguez et al. [2019]). Por último, como actor principal en el abastecimiento de alimentos para todos los seres vivos del planeta, donde sólo considerando a los seres humanos, se tiene una demanda para la alimentación de 8 billones de personas, cifras que según Zeifman et al. [2022] se encuentra con la tasa de crecimiento más alta en décadas, en consecuencia, la agricultura se ve obligada a crecer en números similares e incluso, superiores.

Dados los antecedentes expuestos, la Agricultura de Precisión (PA) es una disciplina muy relevante en la actualidad, su objetivo es incrementar la eficiencia, productividad y al mismo tiempo, disminuir el impacto ambiental de las prácticas agrícolas. A través de la aplicación de tecnologías y métodos de análisis de datos brindan a los agricultores información para aplicar nutrientes y fertilizantes en sitios y cantidades específicas según la necesidad de los cultivos, también permite realizar análisis económicos y de riesgo basado en la variabilidad temporal y espacial de las propiedades medidas (Zhang et al. [2002]), permite detectar oportunidades y amenazas formando un sistema de apoyo a la toma de decisiones para un manejo de las producciones agrícolas (Albornoz [2023]).

El primer paso para la implementación de técnicas de PA es la recolección de los datos, esto se puede realizar a través de sensores que permiten medir diferentes propiedades del suelo. Estos

sensores pueden estar en maquinarias utilizadas de forma móvil en los predios, existen bandas que a través de sensores infrarrojo permiten extraer información, sensores referenciados con GPS e imágenes satelitales y también sensores que van fijos en el suelo y permiten obtener información a tiempo real de diferentes propiedades. Este último tipo, mide de forma local en un punto del suelo, y una técnica muy utilizada en su uso es la interpolación, entonces se obtienen mediciones para puntos específicos del suelo de forma discreta y el tratamiento de datos a través de sistemas de información geográficas (por ejemplo QGIS) permiten obtener información de forma continua para todos los puntos del terreno en estudio. Por esta razón, resulta interesante encontrar una forma de localizar los sensores aumentando la representatividad de la información obtenida, y a la vez, intentar usar la menor cantidad posible, intentando tener muestras eficientes a menor costo.

En este sentido, la información del suelo que se puede obtener y tiene relevancia para agricultores corresponde a las propiedades físicas y químicas del suelo, como el Fósforo (P), Potasio (K), Calcio (Ca), Magnesio (Mg), Materia Orgánica, entre otros (Foth et al. [1988]), donde a partir de estas propiedades se puede analizar su variación temporal y espacial. Además, se pueden considerar las propiedades del cultivo representadas a través de índices como el Índice de Vegetación Verde (GVI) y el Índice de Vegetación de Diferencia Normalizado (NDVI). Estas propiedades son obtenidas a través de bandas de medición infrarroja (NIR) y bandas espectromagnéticas, que miden intensidad y frecuencia de la radiación que la vegetación emite (o refleja), su valor da una medición acerca de la salud de los cultivos, permite determinar la tasa de crecimiento, tomando en cuenta efectos climáticos como sequía, calor, heladas y por último, estima rendimiento y calidad de estos (Albornoz et al. [2022]).

Algunos índices son más simples de medir y se puede contar con los datos sin la necesidad de acceder a sofisticadas tecnologías, sin embargo, no contienen toda la información necesaria para el manejo de los cultivos. Por esta razón, se propone utilizar una propiedad ya medida del suelo para obtener una representación de la variabilidad, que permita localizar de manera óptima sensores complejos, dispositivos más costosos que permiten obtener otra información de interés del suelo y de los cultivos.

Las zonas de manejo son subregiones de un terreno, donde cada una de ellas es homogénea y entre ellas son heterogéneas. Estas zonas permiten la aplicación de prácticas agronómicas eficientes como el manejo específico para la siembra, aplicación de fertilizantes, nutrientes, riego, entre otros, de los cuales cada zona necesita una cantidad específica de insumos (Roudier et al. [2008]). De esta forma, la delimitación de zonas de manejo aumenta la eficiencia de los cultivos, aumentando la calidad y productividad al mismo tiempo que se reduce el impacto ambiental de las prácticas agrícolas.

En este sentido, al tratarse de zonas homogéneas y además la medición de sensores en puntos específicos se puede interpolar para obtener información del terreno de forma continua, resulta interesante dividir el terreno en zonas de manejo de manera en que cada zona pueda ser estimada por sólo un sensor.

1.2. Objetivos

El objetivo de esta investigación es proponer una metodología basada en el uso de modelos de optimización en variables enteras - binarias, que permitan resolver el problema de localización óptima para un número limitado de sensores, aumentando la representatividad de la muestra obtenida. Para ello, se presentarán dos alternativas, ambas utilizarán información de propiedades del suelo conocidas previamente y georreferenciadas. Con la implementación se encontrarán las ubicaciones de sensores fijos en el suelo, que posteriormente medirán propiedades específicas de interés para algún cultivo. Ambas alternativas serán comparadas en un caso de estudio, observando los resultados obtenidos, beneficios y consideraciones de cada una.

1.2.1. Objetivos Generales

Proponer una metodología basada en modelos de optimización lineal, que permita encontrar la ubicación óptima de un número limitado de sensores para la medición de propiedades físicas de terrenos agrícolas, de manera que al interpolar las mediciones se obtenga información representativa del suelo en toda su extensión.

1.2.2. Objetivos Específicos

- Aplicar el modelo de zonificación para encontrar una partición que permita dividir el terreno en zonas homogéneas, las que serán utilizadas como los puntos factibles para un modelo de localización discreto.
- Aplicar un modelo de localización del tipo p-mediana para encontrar la ubicación óptima de sensores utilizados en la agricultura de precisión .
- Encontrar un número de sensores óptimos a localizar para obtener una muestra representativa.
- Implementar y comparar dos modelos en un caso de estudio para demostrar su eficiencia y beneficios en la solución del problema presentado.

1.3. Alcances y Límites

El alcance del presente estudio se enfoca en la formulación, implementación y análisis de dos esquemas de solución al problema presentado, basados en modelos de programación lineal en variables binarias.

La metodología propuesta se analiza en un caso de estudio aplicado a puntos muestrales obtenidos de un terreno ubicado en el sur de Chile, resolviendo ambos esquemas se comparan sus fortalezas y limitaciones. La implementación se ve limitada por capacidades de procesador del equipo utilizado dado que la resolución de ambos modelos se realiza de manera directa utilizando el método de Branch and Cut, este método es aplicado por defecto en problemas con variables enteras binarias, y en este trabajo no se exploran otras metodologías de optimización.

En esta oportunidad no se aborda la posibilidad de implementar un sistema de apoyo a la toma de decisiones cómo lo sería un software con una interfaz de usuario, más bien se aborda una solución al problema de localización de sensores desde el punto de vista de la programación lineal.

1.4. Organización del trabajo

A continuación, el capítulo 2 presenta una revisión de la literatura y marco teórico que sustentan la investigación. En el capítulo 3 se revela la metodología para resolver el problema presentado, a través de dos modelos. La sección 4 muestra la implementación computacional de ambos modelos, demostrando su efectividad, implicancia, beneficios y desafíos. Finalmente, en el capítulo 5 se encuentran las conclusiones y desafíos futuros. Además, en la sección 6 se incluyen como anexos los códigos utilizados en Julia JuMP y Gurobi para la implementación computacional.

2 — Marco Teórico

2.1. Zonas de manejo

El manejo de zonas específicas es un área de interés de la agricultura de precisión con alta evidencia en la mejora de los resultados de los cultivos, donde la idea general es caracterizar la variación temporal y espacial de las propiedades del suelo dentro del terreno, identificar los factores que limitan el rendimiento de un cultivo específico y clasificar áreas homogéneas en zonas, de esta forma poder gestionarlas por separado para mejorar la producción y mejorar el uso eficiente de los insumos y los retornos económicos (Khosla et al. [2010]).

En la investigación desarrollada por Thom et al. [2003] se encuentran los primeros esfuerzos por delimitar zonas específicas de un terreno, a través de la determinación de celdas o grillas para subdividir el espacio en celdas más pequeñas. El objetivo general de este trabajo es determinar una técnica de muestreo representativa, al introducir la división en grillas se permite tomar una medición de cada zona en específico y caracterizar de mejor manera la variabilidad espacial. En estudios posteriores, las técnicas de manejo específico son aplicadas a el manejo de fertilizantes y cuidado de las propiedades del suelo, según los requerimientos de cada cultivo. Las herramientas, tecnologías y métodos para delimitar las zonas de manejo se van puliendo, de manera que permita capturar de mejor forma la variabilidad temporal y espacial de los terrenos, o según la necesidad específica de los usuarios.

Para delimitar las zonas de manejo, en la literatura se encuentran varias alternativas estudiadas anteriormente. En el trabajo de Ortega and Santibanez [2007] se presentan tres métodos para determinar zonas de manejo: análisis de clúster, análisis del componente principal y método del coeficiente de variación. En este trabajo, los datos para aplicar el método provienen 6 propiedades del suelo que son de interés para agricultores el pH, la conductividad eléctrica, la materia orgánica

(OM), nitrógeno, fósforo y potasio. Las tres alternativas presentadas, con sus diferencias y beneficios discutidos en el paper, dan como resultado zonas de manejo no regulares que adoptan la forma del cluster, tomando una figura ovalada que imposibilita controlar el tamaño de ellas.

Una segunda alternativa para encontrar las zonas que conforman la partición óptima es a través de modelos de optimización. En el trabajo de [Cid-Garcia et al. \[2013\]](#) se presenta un modelo de optimización lineal en variables entera binarias, que, minimizando la cantidad de zonas en la partición óptima, permite encontrar la combinación que alcanza el nivel de homogeneidad requerido, y a la vez, es posible controlar la cantidad de zonas obtenidas en la partición. Esta alternativa propone la delimitación en zonas de manejo rectangulares y se discuten varias ventajas, entre ellas el uso de maquinarias y aplicación real de los nutrientes, ya que, se vuelve más complejo delimitar zonas con formas irregulares versus una rectangular.

La delimitación de zonas en el modelo requiere garantizar la homogeneidad en cada zona, y que entre ellas (zonas diferentes) sean heterogéneas logrando la partición óptima. Esto se logra a través de la Varianza Relativa (RV), concepto introducido por [Ortega and Santibanez \[2007\]](#).

La varianza relativa es un parámetro agregado que considera la varianza de cada una de las zonas de la partición, y la varianza total del terreno. El parámetro toma valores entre 0 y 1, cuanto más cercano a 1 la partición contiene zonas con mayor homogeneidad y entre zonas distintas la diferencias es máxima, y cuando este se acerca al 0, la partición contiene zonas menos homogéneas.

Se considera un terreno con N puntos muestrales y una partición en Q zonas de manejo, cada zona z incluye n_z de los N puntos muestrales y su varianza se denota por σ_z^2 finalmente, la varianza de todo el terreno σ_T^2 , la varianza relativa se define como:

$$RV = 1 - \frac{\sum_{z=1}^Q ((n_z - 1)/(N - Q)) \sigma_z^2}{\sigma_T^2} \quad (2.1)$$

Este concepto ha sido aplicado en diferentes trabajos para determinar una partición óptima del terreno en zonas de manejo homogéneas. En el trabajo de [Ortega and Santibanez \[2007\]](#) se utiliza de manera directa para maximizar la homogeneidad. También, es aplicado en la investigación de

Cid-Garcia et al. [2013] y posteriores donde se busca determinar la partición a través de un modelo de optimización lineal utilizando una linealización de la expresión 2.1.

Para la zonificación a través de la optimización lineal se requiere el conjunto de datos de todos las zonas potenciales, y a través de los modelos mencionados se escogen aquellas zonas que conforman la partición óptima según el objetivo a optimizar. En el algoritmo 1 descrito por Cid-Garcia et al. [2013] se proponen las instrucciones que generan zonas rectangulares a partir de un conjunto de puntos muestrales georreferenciados, tomando zonas rectangulares con puntos que son colindantes entre si.

Este algoritmo requiere como variables de entrada los puntos muestrales y los parámetros de ancho y largo, mínimo y máximo que se considerarán para cada zona.

Algorithm 1 Algoritmo para generación de zonas potenciales rectangulares (Fuente: Cid-Garcia et al. [2013])

```

Input WidthF, LengthF, MinWidthQ, MinLengthQ, Soil samples
for  $j = \text{MinWidthQ}$  To  $\text{WidthF}$  do
  for  $l = 0$  To  $(\text{WidthF} = 1)$  do
    if  $(j + l) \leq \text{WidthF}$  then
      for  $i = \text{MinLengthQ}$  To  $\text{LengthF}$  do
        for  $k = 0$  To  $(\text{LengthF} = 1)$  do
          if  $(k + i) \leq \text{LengthF}$  then
            Creation of a new quarter
          end if
        end for
      end for
    end if
  end for
end for

```

En estudios posteriores el manejo de zonas rectangulares se ha aplicado a diferentes problemas, en Cid-Garcia and Ibarra-Rojas [2019] proponen un modelo integrado para la delimitación de zonas de manejo y la asignación de cultivos en un horizonte de planificación utilizando un modelo de optimización no lineal. Por otro lado, en Albornoz et al. [2019] se aborda el problema de delimitación de zonas de manejo como un problema de optimización robusto bajo incertidumbre y desarrolla un método eficiente basado en la generación de columnas para resolver el modelo.

En la publicación de Albornoz [2023] se presentan distintas aplicaciones de la zonificación a través de optimización y su rol en el apoyo a la toma de decisiones, de lo mencionado se destaca

el trabajo de [Haghverdi et al. \[2015\]](#) dónde se aplican diferentes métodos de zonificación a la gestión de riego en la agricultura de precisión, incluyendo programación entera lineal y métodos de cluster.

2.2. Problemas de Localización

En la literatura, los problemas de localización se han estudiado desde hace varios siglos. Los primeros registros se observan con el problema propuesto por el matemático Pierre Fermat (1601 - 1665) [Brimberg \[1995\]](#) y resuelto por el matemático Battista Cavalieri (1598 - 1647) quienes buscaban encontrar la ubicación de un punto que minimizaba la suma de la distancia euclidiana entre los vértices de un triángulo y el punto a localizar.

El primer problema de localización de servidores fue introducido por [Weber \[1909\]](#) que puede consultarse igualmente en [Drezner et al. \[2002\]](#) se busca la ubicación que minimiza la suma de los costos de transporte desde el servidor a varios puntos de demanda.

$$\min W(x, y) = \sum_{i=1}^n w_i d_i(x, y) \quad (2.2)$$

Donde $d_i(x, y)$ es la función que representa la distancia entre el servidor localizado en los puntos (x, y) que optimizan la expresión, y a cada uno de los puntos (a_i, b_i) que se buscan abastecer a un costo (o peso) w_i .

Este problema presenta un punto de partida para varios desarrollos posteriores, tanto de variaciones del problema de localización, como aplicaciones en distintas industrias. Actualmente, en la literatura se encuentran dos grandes clasificaciones de los modelos de localización diferenciándose en los datos de entrada para cada uno, en el tipo de decisión que se debe tomar y también en la función a optimizar.

En primer lugar, se encuentran los modelos de *localización continua*, donde los datos de entrada corresponden a una región factible y el (los) punto (s) a localizar se encuentran a través de distancias y pueden tomar cualquier valor dentro de la región dada. La ventaja de este tipo de modelos es que la función a optimizar encuentra siempre un mínimo (o máximo, según corresponda), independiente si en la realidad este punto es o no factible, dado que preliminarmente, la variable

de decisión podría tomar cualquier valor en una región continua, y sólo considera como condición minimizar la distancia a los puntos de demanda.

En segundo lugar están los modelos de *localización discreta* donde los datos de entrada son puntos factibles, con distintas características que permitirán encontrar cuál (cuales) de ellos optimizan una función objetivo. Este tipo de modelos, tiene la ventaja de poder restringir de manera más real los puntos factibles para ubicar las instalaciones. En la industria se encuentran variadas aplicaciones para los modelos de localización discreta, desde localización de negocios, servicios, centros de distribución, y otros.

Dentro de los modelos de localización discreta, se encuentran diferentes clasificaciones de problemas, según el objetivo que se desee resolver, entre ellas: el modelo de *Máxima Cobertura* (Church et al. [1974]) donde la función objetivo buscará maximizar los puntos atendidos con un número limitado de servidores. El modelo de *Cobertura Total* (Toregas et al. [1971]) busca minimizar el número de servidores requeridos para atender a toda la demanda. También, se encuentra el modelo de *P-Mediana* (Hakimi [1964]) donde se busca minimizar el costo global de atender a todos los puntos de demanda con un número limitado de servidores.

El modelo de la P-mediana requiere de dos conjuntos de entrada, los puntos factibles para localizar los servidores, y las ubicaciones de los nodos de demanda, donde cada nodo de demanda tiene un peso, que puede ser representado por su tamaño, importancia, u otro factor, y se calcula la distancia entre todos los nodos factibles y los nodos de demanda. La distancia ponderada por el peso de cada nodo es el factor con el que el modelo buscará la combinación óptima de localización de p servidores. Una variación a este modelo, fue presentada por ReVelle and Swain [1970] donde se utilizó el mismo conjunto de nodos de demanda como nodo factibles para la localización de servidores, bajo la lógica de que los servidores serían instalados en las mismas comunidades que atenderán.

Los modelos de localización tienen múltiples aplicaciones dónde destacan la localización de servicios, plantas, centros de distribución, e incluso instalaciones que no tienen en su objetivo la atención de una demanda, como la localización de maquinarias, sensores, entre otros. Estos modelos buscan entregar la mejor solución para un objetivo específico según el problema, como ejemplo, la disminución de costos o aumentar la cobertura para alguna demanda.

En el trabajo de Ashraf et al. [1997] se desarrolla una metodología basada en la geoestadística

para localizar redes meteorológicas parciales, en este tipo de redes uno de los parámetros, la evotranspiración, es estimada a través de técnicas de interpolación utilizando los datos medidos en otros puntos de la red. Este tipo de implementación simplifica la red y disminuye costos, sin embargo, se requiere de una buena medición para asegurar la confiabilidad del parámetro interpolado. El objetivo de la investigación presentada es determinar las ubicaciones óptimas de las estaciones, aumentando la confiabilidad de la interpolación del parámetro, y con ello obteniendo una muestra tan buena como una red meteorológica total con los beneficios en costos y simplificación de una red meteorológica parcial.

Las aplicaciones de los problemas de localización a la agricultura son variadas, donde se encuentra desde localización de maquinarias, estaciones de riego y/o servicios para trabajadores, localización óptima de cultivos y también centros de distribución para las cosechas. En el trabajo de [Vera et al. \[2003\]](#) se busca localizar K tipos de maquinarias para la cosecha en bosques forestales. Se aplica un modelo de localización lineal del tipo cobertura total, minimizando tanto costos de inversión, fijos y variables de la instalación óptima. Para esta aplicación el terreno se divide en celdas cuadradas de igual tamaño, representando las zonas a ser cosechadas y también los puntos factibles para instalar la maquinaria. Además, se propone un modelo de resolución eficiente a través de la relajación lagrangeana y generación de columnas que permite abordar grandes instancias en cortos tiempos de resolución.

En específico, la localización de sensores ha sido estudiada previamente, como en el trabajo de [Hamouda and Msallam \[2020\]](#) quienes desarrollaron un sistema para la agricultura de precisión con intervalos de muestreo variables (VSI-PA) que permitía encontrar la combinación óptima de intervalos de muestreo a medir para redes de sensores inalámbricos limitando la energía de los nodos muestreados. Este algoritmo, se calcula de forma adaptativa en el desarrollo y se asignan sensores activos a cada área según las necesidades de la planta, mejorando el manejo específico de cultivos, a la vez que intenta disminuir el consumo energético de la red de sensores. Finalmente este modelo, requiere que todos los sensores sean inicialmente instalados, y el algoritmo seleccionará los necesarios de forma iterativa.

En el trabajo de [Uyeh et al. \[2022\]](#) a través de machine learning (K-means++) se desarrolla un algoritmo que permite seleccionar la ubicación óptima de sensores para la temperatura y humedad relativa dentro de invernaderos. La metodología utilizada, a través de cluster de grupos homogéneos, selecciona la ubicación óptima de sensores basado en la distancia euclideana multidimensional. De esta forma, a partir de una gran cantidad de puntos factibles iniciales, el algoritmo agrupa por similitud de resultados. La cantidad de grupos representa el número de sensores necesarios para obtener una muestra de los factores de interés dentro del invernadero.

3 — Metodología

Para resolver el problema de localización de sensores propuesto en las secciones anteriores se proponen dos modelos de programación entera binaria que permiten decidir dónde localizar un número limitado de sensores en un predio. En primer lugar, el terreno se divide en zonas de manejo homogéneas, donde el centroide de cada una es utilizado como punto factible para el modelo de localización discreto p -mediana. La ubicación óptima resulta de la minimización de la distancia ponderada, donde el peso será la varianza de la propiedad utilizada en cada zona de manejo, y la distancia será calculada entre centroides.

3.1. Generación de zonas potenciales

El objetivo de el primer paso es la obtención y tratamiento de datos para todos los conjuntos y parámetros requeridos en los modelos que serán propuestos a continuación.

El Algoritmo 6.1 implementado en Julia JuMP se basa en la secuencia 1 propuesta por [Cid-García et al. \[2013\]](#), con las modificaciones necesarias para el modelo presentado en este estudio.

El algoritmo determina todas las zonas rectangulares potenciales a partir de puntos muestrales aledaños, tal como muestra la figura 3.1. Además, se calcula la varianza de cada una y la ubicación de su centroide. El modelo también calcula la varianza de la zona que se obtiene de considerar todos los puntos muestrales, es decir, de todo el terreno y las distancias euclidianas entre los centroides de todas las combinaciones de zonas.

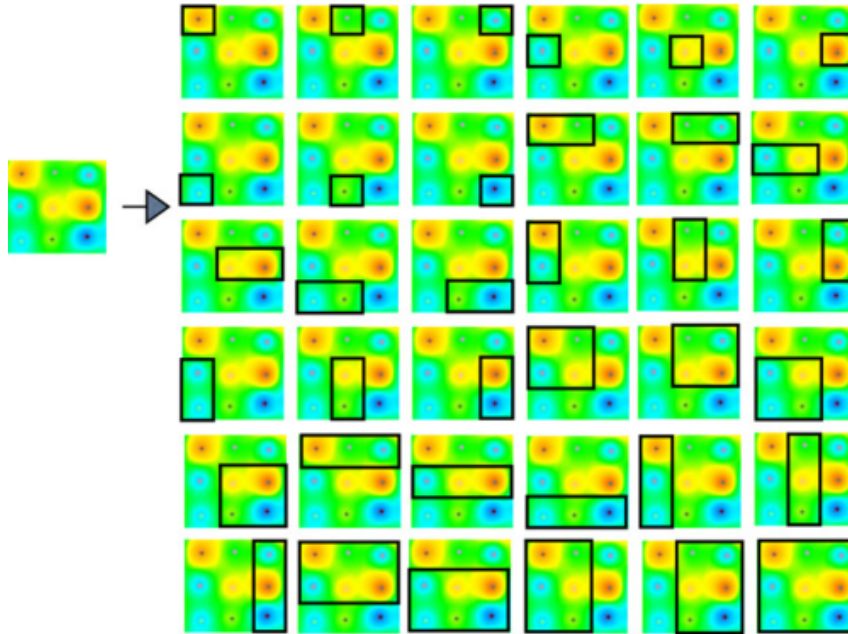


Figura 3.1: Conjunto de zonas potenciales en una muestra de 9 puntos
(Fuente: Cid-García et al. [2013])

3.2. Esquema Jerárquico

El esquema jerárquico presenta un algoritmo que comienza resolviendo el problema de zonificación, este problema tiene como entrada todas las potenciales zonas de manejo, y escoge la partición óptima con menor cantidad de zonas, que logra cumplir con el nivel de homogeneidad requerido, este nivel se obtiene utilizando una linealización de RV (2.1).

Posterior a ello, y utilizando sólo las zonas obtenidas de la partición óptima se resuelve el problema de localización discreto, a través de un modelo basado en el propuesto por [ReVelle and Swain \[1970\]](#).

3.2.1. Modelo de zonificación

Este modelo fue propuesto inicialmente por [Cid-García et al. \[2013\]](#), utilizando como datos de entrada algún índice que pueda ser medido de forma preliminar a la localización de los sensores, y también, apropiado para las propiedades que interesan medir posteriormente.

Este modelo tiene como objetivo dividir el terreno en celdas rectangulares, con la partición que tenga el mínimo de zonas de manejo y además cumpla con el nivel de homogeneidad requerido.

Conjuntos:

- S : Conjunto de puntos muestrales del terreno $s \in S$
- Z : Conjunto de todas las zonas de manejo rectangulares potenciales $z \in Z$.

Parámetros:

- $c_{s,z}$: Coeficiente binario que representa a las zonas, toma el valor 1 cuando el punto muestral s pertenece a la zona z , y 0 de lo contrario.
- σ_z^2 : Varianza del índice muestreado en la zona z .
- n_z : Cantidad de puntos muestrales que pertenecen a la zona z .
- α : Nivel de homogeneidad requerido.
- σ_T^2 : Varianza del índice muestreado de todo el terreno
- N : Total de puntos muestrales.
- LI : Límite inferior de zonas en la partición óptima.
- LS : Límite superior de zonas en la partición óptima.

Variables de decisión:

- q_z : Variable binaria que toma el valor 1 si la zona z pertenece a la partición óptima y 0 de lo contrario.

Formulación del Modelo:

$$\min \sum_{z \in Z} q_z \quad (3.1)$$

$$s.t. \sum_{z \in Z} c_{s,z} q_z = 1 \quad \forall s \in S \quad (3.2)$$

$$\sum_{z \in Z} \left[(n_z - 1) \sigma_z^2 + (1 - \alpha) \sigma_T^2 \right] q_z \leq N(1 - \alpha) \sigma_T^2 \quad (3.3)$$

$$LI \leq \sum_{z \in Z} q_z \leq LS \quad (3.4)$$

$$q_z \in \{0, 1\} \quad z \in Z \quad (3.5)$$

La función objetivo busca minimizar el número de zonas de manejo que forman la partición óptima. La restricción (3.2) obliga a que cada punto muestral s pertenezca sólo a 1 zona en la partición, la restricción (3.3) garantiza el nivel de homogeneidad requerido con el parámetro α , a través, de la linealización del coeficiente RV (2.1), y descrito en términos de los parámetros y variables descritas en el modelo. La restricción (3.4) limita el número de zonas en la partición óptima con una cota inferior y una superior. Y finalmente, la última restricción (3.5) define la naturaleza binaria de las variables.

El tamaño de este problema depende directamente de la cantidad de puntos muestrales que se utilicen en cada iteración, en cuanto a las variables, habrán tantas como zonas potenciales tenga el conjunto $|Z|$. Para las restricciones, se tendrán la cantidad de puntos muestrales más 3. Considerar una muestra de un terreno rectangular con ancho m puntos y largo n puntos muestrales, si se consideran zonas potenciales factibles de tamaño 1×1 hasta $m \times n$ puntos muestrales, el tamaño del modelo está dado por:

Conjunto de zonas potenciales	$ Z = \frac{m \cdot (m+1) \cdot n \cdot (n+1)}{4}$
Cantidad de variables	$\sum_{z \in Z} q_z = Z $
Cantidad de restricciones	$m \cdot n + 3$

3.2.2. Modelo de localización

Este modelo está basado en el modelo de localización de instalaciones desarrollado por [ReVelle and Swain \[1970\]](#), como datos de entrada considera sólo las zonas de manejo en la partición óptima obtenida de resolver el problema de zonificación (3.2.1).

El objetivo es encontrar la ubicación óptima de sensores que logra una muestra representativa, esto es, minimizando la varianza. Es por ello, que se propone la minimización de la suma ponderada, utilizando la varianza de cada zona como los pesos.

Conjuntos:

- Z^* : Conjunto de las zonas de manejo en la partición óptima del modelo (3.1)–(3.5).

Parámetros:

- σ_i^2 : Varianza de la propiedad muestreada inicialmente en cada zona $i \in Z^*$
- $d_{i,z}$: Distancia entre los centroides de las zonas $i \in Z^*$ y $z \in Z^*$
- p : Límite de sensores a localizar.

Variables de decisión:

- $X_{i,z}$: Variable binaria que toma el valor 1 si la zona $i \in Z^*$ es asignada al sensor ubicado en la zona $z \in Z^*$.

Formulación del modelo:

$$\min \sum_{i \in Z^*} \sum_{z \in Z^*} \sigma_i^2 d_{i,z} X_{i,z} \quad (3.6)$$

$$s.t. \sum_{z \in Z^*} X_{i,z} = 1 \quad \forall i \in Z^* \quad (3.7)$$

$$X_{i,z} \leq X_{z,z} \quad \forall i, z \in Z^* \quad (3.8)$$

$$\sum_{z \in Z^*} X_{z,z} \leq p \quad (3.9)$$

$$X_{i,z} \in \{0, 1\} \quad \forall i, z \in Z^* \quad (3.10)$$

El objetivo del modelo presentado (3.6)–(3.10) es minimizar la distancia ponderada entre todas las zonas de la partición óptima y la ubicación de los sensores. Las restricciones (3.7) y (3.8) imponen que cada zona $i \in Z^*$ debe ser asignada a un sensor ubicado en la zona $z \in Z^*$. Notar que cuando $X_{z,z}$ toma el valor 1, un sensor se localiza en la zona z y además, esa zona es asignada a sí misma. La restricción (3.9) limita los sensores a localizar y finalmente (3.10) define la naturaleza binaria de las variables.

El tamaño de este problema depende directamente del número de zonas en la partición óptima obtenidas de la solución del modelo de zonificación. La cantidad de variables está dada por el cuadrado de la norma del conjunto de zonas óptimas. A pesar de que el cuadrado de la cantidad de elementos de un conjunto suena mucho mayor a las variables consideradas en el modelo anterior, cabe destacar que el conjunto de zonas óptimas es un subconjunto mucho más pequeño que el conjunto de zonas potenciales, por lo que, este modelo es de todas formas, más pequeño que el modelo de zonificación. Sea s la cantidad óptima de zonas, el tamaño del modelo de localización está dado por:

$$\begin{aligned} \text{Cantidad de variables} & \quad \sum_{i \in Z^*, z \in Z^*} X_{i,z} = s \cdot s \\ \text{Cantidad de restricciones} & \quad 2 \cdot s \cdot s + 1 + \end{aligned}$$

El esquema jerárquico es implementado en Julia JuMP a través de las instrucciones presentadas en el anexo 6.2, donde ambos modelos son resueltos secuencialmente, guardando la solución óptima del modelo de zonificación para utilizar como conjunto en el modelo de localización. En este algoritmo también se itera para distintos valores de p , es decir, para distintas cantidades de dispositivos disponibles a localizar.

3.3. Modelo Integrado

Como segunda opción, se presenta un modelo de optimización lineal en variables binarias que permite, simultáneamente, obtener una partición del terreno y localización óptima de los dispositivos en un Modelo Integrado.

Este modelo considera los conjuntos y parámetros descritos anteriormente en el esquema jerárquico, tanto del modelo de zonificación como del modelo de localización. A diferencia de la instancia anterior, esta propuesta toma todas las zonas de manejo potenciales para decidir tanto la partición como la localización óptima de los dispositivos al mismo tiempo, por lo que, no se utiliza el conjunto Z^* con las zonas de la partición óptima para la implementación de un modelo de localización.

El objetivo es localizar los sensores. El modelo de zonificación se utiliza como un medio para lograrlo, puesto que los centroides de las zonas representan a los puntos factibles del modelo de localización discreto. Por esta razón, el objetivo de este Modelo Integrado, al igual que en el modelo de localización (3.2.2) es minimizar la distancia ponderada.

VARIABLES DE DECISIÓN:

- q_z : Variable binaria que toma el valor 1 si la zona $z \in Z$ está en la partición y 0 de lo contrario.
- $X_{i,z}$: Variable binaria que toma el valor 1 si la zona $i \in Z$ es asignada a un sensor localizado en la zona $z \in Z$.

Formulación del modelo:

$$\text{mín} \sum_{i \in Z} \sum_{z \in Z} \sigma_i^2 d_{i,z} X_{i,z} \quad (3.11)$$

$$\text{s.t.} \sum_{z \in Z} c_{s,z} q_z = 1 \quad \forall s \in S \quad (3.12)$$

$$\sum_{z \in Z} [(n_z - 1) \sigma_z^2 + (1 - \alpha) \sigma_T^2] q_z \leq N(1 - \alpha) \sigma_T^2 \quad (3.13)$$

$$LI \leq \sum_{z \in Z} q_z \leq LS \quad (3.14)$$

$$X_{z,z} \leq q_z \quad \forall z \in Z \quad (3.15)$$

$$\sum_{z \in Z} X_{i,z} = q_i \quad \forall i \in Z \quad (3.16)$$

$$X_{i,z} \leq X_{z,z} \quad \forall i, z \in Z \quad (3.17)$$

$$\sum_{z \in Z} X_{z,z} \leq p \quad (3.18)$$

$$q_z \in \{0, 1\} \quad \forall z \in Z \quad (3.19)$$

$$X_{z,z} \in \{0, 1\} \quad \forall z \in Z \quad (3.20)$$

La función objetivo (3.11), siguiendo la idea presentada en el modelo de localización (3.2.2) minimiza la distancia ponderada entre las zonas de la partición óptima y la ubicación de los sensores. En este caso, la suma se realiza en el conjunto de todas las zonas, por lo que, el modelo podrá encontrar la combinación óptima de ubicaciones de sensores considerando todas las posibilidades, en vez, de limitarse a una partición dada.

Las restricciones (3.12)–(3.14) corresponden a las restricciones presentadas en el modelo de zonificación (3.2.1). La restricción (3.15) se encarga de que un sensor pueda ser localizado ($X_{z,z} = 1$) en la zona $z \in Z$ sólo si esa zona pertenece a la partición ($q_z=1$), esta restricción realiza una tarea similar a lo presentado en el esquema jerárquico, que es utilizar las zonas de la partición óptima como los puntos factibles para el modelo de localización.

Las restricciones (3.16) – (3.18) corresponden a las presentadas en el modelo de localización (3.2.2), pero con la variación en la restricción (3.16), en este caso la restricción de asignación se

limita a las zonas que pertenecen a la partición óptima, entonces serán asignadas (lado derecho igual a 1) sólo si la zona $z \in Z$ pertenece a la partición óptima, de lo contrario el lado derecho será igual a 0. Finalmente las restricciones (3.19) – (3.20) corresponden a la naturaleza binaria de las variables tanto de zonificación como de localización.

El tamaño de este problema depende directamente del tamaño de la muestra utilizada. La cantidad de variables será mayor al cuadrado de la cantidad de zonas potenciales, siendo este número considerablemente mayor a la cantidad en ambos modelos del esquema jerárquico. Considerar una muestra de un terreno rectangular con ancho m puntos y largo n puntos muestrales, si se consideran zonas potenciales factibles de tamaño 1×1 hasta $m \times n$ puntos muestrales, el tamaño del modelo está dado por:

Conjunto de zonas potenciales	$ Z = \frac{m \cdot (m+1) \cdot n \cdot (n+1)}{4}$
Cantidad de variables	$\sum_{z \in Z} q_z = Z ^2 + Z $
Cantidad de restricciones	$m \cdot n + 3 + 2 \cdot Z + Z ^2$

El modelo integrado es implementado en Julia JuMP a través de las instrucciones presentadas en el anexo 6.3, donde en primer lugar se fija un LS apropiado y se itera para distintos valores de p .

Los modelos presentados en este capítulo tienen muchas similitudes, puesto que se intenta resolver el mismo problema de localización de sensores. Sin embargo, una de las principales diferencias a encontrar es la carga computacional, dado que el modelo integrado tiene una cantidad de variables y restricciones considerablemente mayor a la que se puede encontrar en el esquema jerárquico. Esta carga se explica dado que el modelo integrado permite evaluar todas las posibles combinaciones de partición, con el objetivo de minimizar la distancia ponderada de la localización de dispositivos, a diferencia del esquema jerárquico, que minimiza la distancia ponderada con una partición fija.

4 — Caso de estudio y resultados

4.1. Implementación Computacional

La presente sección muestra la implementación de la metodología propuesta en el capítulo 3 en un caso de estudio. La implementación computacional fue obtenida utilizando Gurobi en Julia JuMP.

Julia JuMP es un lenguaje de programación de alto rendimiento, tiene una sintaxis de alto nivel y es de fácil uso para implementar tanto el procesamiento de datos de entrada como los modelos presentados ([Lauwens and Downey \[2019\]](#)). Gurobi Optimizer es el solucionador más rápido y potente para el apoyo a la toma de decisiones hoy en día ([Gurobi Optimization, LLC \[2023\]](#)).

Las instancias resueltas mostradas en esta sección fueron implementadas en un computador con las siguientes características Intel Core i5-10210U CPU 1.60 GHz processor and 16 GB RAM, y las soluciones fueron obtenidas con un GAP de optimalidad menor a 10^{-4} .

Los modelos descritos en el capítulo 3 fueron aplicados en un caso de estudio para una data obtenida del sur de Chile, el índice utilizado para los datos de entrada fue el NDVI. Este índice toma valores entre -1.0 y +1.0, a continuación la imagen 4.1 explica la interpretación del rango de valores que puede tomar:

El conjunto de datos disponible corresponde a 1000 puntos muestrales del índice NDVI en un rango entre 0.7594 – 0.7725, por lo que, la vegetación muestreada se encuentra en el rango de plantas muy sanas. Debido a que el rango es muy pequeño y que los datos presentan una gran cantidad de decimales, para evitar la pérdida de información por cifras significativas se decide amplificar

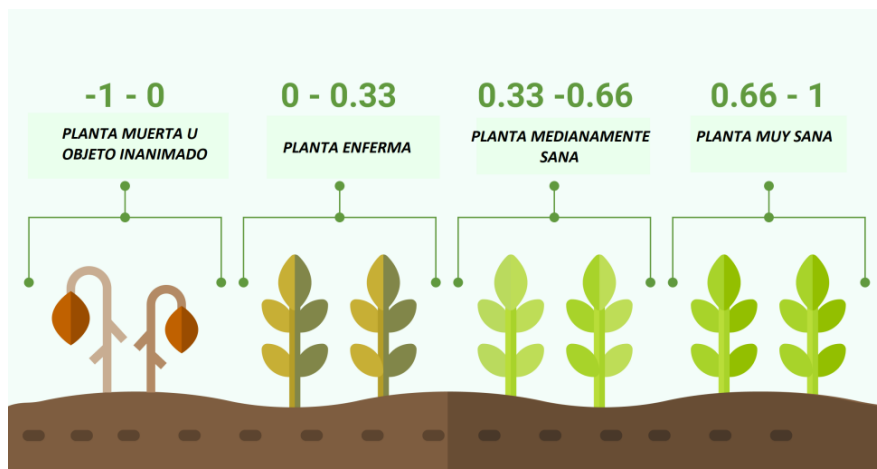


Figura 4.1: Valores del NDVI y su interpretación según el estado de la vegetación. (Fuente: Toribio [2022])

todos los datos por 1000, implementando todo lo que sigue a continuación con el índice NDVI multiplicado por este valor. En la tabla 4.1 se muestra una parte de los datos mostrado como una grilla de 6x10 utilizados para implementar los modelos.

Tabla 4.1: Sección de 6x10 puntos muestrales de la data de NDVI amplificada

766.7	767.2	767.8	768.8	769	768.3	768.5	769.2	769	769.4
766.7	764.9	767.7	767.9	768.6	769.1	769.2	769.6	769.9	769.7
767.5	767.9	768.4	767.7	767.6	768.2	768.3	769.6	769.5	769.7
767.9	771.9	769	768	769	768.1	767.2	767.7	768.8	769.5
767.3	768.7	769.2	769.1	769.7	770.2	769.6	767.3	768.4	769.2
766	768.9	769.4	769.7	770.6	770.1	769.6	768.1	768.1	769.2

A continuación, la figura 4.2 muestra el comportamiento del índice en un mapa de calor del terreno muestreado. Las zonas rojas muestran los puntos que alcanzan mayores valores, mientras que las zonas azules corresponden a los puntos que alcanzan los menores valores del índice. Aplicando el algoritmo (6.1) se obtienen todas las combinaciones de zonas rectangulares a partir del conjunto de puntos.

En esta sección, varias instancias serán resueltas utilizando el mismo set de datos, o subconjuntos más pequeños de este, para ambos modelos presentados en el capítulo 3. En particular para el esquema jerárquico, no se impone restricción sobre los límites inferiores y superiores (LI y LS) de la cantidad de zonas que debe tener la restricción óptima del modelo de zonificación (3.2.1), esto se debe a que este modelo está diseñado para buscar la mejor combinación que cumple con el nivel de homogeneidad y a la vez, con la cantidad mínima de zonas que puede formar esa partición.

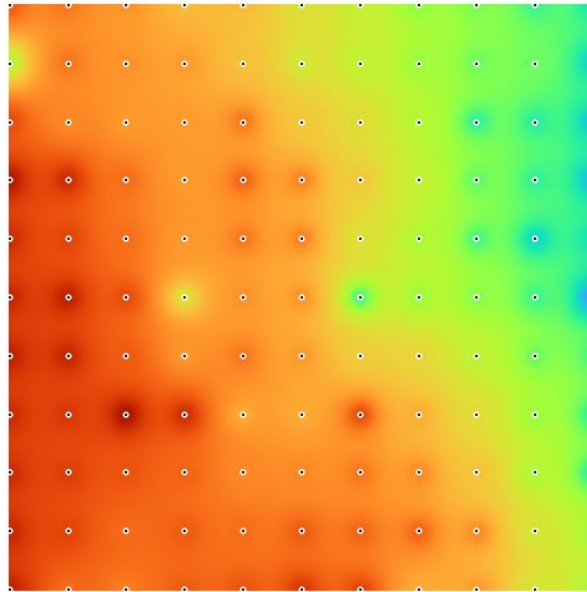


Figura 4.2: Mapa de calor basado en el índice NDVI.
(Fuente: Elaboración propia, uso de software QGIS)

Por otro lado, en el esquema integrado, la función objetivo sólo busca la mejor combinación para minimizar la suma ponderada de las distancias, y a pesar de que, se imponen restricciones sobre la homogeneidad de la partición, nada obliga a que este modelo encuentre la cantidad mínima de zonas que cumpla con lo anterior, así el modelo irrestricto en LI y LS buscará la partición con mayor cantidad de zonas de manejo, y menos puntos en cada una de ellas logrando minimizar el peso definido como la varianza de cada zona. Para remediar esta situación se propone fijar una cota superior (LS) adecuada, que permitiendo alcanzar el nivel de homogeneidad requerido, encuentre el mínimo de zonas que satisfacen las restricciones.

En esta oportunidad, se propone utilizar la solución óptima del modelo de zonificación (3.2.1) dado que corresponde a una solución factible para alcanzar el nivel de homogeneidad, con ello fijar el límite superior (LS) de zonas en el modelo integrado y comparar si este último, en las mismas condiciones que el esquema jerárquico, logra encontrar una mejor localización para los sensores.

La primera instancia resuelta considera una porción del terreno con 42 puntos muestrales. El nivel de homogeneidad requerido es de $\alpha = 0,9$. Considerando ello, en primer lugar fue resuelto el modelo de zonificación (3.2.1), arrojando una partición óptima con 12 zonas de manejo rectangulares, por lo que, LS fue fijado en ese valor. A continuación la tabla 4.2 muestra la comparación en las localizaciones de dispositivos para diferentes valores de sensores disponibles (p).

Tabla 4.2: Resultados para ambos modelos con diferentes valores de p

Instance	$ S $	$ Z $	p	z_{HA}	z_{IA}
1	42	588	12	0,0	0,0
2	42	588	11	0,0	0,0
3	42	588	10	0,0	0,0
4	42	588	9	0,0	0,0
5	42	588	8	0,0	0,0
6	42	588	7	512,8	0,0
7	42	588	6	1.096,6	33,3
8	42	588	5	1.771,7	119,4
9	42	588	4	3.463,1	407,8
10	42	588	3	7.426,9	1.075,9
11	42	588	2	11.256,6	3.231,4
12	42	588	1	17.585,2	9.751,6

La tabla 4.2 muestra en la columna Z_{HA} los resultados de la distancia ponderada resolviendo el modelo de localización del esquema jerárquico, mientras que la columna Z_{IA} muestra la distancia ponderada de localizar los dispositivos con el esquema integrado. En todas las instancias la partición de ambos modelos tiene 12 zonas de manejo rectangulares en la partición óptima.

A partir de la tabla 4.2 se observa que para todas las instancias el modelo integrado presenta una solución con un valor óptimo menor que el encontrado por el esquema jerárquico, esto se explica porque el modelo integrado siempre puede evaluar la solución en todas las zonas, optimizando la distancia ponderada como objetivo, y a pesar de tener la misma cantidad de zonas en el óptimo, esta decisión es distinta a la tomada en el esquema jerárquico, dónde primero se evalúa el menor número de zonas que puede tener una partición bajo un nivel de homogeneidad requerido y posteriormente, con la partición fija, escoge la posición de los dispositivos encontrando una solución óptima peor que la del modelo integrado.

Cuando se localizan más de 8 dispositivos, en ambos modelos se observa que la distancia ponderada se va a 0, siendo este el menor valor que puede tomar. Esto se explica porque la decisión de localización se toma instalando los dispositivos cercano a las zonas de mayor varianza, así se disminuye uno de los términos de la suma considerada, entonces cuando se tiene cierta cantidad de dispositivos la solución no puede mejorar puesto que las zonas de mayor varianza están siendo cuantificadas por un sensor. Estas instancias, de distancia ponderada 0 serán llamadas óptimas, por lo que, en este caso para tener una muestra óptima del terreno en el esquema jerárquico se requieren al menos 8 dispositivos, siendo 8 el número eficiente, ya que con 9 seguirá siendo 0, mientras que para el modelo integrado se requieren al menos 7. Considerando ello, el esquema integrado

sobredimensiona la cantidad de dispositivos requerida, ya que para obtener la misma solución, se requiere de más sensores.

A continuación, se muestran instancias más grandes, el nivel de homogeneidad sigue siendo $\alpha = 0,9$. En primer lugar, fue resuelto el modelo de zonificación, con este valor se fija el número de zonas en la partición óptima. Los gráficos de las Figuras 4.3 ,4.4 y 4.5 muestran el valor de la distancia ponderada tanto para el esquema jerárquico como para el modelo integrado.

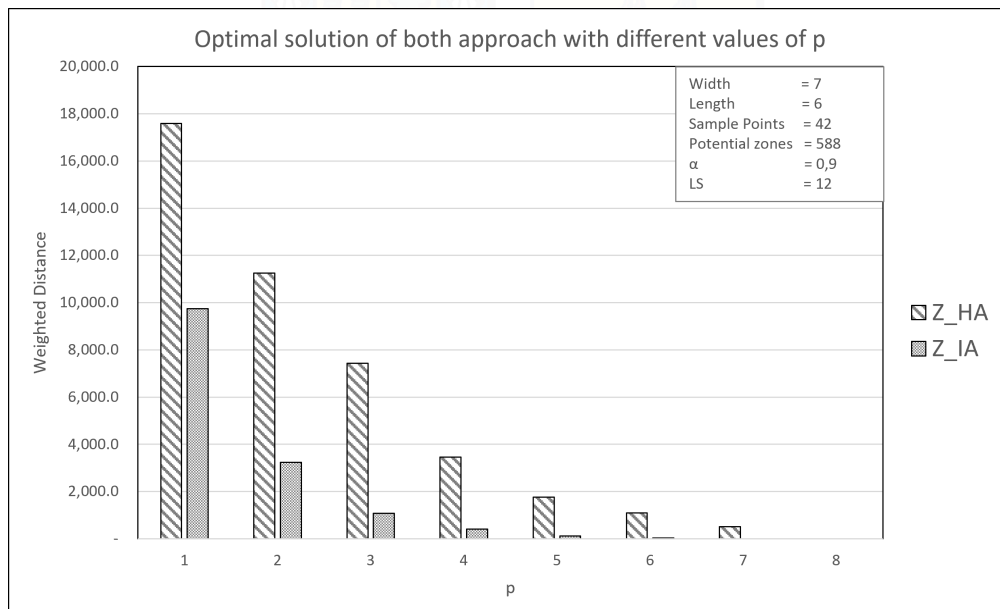


Figura 4.3: Resultados de ambos modelos para una instancia de 42 puntos muestrales. (Fuente: Elaboración propia)

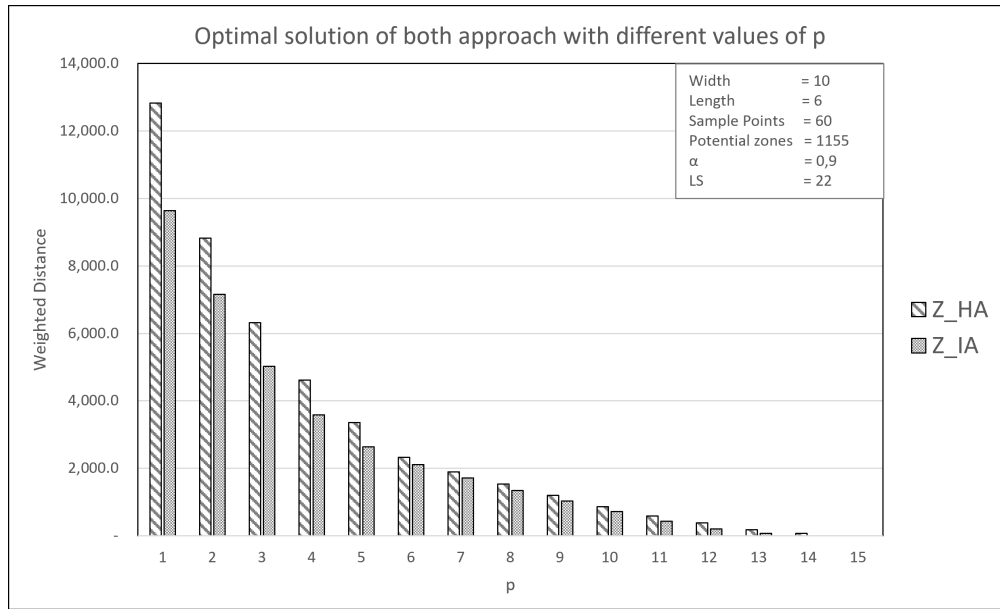


Figura 4.4: Resultados de ambos modelos para una instancia de 60 puntos muestrales.
(Fuente: Elaboración propia)

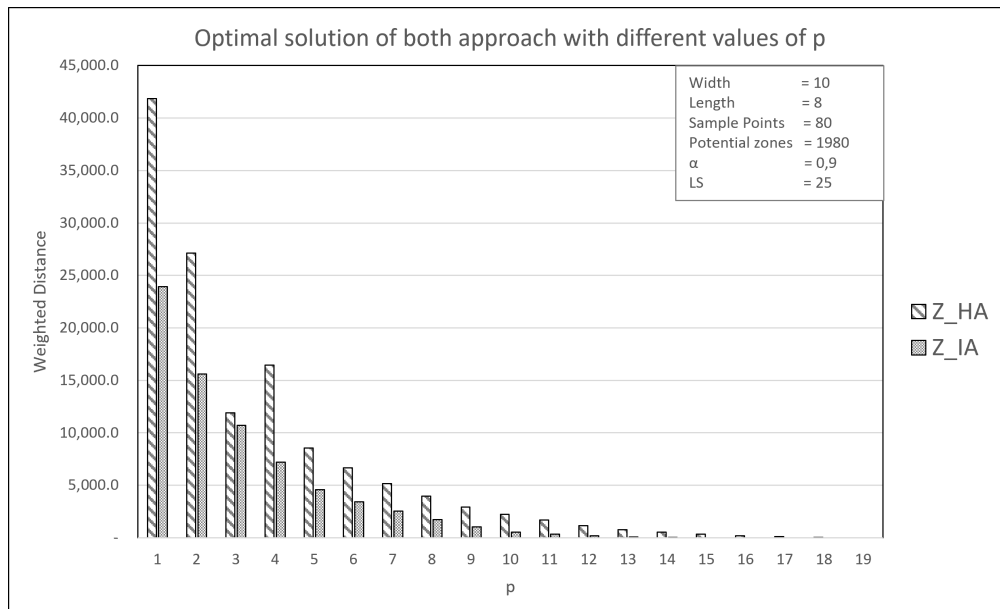


Figura 4.5: Resultados de ambos modelos para una instancia de 80 puntos muestrales.
(Fuente: Elaboración propia)

Los gráficos 4.3 ,4.4 y 4.5, muestran que aunque la instancia crezca nuevamente ocurre lo observado en la tabla 4.2 la solución encontrada por el modelo integrado siempre mejora respecto de la que puede encontrar el esquema jerárquico.

La superioridad mencionada también se evidencia en la tabla 6.2, esta tabla presenta los datos

con los que fue construido el gráfico 4.5 y la diferencia porcentual de la solución óptima entregada por ambos modelos en cada caso. La instancia de $p=3$ dispositivos a localizar, es la instancia donde las respuestas se encuentran más cercanas entre si, y el esquema jerárquico entrega una solución un 10 % superior a la entregada por el modelo integrado, a excepción de la instancia $p=19$ dispositivos donde ambos modelos llegan a 0 en su función objetivo, sin embargo, el modelo integrado alcanza este cero en varias iteraciones antes que el esquema jerárquico.

La ventaja de usar zonas de manejo para representar los puntos factibles, y para determinar los puntos a estimar, por sobre otro tipo de partición del terreno viene dada porque es posible fijar un nivel α de homogeneidad de las zonas, permitiendo medir en un sólo punto de la zona para obtener una estimación de la propiedad en toda la extensión, de esta forma una solución factible para el modelo de localización, es utilizada la cantidad de dispositivos requeridos para una muestra de todo el espacio sólo utilizando tantos sensores como cantidad de zonas en la partición óptima. Por lo que, es esperable que en la decisión óptima se escojan a lo más tantos dispositivos como zonas tenga la partición óptima, a pesar de que puedan haber más dispositivos disponibles.

Los gráficos 4.3 ,4.4 y 4.5 se encuentra una frontera de pareto, o puntos eficientes que corresponde al nivel de distancia ponderada que se puede alcanzar con la cantidad de sensores disponibles, entonces dado cierto número de sensores si se decide localizar en los puntos óptimos que encuentran los modelos (según el que se decida utilizar) se tendrá cierto nivel de distancia ponderada, y por otro lado, si se restringe el nivel de objetivo que se desea alcanzar, se puede encontrar la cantidad de dispositivos que ofrece esa solución.

Para cada instancia resuelta, se encuentra el nivel en que la función objetivo alcanza su mejor solución, como ya se mencionó, esta es 0. Con ello, se puede conocer la cantidad óptima de sensores a localizar para cada instancia en específico.

4.2. Análisis Gráfico

A continuación se presenta la solución gráfica del modelo de zonificación para distintos tamaños de instancias y un nivel de homogeneidad $\alpha = 0,9$.

De las figuras anteriores 4.6 se observa que cuanto menos puntos muestrales contiene el terreno, la partición óptima considera zonas más pequeñas, es decir, con menos puntos muestrales en cada una, y a medida que el terreno considerado crece, las zonas consideradas son más grandes, y de hecho, escoge menos zonas que en la instancia más pequeña. Incluso, se observa a través del mapa

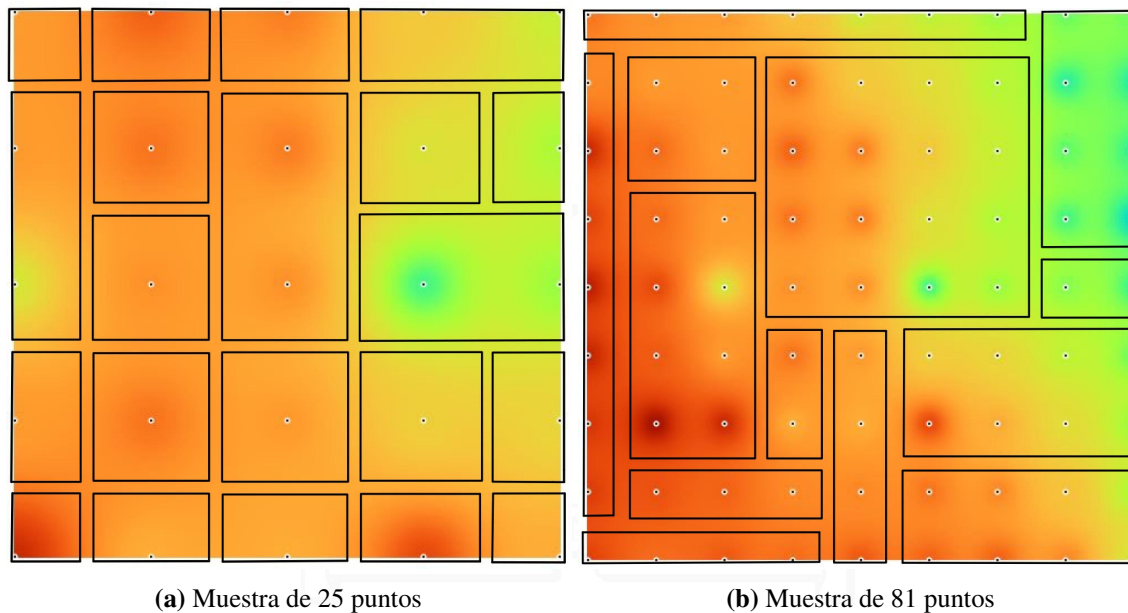


Figura 4.6: Solución gráfica del modelo de zonificación para dos tamaños de muestras

de calor que el parámetro observado en la figura 4.6a es mucho más homogéneo que lo presentado en la figura 4.6b.

Esto se debe a que, el nivel de homogeneidad utilizando la Varianza Relativa (2.1) considera todo el terreno, entonces con menos puntos, el nivel de α aumenta considerablemente con pequeñas variaciones de la varianza de cada una de las zonas de manejo, a diferencia de muestras más grandes, donde cada una de las tiene un aporte marginal mucho más bajo al coeficiente, y al intentar minimizar la cantidad de zonas en la partición óptima, se permite tomar zonas más grandes.

En las figuras 4.7 y 4.8 se presentan dos iteraciones del modelo, con 49 y 81 puntos muestrales respectivamente, donde la muestra más grande contiene en el centro los 49 puntos de la muestra pequeña y se agregan los puntos colindantes a los bordes. La metodología explicada se realiza con el objetivo de observar qué ocurre con el resultado obtenido cuando el modelo cuenta con más información en la frontera de la muestra.

La partición óptima del terreno varía en gran cantidad en cada iteración, obteniendo zonas de manejo muy diferentes en ambas instancias, y también distinta respecto de usar el esquema jerárquico y el modelo integrado. Si se considera el uso de sólo un modelo, para explicar la diferencia entre la partición obtenida en la iteración 4.7a y 4.7b y en las iteraciones 4.8a y 4.8b se debe recordar lo mencionado en el capítulo 2, tanto el modelo de zonificación en el esquema jerárquico, como el

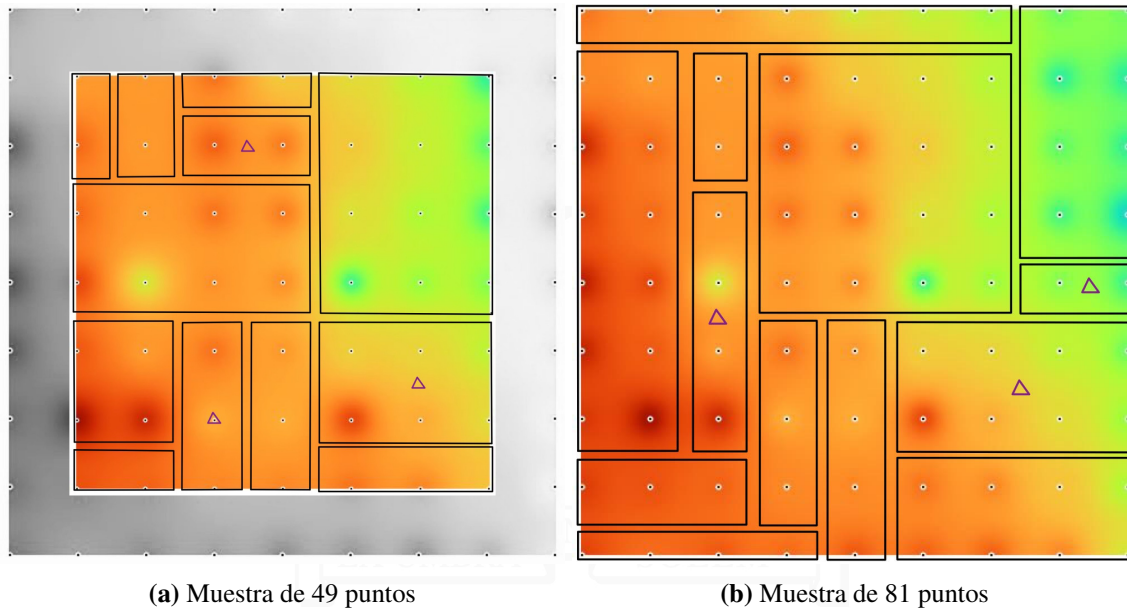


Figura 4.7: Localización de 3 sensores utilizando el esquema jerárquico para distintos tamaños de muestras

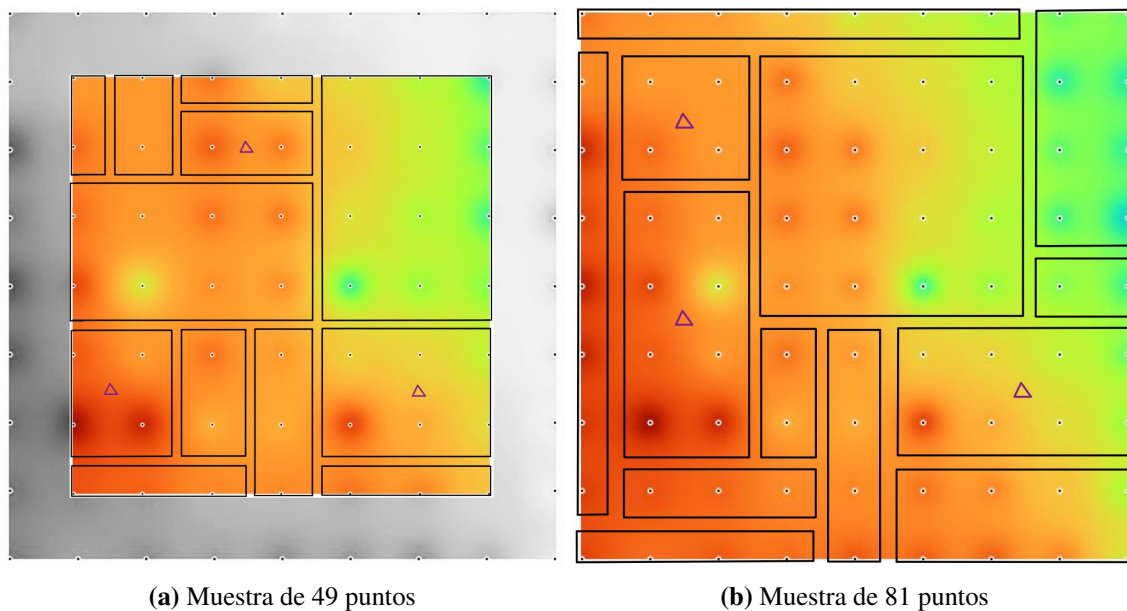


Figura 4.8: Localización de 3 sensores utilizando el modelo integrado para distintos tamaños de muestras

modelo integrado utilizan la varianza relativa 2.1 para garantizar la homogeneidad de la partición, este parámetro considera de forma integrada la varianza de todas las zonas, a través de la sumatoria, y también la cantidad de puntos muestrales considerados, por lo que, a pesar que el nivel de homogeneidad se mantenga constante, al modificar la cantidad de puntos muestrales el nivel α se alcanza con una distribución distinta de zonas de manejo.

Por otro lado, si se considera la diferencia de particiones encontradas a través de las dos propuestas para resolver el modelo, es decir, la partición obtenida en la iteración 4.7a y 4.8a y, por otro lado, con las iteraciones 4.7b y 4.8b se encuentra que con ambos modelos tienen la misma cantidad de zonas, 12 para la iteración de 49 puntos y 13 para la iteración de 81 puntos, sin embargo, la distribución de zonas de manejo es distinta. Esto se debe a que en primer lugar se resuelve el modelo de zonificación, donde se encuentra la partición óptima con la homogeneidad requerida $\alpha = 0,9$ que contiene la menor cantidad de zonas, sabiendo ello se utiliza este mínimo como límite superior para resolver el modelo integrado, el cual, tiene como solución factible la misma partición encontrada por el modelo de zonificación, sin embargo, el objetivo de este modelo es minimizar la distancia ponderada de la localización de dispositivos, por lo que, la partición, aunque con la misma cantidad de zonas, podría tener distinta distribución.

A pesar de la diferencia encontrada en las particiones, se observa que la localización de los dispositivos cuando se agregan puntos colindantes sigue una tendencia. Para estas iteraciones en específico, es ubicar dos dispositivos en la parte inferior del terreno y uno en la parte superior. En el caso del modelo integrado 4.8 no se observa gran diferencia en las ubicaciones escogidas cuando se agregan los puntos colindantes a los bordes, es decir, entre la figura 4.8a y 4.8b. Con ello, es posible observar un modelo más consistente, el cual no se ve perturbado por la información más extrema de la muestra. A diferencia de lo que ocurre en el esquema jerárquico 4.7, en este modelo se percibe la tendencia mencionada, sin embargo, al considerar los puntos colindantes a la frontera, la ubicación óptima encontrada para los dispositivos se encuentra más agrupada hacia el centro y la derecha, por lo que, este modelo sí se perturbó con la información de los puntos extremos.

4.3. Relajación en el límite superior LS

Según lo mencionado anteriormente, las soluciones óptimas fueron obtenidas al resolver ambos modelos fijando LS como la solución óptima del modelo de zonificación, sin embargo, esto representa un esfuerzo superior al intentar resolver el modelo ya que implica tener que resolver un modelo adicional de manera preliminar para obtener los parámetros de entrada del modelo propuesto.

De manera óptima para efectos de comparación entre la propuesta jerárquica y la integrada, se debe utilizar la cota de LS mencionada. A continuación, se observa como cambia la solución y en cuánto se subestima el óptimo cuando LS es relajado. El gráfico 4.9 muestra la solución para varias iteraciones utilizando 60 puntos muestrales. El modelo de zonificación encuentra una partición óptima con 22 zonas de manejo, las barras verticales muestran la distancia ponderada del modelo integrado utilizando $\alpha = 0,9$ y $p=3$, para distintos valores de la cota superior en un rango entre 22 y 32. También se grafica línea de tendencia, la cual es decreciente y cuanto más aumenta el límite superior, la tasa de decrecimiento disminuye.

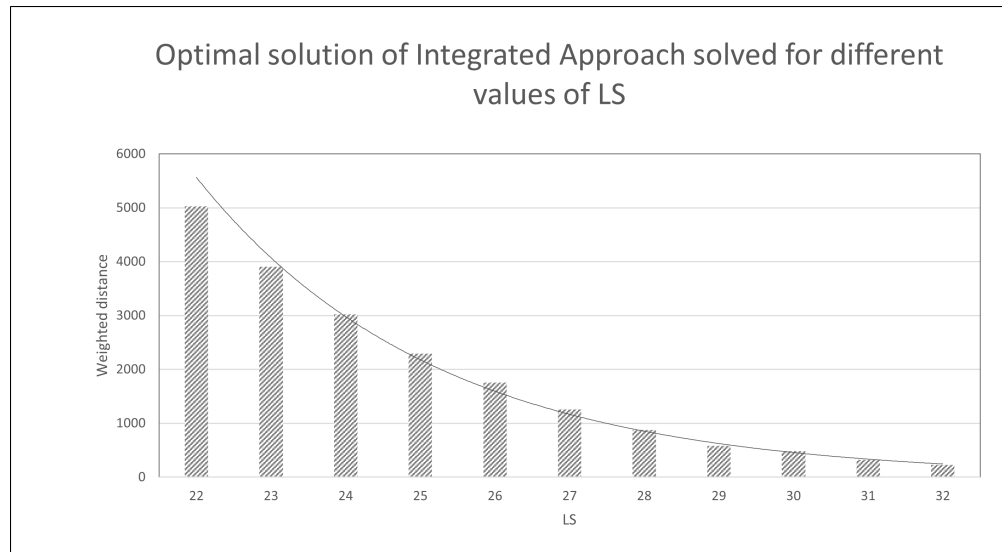


Figura 4.9: Distancia ponderada para el modelo Integrado con distintos valores del límite superior.
(Fuente: Elaboración propia)

Además, cabe mencionar que cuando el límite se fija a un valor menor que el óptimo del problema de zonificación, el modelo integrado es infactible. Esto se debe a que para cierto nivel de homogeneidad, la zonificación encuentra la partición con el mínimo de zonas de manejo. Por otro lado, si el número es superior a este óptimo, la partición tendrá zonas más pequeñas y la distancia ponderada de la localización de sensores será infravalorada, sin embargo, es tolerable ya que esta distancia no tiene una interpretación por sí misma y lo realmente interesante es las ubicaciones que

escoge para los dispositivos.

4.4. Costo Computacional

En secciones anteriores se evidencia la superioridad en el modelo integrado para encontrar la ubicación óptima de sensores, sin embargo, un aspecto importante para la comparación es el costo computacional de la implementación de los modelos.

La tabla 4.3 muestra la cantidad de variables y restricciones que tiene cada modelo para distintos tamaños de muestra. No se evidencia el tiempo de solución de las instancias puesto que, dado los tamaños de los modelos no fue posible resolver instancias mayores a 121 puntos muestrales para el caso del esquema integrado. De la tabla 4.3 se infiere la gran diferencia de tamaño que tienen ambos modelos, y si bien, la cantidad de variables y restricciones crece para ambos modelos, en el caso del modelo integrado, la tasa de crecimiento es muy alta, e incluso en instancias pequeñas la cantidad de variables y restricciones que tiene es mucho más grande que las que considera el esquema jerárquico sumando las de ambos modelos (zonificación y localización).

Tabla 4.3: Costo computacional de ambos modelos para distintos tamaños de muestras

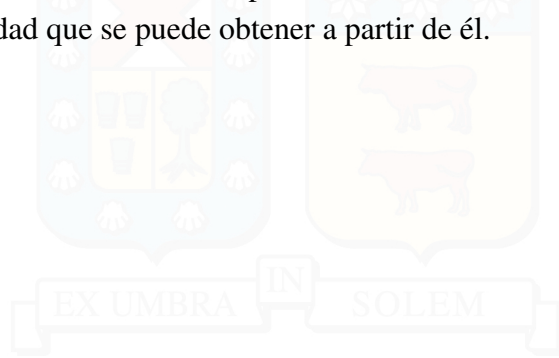
S	Z	Esquema Jerárquico				Modelo Integrado	
		Zonificacion		Localizacion		Variables	Restricciones
		Variables	Restricciones	Variables	Restricciones		
42	588	588	45	144	157	346.332	346.966
60	1.155	1.155	63	484	507	1.335.180	1.336.399
80	1.980	1.980	83	625	651	3.922.380	3.924.444
100	3.025	3.025	103	841	871	9.153.650	9.156.779
400	44.100	44.100	403	1.764	1.807	1.944.854.100	1.944.898.604
600	97.650	97.650	603	4.900	4.971	9.535.620.150	9.535.718.404

Lo anterior se explica porque el modelo integrado considera para cada una de las zonas la localización de dispositivos, a diferencia del esquema jerárquico que en primer lugar resuelve sólo el modelo de zonificación, y luego tomando sólo las zonas óptimas, resuelve la localización de dispositivos.

La cantidad de zonas de manejo potenciales a partir de N puntos muestrales, obtenidos de un terreno rectangular de ancho n y largo m serán $|Z| = n \cdot (n + 1) \cdot m \cdot (m + 1)$, a partir de ello, la cantidad de variables que tendrá el esquema jerárquico, en el modelo de zonificación q_z será igual a la cantidad de zonas potenciales, y para el modelo de localización $X_{i,z}$ será el cuadrado de las zonas de la partición óptima. Por otro lado, en el modelo integrado la cantidad de variables está dada por

las variables $q_z = |Z|$ más $X_{i,z} = |Z|^2$ explicando su alta tasa de crecimiento.

A partir de lo anterior, el esquema jerárquico, a pesar de sobreestimar la distancia ponderada, y encontrar una solución algo peor que el modelo integrado, continúa siendo una buena opción para resolver grandes instancias, donde el costo computacional del modelo integrado es muy alto respecto de la ganancia en optimalidad que se puede obtener a partir de él.



5 — Conclusiones

La presente investigación propone dos alternativas para resolver el problema de localización de sensores en la agricultura de precisión. Ambas están basadas en utilizar la partición óptima de un modelo de zonificación como los puntos factibles para un modelo de localización discreta, dado que, al fijar un nivel de homogeneidad, el valor que se mida en cualquier punto puede ser utilizado para estimar la propiedad en todos los puntos de la zona de manejo.

En primer lugar se presenta el esquema jerárquico, este modelo comienza con aplicar el modelo de zonificación y con esa partición óptima se resuelve el modelo de localización discreta tipo p-mediana. Este modelo tiene la ventaja de un bajo costo computacional, lo que permite resolver grandes instancias en un corto periodo de tiempo, pero la solución obtenida se ve perjudicada al fijar la partición óptima.

La solución mejora al utilizar el modelo integrado, en este modelo se evalúa la localización de dispositivos utilizando distintas combinaciones de partición del terreno, dado que se resuelven de forma simultánea, lo que significa un mayor costo computacional.

El modelo integrado requiere de una cota superior para limitar la cantidad de zonas en la partición óptima, una posibilidad es utilizar la solución óptima del modelo de localización, siendo esta la partición con menor cantidad de zonas que se puede obtener para cierto nivel de homogeneidad, y a medida que esta cota se relaja la distancia ponderada de la localización de dispositivos es subestimada, ya que se tendrán particiones con mayor cantidad de zonas en ella y menor varianza por cada zona.

Cuando se dispone de más dispositivos a localizar la solución mejora y se obtienen muestras

más confiables, sin embargo, el número de sensores óptimo no puede aumentar ilimitadamente, dado que existe un punto en que la función objetivo no puede mejorar, esta cantidad es el número óptimo de sensores a localizar en una determinada muestra.

La determinación de las zonas de manejo se realiza a través de la varianza relativa, este parámetro al ser agregado presenta diferencias en la determinación de la partición cuando se consideran muestras de distintos tamaños. Al considerar pocos puntos muestrales, como resultado la partición tendrá una alta cantidad de zonas de manejo con pocos puntos en cada una de ellas, incluyendo zonas con sólo 1 punto para alcanzar un nivel alto de homogeneidad.

Cuando se compara soluciones obtenidas en muestras de diferentes tamaños el modelo integrado presenta una solución consistente, localizando los sensores con la misma tendencia en instancias grandes y pequeñas, mientras que el esquema jerárquico se ve perturbado por los puntos muestrales agregados en cada iteración. Esto significa que el esquema jerárquico da mucha importancia a los puntos de borde o frontera, representando una desventaja del modelo dado que estos puntos representan una porción pequeña respecto de todo el terreno que interesa medir.

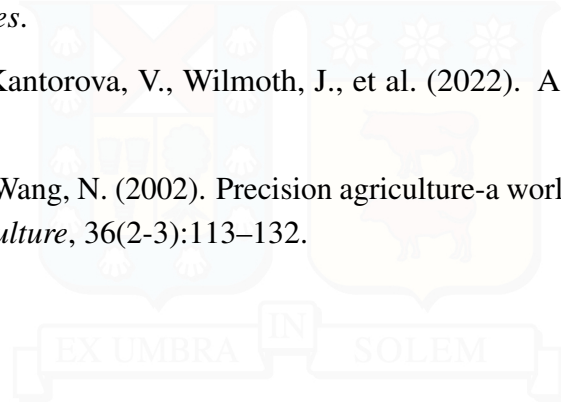
Trabajos futuros se pueden enfocar en dos áreas, en primer lugar el desarrollo de un algoritmo de solución eficiente para grandes instancias, y en segundo lugar la extensión estocástica de este modelo, que en esta ocasión se consideró determinista.

Bibliografía

- Albornoz, V. M. (2023). Management zones by optimization. *Encyclopedia of Digital Agricultural Technologies*, pages 785–791.
- Albornoz, V. M., Araneda, L. C., and Ortega, R. (2022). Planning and scheduling of selective harvest with management zones delineation. *Annals of Operations Research*, 316(2):873–890.
- Albornoz, V. M., Ñanco, L. J., and Sáez, J. L. (2019). Delineating robust rectangular management zones based on column generation algorithm. *Computers and Electronics in Agriculture*, 161:194–201.
- Ashraf, M., Loftis, J. C., and Hubbard, K. (1997). Application of geostatistics to evaluate partial weather station networks. *Agricultural and forest meteorology*, 84(3-4):255–271.
- Brimberg, J. (1995). The fermat-weber location problem revisited. *Mathematical programming*, 71:71–76.
- Church, R., Velle, C. R., et al. (1974). The maximal covering location problem. *Papers in Regional Science*, 32(1):101–118.
- Cid-García, N. M., Albornoz, V., Rios-Solis, Y. A., and Ortega, R. (2013). Rectangular shape management zone delineation using integer linear programming. *Computers and Electronics in Agriculture*, 93:1–9.
- Cid-García, N. M. and Ibarra-Rojas, O. J. (2019). An integrated approach for the rectangular delineation of management zones and the crop planning problems. *Computers and Electronics in Agriculture*, 164:104925.
- Domínguez, J. I., Vergara, M. M., Aguirre, R., Barrera, D., Montero, J., Cáceres, L., Eguillor, P., Espinoza, A., García, A., Reyes, A., Guillermo Pino, M. J. P., Tapia, B., Acuña, D., Laval, E., Yañez, L., Muñoz, M., Cartes, G., Contreras, P., Valdés, A., and Galán, M. (2019). Panorama de la agricultura chilena. *ODEPA, Ministerio de Agricultura*.

- Drezner, Z., Klamroth, K., Schöbel, A., and Wesolowsky, G. O. (2002). The weber problem. *Facility location: Applications and theory*, pages 1–36.
- Foth, H. D., Ellis, B. G., et al. (1988). *Soil fertility*. John Wiley and Sons Inc.
- Gurobi Optimization, LLC (2023). Gurobi Optimizer Reference Manual.
- Haghverdi, A., Leib, B. G., Washington-Allen, R. A., Ayers, P. D., and Buschermohle, M. J. (2015). Perspectives on delineating management zones for variable rate irrigation. *Computers and Electronics in Agriculture*, 117:154–167.
- Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations research*, 12(3):450–459.
- Hamouda, Y. and Msallam, M. (2020). Variable sampling interval for energy-efficient heterogeneous precision agriculture using wireless sensor networks. *Journal of King Saud University-Computer and Information Sciences*, 32(1):88–98.
- Khosla, R., Westfall, D., Reich, R., Mahal, J., and Gangloff, W. (2010). Spatial variation and site-specific management zones. *Geostatistical applications for precision agriculture*, pages 195–219.
- Lauwens, B. and Downey, A. B. (2019). *Think Julia: how to think like a computer scientist*. O’Reilly Media.
- Ortega, R. A. and Santibanez, O. A. (2007). Determination of management zones in corn (*zea mays* l.) based on soil fertility. *Computers and Electronics in agriculture*, 58(1):49–59.
- ReVelle, C. S. and Swain, R. W. (1970). Central facilities location. *Geographical analysis*, 2(1):30–42.
- Roudier, P., Tisseyre, B., Poilvé, H., and Roger, J.-M. (2008). Management zone delineation using a modified watershed algorithm. *Precision Agriculture*, 9:233–250.
- Thom, W., Schwab, G., Murdock, L., and Sikora, F. (2003). Agr-16: Taking soil test samples. *Director of Cooperative Extension Service, University of Kentucky, Lexington, KY*.
- Toregas, C., Swain, R., ReVelle, C., and Bergman, L. (1971). The location of emergency service facilities. *Operations research*, 19(6):1363–1373.
- Toribio, G. (2022). Ndzi, ¿qué es y para qué sirve? *Cursos Teledetección*.
- Uyeh, D. D., Akinsoji, A., Asem-Hiablie, S., Basse, B. I., Osinuga, A., Mallipeddi, R., Amaizu, M., Ha, Y., and Park, T. (2022). An online machine learning-based sensors clustering system for efficient and cost-effective environmental monitoring in controlled environment agriculture. *Computers and Electronics in Agriculture*, 199:107139.

- Vera, J. R., Weintraub, A., Koenig, M., Bravo, G., Guignard, M., and Barahona, F. (2003). A lagrangian relaxation approach for a machinery location problem in forest harvesting. *Pesquisa Operacional*, 23:111–128.
- Weber, A. (1909). Über den standort der industrien, tübingen. *English Translation: The Theory of the Location of Industries*.
- Zeifman, L., Hertog, S., Kantorova, V., Wilmoth, J., et al. (2022). A world of 8 billion. *United Nations*.
- Zhang, N., Wang, M., and Wang, N. (2002). Precision agriculture-a worldwide overview. *Computers and electronics in agriculture*, 36(2-3):113–132.



6 — ANEXOS

6.1. Algoritmo 1

Algoritmo que procesa los datos de entrada para obtener los conjuntos y parámetros requeridos en la implementación del modelo

```

1  #Crear la data
2
3  Width,Length= size(data)
4
5  Nz= sum((Width - i+1)*(Length - j+1) for i in range(1,Width+1) for j in range(1,Length+1))
6
7  puntos = reshape(collect(1:Width*Length), Width, Length)
8  combinaciones = []
9
10 for ancho in 1:Width
11     for alto in 1:Length
12         if ancho * alto <= Width * Length
13             for fila in 1:Width-ancho+1
14                 for columna in 1:Length-alto+1
15                     combinacion = puntos[fila:fila+ancho-1, columna:columna+alto-1]
16                     push!(combinaciones, combinacion[:])
17                 end
18             end
19         end
20     end
21 end
22 # Eliminar combinaciones duplicadas
23 combinaciones_unicas = unique(combinaciones, dims=1)
24
25 # Generar matriz C de combinaciones únicas
26 C = zeros{Int, length(combinaciones_unicas), Length*Width}
27 for (i, combinacion) in enumerate(combinaciones_unicas)
28     for punto in combinacion

```

```

29     C[i, punto] = 1
30     end
31 end
32
33 println("Matriz C ok")
34
35 data2 = vec(data)
36
37 tethaz=zeros(Nz)
38
39 n=zeros(Nz)
40
41 tethat=0
42 for k in range(1,Nz)
43     println("calculando varianza zona$k")
44     prom=0
45     cont=0
46     suma=0
47     for i in range(1,Width * Length )
48         prom=prom+data2[i]*C[k,i]
49         cont=cont+C[k,i]
50         println(cont)
51     end
52     prom=prom/cont
53     for i in range(1,Width *Length )
54         if C[k,i]==1
55             point=(data2[i]-prom)^2
56             suma=suma+point
57         end
58     end
59     var=suma*1000/cont
60     tethaz[k]=var
61     n[k]=cont
62     if cont==42
63         tethat=float(var)
64     end
65 end
66 println("varianza ok")
67
68 #centroides
69 centr=["x" "y"]
70 for k in range(1,Nz)
71     x=0
72     y=0
73     contar=1
74     for i in range(1,Width*Length)
75         x=x+ubi[i,1]*C[k,i]
76         y=y+ubi[i,2]*C[k,i]
77         contar=contar+ C[k,i]

```

```
78     end
79     x=x/contar
80     y=y/contar
81     centr=vcat(centr, [x y])
82 end
83
84 #distancias
85 d=zeros((Nz,Nz))
86 for k in range(1,Nz)
87     print("calculando distancia fila $k")
88     di=zeros(Nz)
89     for i in range(1,Nz)
90         if d[i,k] == 0
91             X1=float(centr[k+1,1])
92             X2=float(centr[i+1,1])
93             Y1=float(centr[k+1,2])
94             Y2=float(centr[i+1,2])
95             dist=sqrt((X1-X2)^2 + (Y1-Y2)^2)
96             d[i,k]=dist
97             d[k,i]=dist
98         end
99     end
100 end
```

6.2. Implementación esquema jerárquico

Código utilizado para la implementación y solución del esquema jerárquico presentado en el capítulo 3.2 en Julia JuMP.

```

1  #Zonificacion
2
3  using JuMP, Gurobi, DataFrames, DelimitedFiles
4  #for ins in 1:9
5  # x=string(ins)
6  # println(x)
7  #end
8
9  FO=[]
10
11 alpha=0.9
12 v=tethaz
13 vT=tethat
14 c=transpose(C)
15 N=Width*Length
16 LI=0
17 LS=N
18 ins=1
19 ZON = Model(Gurobi.Optimizer)
20
21     #variables
22 @variables(ZON, begin
23     Q[1:Nz] >= 0, Bin
24 end)
25
26 for s in 1:N
27     @constraint(ZON, sum(c[s,z]*Q[z] for z in 1:Nz)==1)
28 end
29
30     #Homogeneidad de la particion
31 @constraint(ZON, sum(((n[z]-1)*v[z] + (1-alpha)*vT)*Q[z] for z in 1:Nz)<= vT*N*(1-alpha))
32
33     #Particion optima
34 @constraint(ZON,sum(Q[z] for z in 1:Nz)>=LI )
35 @constraint(ZON,sum(Q[z] for z in 1:Nz)<=42 )
36
37 @objective(ZON, Min, sum(Q[z] for z in 1:Nz))
38
39     # Resolver el modelo
40 optimize!(ZON)
41
42
43     # Imprimir resultados
44 println("Estado de la solución: ", termination_status(ZON))

```

```

45 println("Valor objetivo óptimo: ", objective_value(ZON))
46
47 op=Int(objective_value(ZON))
48 ops=JuMP.value.(Q)
49 zonas= zeros(op)
50 j=1
51 for x in 1:Nz
52     if JuMP.value(Q[x])==1
53         zonas[j]=x
54         j=j+1
55     end
56 end
57 Z=length(zonas)
58 x="1"
59 nombre="phZON0"*x*".csv"
60 writedlm(nombre,ops)
61
62 L=Int(Z)
63 for PP in 1:L
64     p= PP
65     LOC = Model(Gurobi.Optimizer)
66     @variables(LOC, begin
67         Xloc[1:L, 1:L] >= 0, Bin
68     end)
69
70     for i in 1:L
71         @constraint(LOC, sum(Xloc[i,z] for z in 1:L)==1)
72     end
73
74     for i in 1:L
75         for z in 1:L
76             @constraint(LOC, Xloc[z,z]>=Xloc[i,z])
77         end
78     end
79
80     @constraint(LOC, sum(Xloc[z,z] for z in 1:L)<=p)
81
82     PMED=0
83     for j in 1:L
84         for k in 1:L
85             i=Int(zonas[j])
86             z=Int(zonas[k])
87             PMED= PMED + v[i]*d[i,z]*Xloc[j,k]
88         end
89     end
90     @objective(LOC, Min, PMED )
91
92     # Resolver el modelo
93     optimize!(LOC)

```

```
94     # Imprimir resultados
95     println("Estado de la solución: ", termination_status(LOC))
96     println("Valor objetivo óptimo: ", objective_value(LOC))
97     push!(FO,[1,p,objective_value(LOC)])
98
99
100    Sensores= zeros(L+1,L+1)
101    for x in 1:L
102        Sensores[1,x+1]=zonas[x]
103    end
104    for x in 1:L
105        Sensores[x+1,1]=zonas[x]
106        for y in 1:L
107            Sensores[x+1,y+1]= JuMP.value.(Xloc[x,y])
108        end
109    end
110    end
111    x=string(Ins)
112    z=string(PP)
113    nombre="phLOC0"*x*z*".csv"
114    writedlm(nombre,Sensores)
115    optimo= objective_value(LOC)
116    nombre="phOP0"*x*z*".csv"
117    writedlm(nombre,optimo)
118 end
119
120 #
```

6.3. Implementación modelo integrado

Código utilizado para la implementación y solución del modelo integrado presentado en el capítulo 3.3 en Julia JuMP.

```

1 using JuMP, Gurobi, DataFrames, DelimitedFiles
2 alpha=0.9
3 v=tethaz
4 vT=tethat
5 c=transpose(C)
6 N=Width*Length
7 LI=0
8 LS=12
9 for insP in 1:LS
10     p=Int(insP)
11
12     REV = Model(Gurobi.Optimizer)
13
14     #variables
15     @variables(REV, begin
16         Xz[1:Nz,1:Nz] >= 0, Bin
17     end)
18     @variables(REV, begin
19         Qz[1:Nz] >= 0, Bin
20     end)
21
22     #Cada punto muestral debe pertenecer a una zona de la particion
23     for s in 1:N
24         @constraint(REV, sum(c[s,z]*Qz[z] for z in 1:Nz)==1)
25     end
26
27     #Homogeneidad de la particion
28     @constraint(REV, sum(((n[z]-1)*v[z] + (1-alpha)*vT)*Qz[z] for z in 1:Nz)<= vT*N*(1-alpha))
29
30     #Particion optima
31     @constraint(REV,sum(Qz[z] for z in 1:Nz)>=LI )
32     @constraint(REV,sum(Qz[z] for z in 1:Nz)<=LS )
33
34     #Matching Zonificacion y Localizacion
35     for z in 1:Nz
36         for j in 1:Nz
37             if z==j
38                 @constraint(REV, Xz[j,z]<=Qz[z])
39             end
40         end
41     end
42
43     #Yi debe ser estimada por algun z si esa zona i es escogida Q[i]==1

```

```

44   for i in 1:Nz
45       @constraint(REV, sum(Xz[i,z] for z in 1:Nz)==Qz[i])
46   end
47
48       #una zona es estimada solo por un sensor instalado
49   for i in 1:Nz
50       for z in 1:Nz
51           @constraint(REV, Xz[z,z]>=Xz[i,z])
52       end
53   end
54
55       #instalar un maximo de p dispositivos
56   @constraint(REV, sum(Xz[z,z] for z in 1:Nz)<=p)
57
58
59   @objective(REV, Min, sum(v[i]*d[i,z]*Xz[i,z] for i in 1:Nz, z in 1:Nz))
60
61       # Resolver el modelo
62   optimize!(REV)
63
64
65       # Imprimir resultados
66   println("Estado de la solución: ", termination_status(REV))
67   a=termination_status(REV)
68   println("Valor objetivo óptimo: ", objective_value(REV))
69   b=objective_value(REV)
70
71       # Obtener los valores de las variables Xz y Qz
72   Xz_values = JuMP.value.(Xz)
73   Qz_values = JuMP.value.(Qz)
74
75       # Escribir los resultados en un archivo CSV
76   BIPLOC="phBIPLOCC"*string(alpha)*string(p)*".csv"
77   BIPZON="phBIPZONN"*string(alpha)*string(p)*".csv"
78   BIPOP="phBIPOPP"*string(alpha)*string(p)*".csv"
79   writedlm(BIPLOC,Xz_values)
80   writedlm(BIPZON,Qz_values)
81   writedlm(BIPOP,objective_value(REV))
82
83 end

```

6.4. Datos

n	Coordenada x	Coordenada y	Índice	n	Coordenada x	Coordenada y	Índice
1	80	170	767.5	21	140	190	765.3
2	90	170	767.3	22	80	200	767.6
3	100	170	766.9	23	90	200	765.6
4	110	170	766.9	24	100	200	766.8
5	120	170	767.1	25	110	200	766.8
6	130	170	766.9	26	120	200	764.2
7	140	170	766.2	27	130	200	764.9
8	80	180	768.6	28	140	200	764.7
9	90	180	768.1	29	80	210	767.2
10	100	180	766.5	30	90	210	766.7
11	110	180	766.5	31	100	210	767.1
12	120	180	767.6	32	110	210	767
13	130	180	766.5	33	120	210	765.7
14	140	180	765.8	34	130	210	765.1
15	80	190	767.4	35	140	210	764.2
16	90	190	766.7	36	80	220	767.2
17	100	190	767.1	37	90	220	766.7
18	110	190	766.7	38	100	220	767.3
19	120	190	766	39	110	220	767
20	130	190	765.9	40	120	220	766.1

Tabla 6.1: Sección de 40 puntos de la muestra utilizada para implementación computacional

6.5. Resultados implementación computacional

La siguiente tabla muestra los resultados obtenidos al resolver ambos modelos en una instancia de 80 puntos para distintos valores de p con los que fue construido el gráfico 4.5.

Instancia	$Z_{jerarquico}$	$Z_{integrado}$	Diferencia porcentual
1	41,838.1	23,921.4	43 %
2	27,132.1	15,610.6	42 %
3	11,907.7	10,713.6	10 %
4	16,467.9	7,215.0	56 %
5	8,539.7	4,579.3	46 %
6	6,652.5	3,413.7	49 %
7	5,153.4	2,558.7	50 %
8	3,979.5	1,733.5	56 %
9	2,925.0	1,028.3	65 %
10	2,219.8	532.1	76 %
11	1,678.6	347.2	79 %
12	1,146.1	183.4	84 %
13	759.1	78.2	90 %
14	527.4	14.7	97 %
15	342.5	-	100 %
16	178.7	-	100 %
17	98.8	-	100 %
18	35.4	-	100 %
19	-	-	0 %

Tabla 6.2: Resultados de ambos modelos para una instancia de 80 puntos muestrales