

Predicción computacional para la reubicación física de apiarios en el estado de Jalisco

Héctor Sierra¹, Rosa Michel², María Puga², Karla Puga³, María Cárdenas⁴ y Gloria Cárdenas⁴
Maestría en Ciencias de la Computación¹, Sistemas y Computación², Ciencias Básicas³, Ciencias Económico
Administrativas⁵
Instituto Tecnológico de Ciudad Guzmán
Ciudad Guzmán, Jal.; México
h.sierra.gtz@gmail.com, [michel91_3, kenapn, karlalpn4, noniss]@hotmail.com, gloriae7cg@yahoo.com.mx

Abstract— The present work approaches the problematic of the mortality of bees in the state of Jalisco. Which is why, with this research we are looking through Data Mining using the decision tree tool, to predict the risk levels for bees, derivate from protected agriculture crops, irrigation agriculture and annual cycle temporary agriculture, existing in the closeness of the apiaries; obtaining by a software application, the risk level in the area, as a mapping of the places that represent a minor risk for bees, with the goal of contributing to decrease this problem that impacts economically, ecologically and socially. The present work addresses the problem regarding the mortality of bees in the state of Jalisco.

Keywords— *Data analysis, prediction, risk level, decision trees, apiaries, relocation.*

Resumen— El presente trabajo aborda la problemática referente a la mortandad de las abejas en el estado de Jalisco. Razón por la cual, con esta investigación se busca a través de Minería de Datos mediante la herramienta de árboles de decisión, predecir niveles de riesgo para las abejas, derivados de cultivos de agricultura protegida, agricultura de riego y agricultura de temporal de ciclo anual, existentes en las cercanías de los apiarios; obteniendo mediante una aplicación de software, el nivel de riesgo existente en la zona, así como un mapeo de los lugares que representen menor riesgo para las abejas, con la finalidad de contribuir a disminuir esta problemática que impacta económica, ecológica y socialmente.

Palabras claves— *Análisis de datos, predicción, niveles de riesgo, árboles de decisión, apiarios, reubicación.*

I. INTRODUCCIÓN

El cuidado de las abejas y la apicultura en general, es un tema de suma importancia actualmente a nivel mundial, por lo cual se requiere un alto grado de atención en esta actividad.

Las abejas, incluidas las abejas melíferas, abejorros y abejas solitarias, son el grupo prominente y económicamente más importante de polinizadores en todo el mundo. El 35% de la producción mundial de cultivos alimentarios depende de los polinizadores, lo que representa un valor anual de 153,000 millones de euros (Blacquièrre, Smagghe, van Gestel & Mommaerts, 2012).

El investigador apícola de la Universidad de Guadalajara, Macías (2016), expuso que “el 70 por ciento de las frutas y hortalizas que existen en la tierra dependen de las abejas para su producción, si no existiera este insecto habría escasez de alimentos, porque si ellas no polinizan las flores, no hay frutas”.

Es ahí donde radica la importancia de la apicultura a nivel mundial, ya que sin ella existiría un descontrol en la cadena alimenticia, además de la falta de polinización de las flores (Macías, 2016).

Como una propuesta para combatir la problemática de la mortandad de las abejas, se desarrolla el presente proyecto que contribuye, mediante el análisis de datos con la identificación de las características de vegetación, zonas urbanas, cuerpos de agua y cultivos de aguacate, temporal y agricultura protegida, a generar una herramienta de predicción de niveles de riesgos en los diferentes puntos de la zona de trabajo, para así evaluar su riesgo de acuerdo a la cantidad de hectáreas de cada uno

de los factores, causados por los productos químicos utilizados en los procesos de producción y post producción, que son una de las causas de muerte para las abejas, de tal manera que antes de mover físicamente el apiario a determinada zona, sea posible evaluar su riesgo a través de predicciones computacionales, con técnicas de minería de datos.

Los neonicotinoides son actualmente la clase más utilizada de insecticidas en todo el mundo. Estos pesticidas son cada vez más frecuentes en entornos terrestres y acuáticos. Los neonicotinoides son absorbidos por las plantas y transportados a todos los órganos, incluidas las flores, contaminando así el polen y el néctar, así como cualquier fluido producido por la planta. Hay una creciente preocupación por el impacto de estos pesticidas sistémicos, no solo en organismos no objetivo, especialmente polinizadores como abejas melíferas y abejas silvestres, así como en otros invertebrados terrestres y acuáticos, sino también en vertebrados, incluidos los humanos. (Mitchell et al., 2017).

El presente proyecto impacta de forma general a los apicultores del estado de Jalisco, principalmente a los que cuentan con apiarios en la zona de trabajo establecida de 15,552 kilómetros cuadrados, que se encuentra bajo las coordenadas delimitadas por los puntos de coordenadas geográficas grados decimales que se citan a continuación; punto 1: 20.570193°, -103.908013°, punto 2: 20.570193°, -102.986697°, punto 3: 19.052650°, -102.986697° y punto 4: 19.052650°, -103.908013°. Dichas zonas corresponden a 47 municipios del estado de Jalisco. Esta área de trabajo se dividió en 27 zonas de 24 por 24 kilómetros, con la finalidad de facilitar el análisis de la información y determinar la cantidad de hectáreas de cultivos de agricultura protegida, agricultura de riego (cultivo de aguacate) y agricultura de temporal, que fueron la base de la investigación de este proyecto. Cabe mencionar que la zona 25 no fue considerada, ya que incluye parte del territorio del estado de Colima, y el objeto de esta investigación está centrada solo al estado de Jalisco.

Con el uso del dispositivo GPS (Global Positioning System/Sistema de Posicionamiento Global) se obtuvieron las coordenadas de cada uno de los apiarios, y con la herramienta de representación cartográfica de ArcGIS, se obtuvieron imágenes satelitales provenientes del satélite LANDSAT, en las cuales se ubicaron las coordenadas previamente recabadas, además de la vegetación y cultivos antes mencionados, cuerpos de agua y zonas urbanas, existentes en cada una de las 27 zonas de trabajo establecidas de 576 kilómetros cuadrados. En la fig. 1 se muestra toda la zona de trabajo.

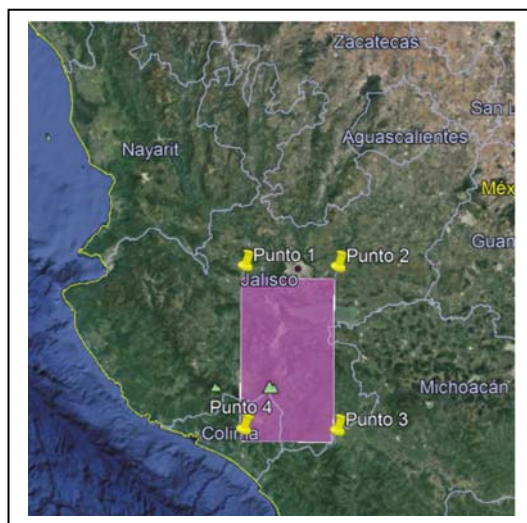


Fig. 1. Zona de trabajo total del proyecto.

En la fig. 2 se muestra la división de las áreas en las que se dividió toda la zona de trabajo.

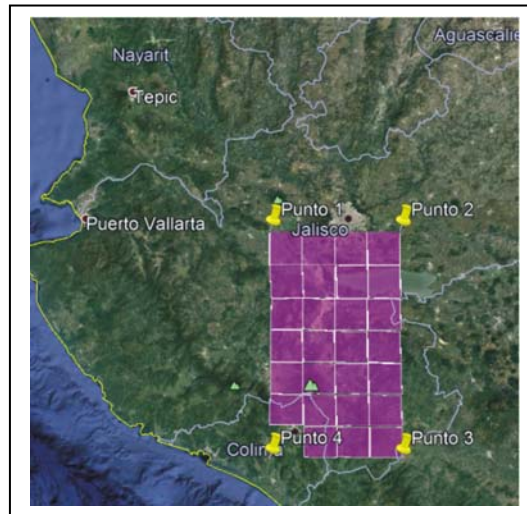


Fig. 2. División de las áreas de toda la zona de trabajo.

Posteriormente se procesó la información para obtener niveles de riesgo al posicionar el apiario en un nuevo punto dentro del área de trabajo de los 15,552 kilómetros cuadrados, teniendo así un panorama del nivel de riesgo que tendría el apiario en esa nueva ubicación, sin necesidad de moverlo físicamente.

La herramienta computacional desarrollada en el presente proyecto está basada en el uso de minería de datos con la técnica de árboles de decisión. La aplicación está conformada por dos módulos.

En uno de ellos se obtiene como resultado el nivel de riesgo de un punto en específico (tomando las coordenadas geográficas en grados decimales: latitud y longitud), así como las características y número de hectáreas de los tipos de cultivos que rodean al punto y que fueron considerados como parte de la entrada de datos de entrenamiento del árbol de decisión. Posteriormente estas coordenadas son procesadas para determinar el riesgo y mostrar la información relacionada con este punto.

El segundo módulo tiene como finalidad permitir al usuario hacer el cambio de ubicación de determinado apiario (seleccionado de una lista). Posteriormente la información relativa a este apiario se muestra como referencia. Se elige un nuevo punto del cual se desea conocer la predicción del nivel de riesgo, para determinar si es conveniente mover al apiario a estas nuevas coordenadas. Una vez hecho el cambio de ubicación, se procede a actualizar la nueva información en la base de datos correspondiente.

La clasificación de los niveles de riesgo está basada en parámetros proporcionados por la empresa Frutimelmex de Zapotlán S.P.R. de R.L. de C.V, la cual brindó la información relacionada con los porcentajes de hectáreas considerados para ellos como riesgo alto, medio y bajo, de los cultivos de agricultura protegida (invernaderos, malla sombra y macro túneles), agricultura de riego (cultivo de aguacate) y agricultura de temporal de ciclo anual, considerando la suma de porcentajes de cada tipo de factor en relación con el total de hectáreas que cuenta cada una de las zonas de trabajo. Para cada factor se le asignó una ponderación de riesgo, quedando de la siguiente manera:

- Cultivos de agricultura protegida (invernaderos, malla sombra y macro túneles): 40%.
- Agricultura de riego (cultivo de aguacate): 40%.
- Agricultura de temporal de ciclo anual: 20%.

De tal manera, que la clasificación quedó delimitada tal como se muestra en la tabla I, en el cual el porcentaje máximo tiende a 40%:

Tabla I. Tabla de clasificación de niveles de riesgos para las abejas.

Nivel de riesgo	Rango
Bajo	0.0% - 11.0%
Medio	11.1% - 27.0%
Alto	27.1% - 40.0%

Para poder lograr dichos resultados, se usaron herramientas informáticas, de minería y análisis de datos y lenguajes de programación, como R, programación web, MySQL, Weka, PHP y JavaScript.

II. TRABAJOS RELACIONADOS

La apicultura en México tiene una gran importancia socioeconómica y ecológica, ya que es considerada como una de las principales actividades pecuarias generadora de ingresos. Según datos del Servicio de Información Agroalimentaria y Pesquera (SIAP, 2016) a nivel nacional en el 2016 se registró una producción de miel de 55,358 toneladas, generando un ingreso de \$2,278'810,000. En el estado de Jalisco se obtuvo una producción anual en el mismo 2016 de 1,438 toneladas.

De manera muy usual esta actividad se asocia únicamente con la producción de la miel, polen, jalea real, propóleos. Sin embargo, las abejas son fundamentales para un equilibrio del medio ambiente, ya que éstas al obtener el alimento de las flores fomentan en las plantas la capacidad de fecundarse. Muestra de ello es que en 2015, la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) reportó que en México había 316 especies de plantas, de las cuales 286 se destinan para la alimentación y 80 como insumos para la industria, de los cuales el 80% depende de un polinizador para su producción.

Actualmente existe un problema muy grave con el alto índice de mortandad de las abejas, que debido a distintas situaciones ambientales y humanas, han perjudicado de manera sustancial a este grupo de fauna. El gran problema que tiene en jaque a los apicultores es el uso desmesurado e irresponsable de los plaguicidas en los campos agrícolas y su falta de regulación, según un estudio del Consejo Nacional de Ciencia y Tecnología (CONACYT) realizado en el año 2016. En estados como Chihuahua, Jalisco, Tamaulipas, San Luis Potosí, Michoacán y Yucatán, se reportaron la pérdida de 53 por ciento de las colmenas, (CONACYT, 2016).

Existen programas y grupos que promueven crear conciencia en la población, así como en los productores agrícolas y ganaderos, para frenar la mortandad de las mismas.

La problemática ha ocasionado principalmente la baja de la producción, aumento en los costos productivos y por ende la reducción significativa en las ventas de las empresas apícolas.

De acuerdo con Contreras (2015), investigadora del Centro Universitario del Sur, ubicado en Zapotlán El Grande, Jalisco, la mortandad de las abejas es un problema que se ha presentado desde hace más de una década, con mayor intensidad en ciudades de Europa y Asia, donde hay poblaciones de abejas que han disminuido hasta en un 85 por ciento. Así mismo, la Dra. Contreras, expone que "México, estaba de espectador, no obstante, en otoño de 2015 se comenzaron a conocer casos de mortalidad de abejas en algunas zonas del estado de entre 35% y hasta un 50%. Respecto al sur Jalisco, se estima que han muerto un 35% de colmenas, situación que es alarmante dado que Jalisco es el

segundo estado productor de miel a nivel nacional y México es de los tres primeros productores que exportan a nivel mundial”.

En el estado de Jalisco, la académica explicó que las autoridades han estado dando diagnósticos a los apicultores quienes han tomado muestras de sus abejas para analizarlas, indicando que las abejas han muerto porque éstos no les dan un adecuado manejo. Reconoció que hasta ahora no se tienen identificada una sola causa para este problema, sino que la suma de factores, como el uso de agroquímicos, el cambio climático, enfermedades de las abejas y la generación de monocultivos.

En el ámbito del análisis de datos y de manera específica en el campo de la minería de datos, se han integrado de manera significativa el uso de estas tecnologías al campo agrícola y pecuario, mediante la inteligencia de negocios y las herramientas de análisis predictivo, lo cual genera factores decisivos en el desarrollo de las potencialidades competitivas de la empresa.

En la investigación realizada se encontró que se ha estado trabajando en aplicaciones relacionadas con el análisis y minería de datos, así como predicción computacional, en el rubro agrícola-pecuario, las cuales están orientadas al uso de la información climática para realizar predicciones meteorológicas ajustadas según territorios y tipos de actividad, para la optimización de los costos de la cadena de suministro según las condiciones del entorno, el arraigamiento de las medidas preventivas que permitan deducir las incidencias y los costos derivados de plagas, entre otros usos. Enseguida se enlistan algunos programas encontrados, así como sus características, enfocados a la incorporación de análisis y minería de datos con sus diferentes herramientas.

Programa de Red de Estaciones Meteorológicas

Es una iniciativa originada en Argentina que se despliega por la provincia de Córdoba con el objetivo de brindar datos de utilidad para la toma de decisiones de los productores agropecuarios, pero también para toda la comunidad. A través de este programa, se intenta colaborar y facilitar la información que resulta de carácter estratégico para el trabajo en el campo, tanto en lo netamente productivo como en lo referido a aplicaciones periurbanas, entre otras utilidades. La información que brindan las estaciones se sube a un servidor en Internet, lo cual permite que los usuarios puedan acceder a los datos utilizando cualquier dispositivo que tenga conectividad, como un celular, laptop o computadora de escritorio. Una vez en el servidor, los datos son presentados en un software que permite visualizar datos actuales, graficar, generar reportes o realizar cálculos que sirven para la elaboración de predicciones meteorológicas, estudios climáticos ajustados a las variables fenológicas de los cultivos para determinar la oportunidad de la aplicación terrestre o aérea de fitosanitarios, entre otros usos. El acceso a los datos es instantáneo, gratuito y sin clave de acceso. En la fig. 3 se aprecia una captura de pantalla del programa antes mencionado:

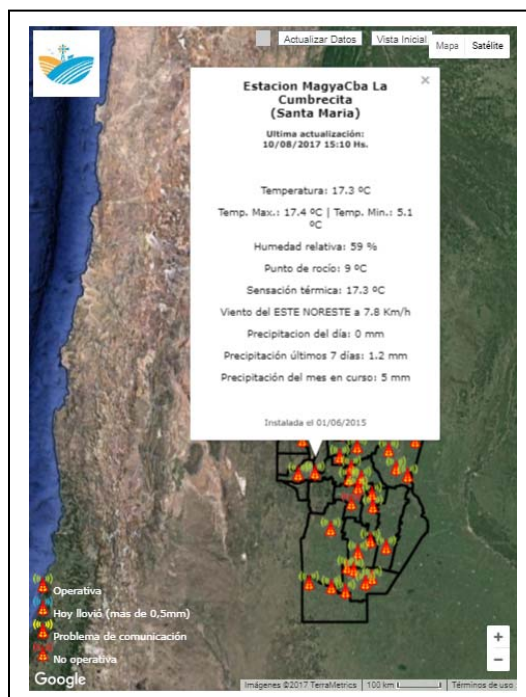


Fig. 3. Captura de pantalla de software APITECNIC.

ApisRAM

Es un modelo detallado de colonia de abejas en fase de desarrollo para EFSA (European Food Safety Authority / Autoridad Europea de Seguridad Alimentaria). Las oportunidades que ofrece este modelo y el enfoque utilizado para desarrollarlo son: integrar impactos de múltiples factores en la salud de las abejas; simular las interacciones entre los componentes de la abeja melífera y sus ambientes; y para predecir la dinámica futura del sistema en condiciones cambiadas de la gestión agrícola, apícola y el ambiente biológico y físico. En la medida en que estas oportunidades se pueden completar, están limitadas por aspectos de disponibilidad y gestión de datos. Los impactos de los factores estresantes o el manejo en las colonias de abejas son difíciles de predecir. Se caracterizan por interacciones de cadena larga y retroalimentaciones no lineales.

Este sistema es capaz de integrar los impactos de múltiples estresores (químicos, biológicos y ambientales), simular interacciones entre componentes (entre las abejas y el ambiente) y predecir sistemas dinámicos complejos (cambios de temperatura y disponibilidad de recursos).

El éxito de ApisRAM está restringido por los datos, esto es, que sin ellos no se puede obtener una simulación real. Estas restricciones de los datos están dadas por: disponibilidad y acceso, recopilación, colaboración y gestión, análisis y comunicación.

En el ámbito concreto aplicado a la predicción de nivel de riesgos en el estado de Jalisco, hasta el momento no se encontró en la investigación aplicaciones o herramientas informáticas que implementen técnicas de minería de datos aplicados a dichos procesos.

En el mercado actual del software y aplicaciones informáticas, existen diversas herramientas que ayudan a mejorar el control de cada una de las fases de los procesos productivos y post productivos apícolas. Cabe mencionar que ninguno de ellos ofrece al usuario la opción de predicción de niveles de riesgo para las abejas.

III. FUNDAMENTO TEÓRICO

La CONABIO reporta que hay 316 especies de plantas de las cuales 286 se destinan para la alimentación y 80 como insumos para la industria, de los cuales el 80% depende de un polinizador para su producción. Por su parte, investigadores de la UNAM identificaron 345 especies de plantas comestibles aprovechadas donde el 86% dependen de la polinización y estimaron el valor de la polinización en 43 mil millones de pesos.

Estos datos dan la certeza de la gran importancia que tienen las abejas en la conservación de la flora y la fauna, y por tal motivo es relevante implementar acciones que permitan dicha conservación.

Para el desarrollo de la investigación, fue indispensable el análisis y tratamiento de la información antes mencionada, mediante herramientas computacionales, que facilitaran su manejo. Por lo que fue necesaria la investigación y estudio de árboles de decisión en todos sus sentidos, partiendo del algoritmo C5.0, el cual es utilizado para generarlos.

Un árbol de decisión es una estructura de diagrama de flujo en la que el nodo interno representa la prueba de un atributo, cada rama representa el resultado de la prueba y cada nodo de hoja representa la etiqueta de clase. Un camino de raíz a hoja representa las reglas de clasificación. En análisis de decisión, un árbol de decisión y el diagrama de influencia estrechamente relacionado, se utiliza como una herramienta de apoyo a la decisión visual y analítica, para los valores esperados (Ran & Wang, 2015).

La metodología de árbol de decisión es un método de minería de datos, comúnmente utilizado para establecer sistemas de clasificación basados en múltiples covariables o para desarrollar algoritmos de predicción para una variable objetivo. Este método clasifica una población en segmentos tipo rama, que construyen un árbol invertido con un nodo raíz, nodos internos y nodos hoja (Song, 2015).

De acuerdo a la literatura revisada y el tipo de problema de investigación realizada, se optó por utilizar el algoritmo C5.0, el cual divide la muestra en función del campo que ofrece la máxima ganancia de información. Las distintas sub muestras definidas por la primera división se vuelven a dividir, por lo general basándose en otro campo, y el proceso se repite hasta que resulta imposible dividir las sub muestras de nuevo. Por último se vuelven a examinar las divisiones del nivel inferior, y se eliminan o podan las que no contribuyen significativamente con el valor del modelo.

Para entrenar un modelo C5.0, debe existir un campo categórico objetivo (por ejemplo, nominal u ordinal) y uno o más campos "Entrada" de cualquier tipo. Se ignorarán los campos establecidos en "Ambos" o "Ninguno". Los tipos de los campos utilizados en el modelo deben estar completamente instanciados. También se puede especificar un campo de ponderación. Los modelos C5.0 son bastante más robustos cuando aparecen problemas como datos perdidos y un número elevado de campos de entrada. Por lo general no precisan de largos tiempos de entrenamiento para calcular las estimaciones. Además, los modelos C5.0 suelen ser más fáciles de comprender que algunos tipos de modelos, ya que la interpretación de las reglas derivadas del modelo es muy directa. C5.0 también ofrece el eficaz método del aumento para obtener una mayor precisión en tareas de clasificación.

Para el desarrollo del trabajo de investigación, fue de suma importancia adquirir conocimientos y datos referentes a las abejas, así como las tecnologías a utilizar para la generación del software. Dichos tópicos fueron: recomendaciones para la instalación de los apiarios, movilidad de los apiarios, minería de datos, árboles de decisión, R language, Weka, PHP (Hypertext Preprocessor), JavaScript y MySQL.

IV. METODOLOGÍA

De acuerdo al tipo de problema que se resolvió con este trabajo, se llevó a cabo una investigación aplicada. Ya que mediante el uso de la tecnología computacional, específicamente con el análisis y minería de datos, se abordó una problemática mundial y de gran importancia, como lo es el riesgo que viven las abejas con su ubicación en zonas rodeadas de cultivos de agricultura protegida (invernaderos, malla sombra y macro túneles), agricultura de riego (cultivo de aguacate) y agricultura de temporal de ciclo anual.

La población está conformada por los apiarios del estado de Jalisco, que son el foco de atención de la investigación. La muestra contemplada serán los apiarios de la empresa Frutimielmex, que se encuentra en dicho estado de la república Mexicana.

Con respecto a las técnicas y herramientas de estadísticas utilizadas en análisis y minería de datos, la que se utilizó en esta investigación es la conocida como árboles de decisión, ya que se tendrá una gran variedad de datos referidos a los factores de riesgos definidos que rodea a las diferentes coordenadas generadas por el software en cuestión, que fue necesario para analizar y evaluar los resultados que se arrojan siguiendo las distintas opciones, para poder predecir el nivel de riesgo en determinado punto, para así el usuario determine si decide reubicar los apiarios.

De acuerdo al tipo de problema que se logró resolver con el trabajo, se llevó a cabo una investigación aplicada. Para Murillo (2008), la investigación aplicada recibe el nombre de “investigación práctica o empírica”, que se caracteriza porque busca la aplicación o utilización de los conocimientos adquiridos, a la vez que se adquieren otros, después de implementar y sistematizar la práctica basada en investigación. El uso del conocimiento y los resultados de investigación que da como resultado una forma rigurosa, organizada y sistemática de conocer la realidad. Es decir, es aquella que se realiza para la adquisición de nuevos conocimientos, dirigida hacia un objetivo o fin práctico, que responda a una demanda específica y determinada, y que en el campo que se está aplicando, que en este caso específico se busca a través de tecnología resolver el problema.

La investigación aplicada está fundamentada en los resultados de la investigación básica.

Los resultados de la investigación aplicada son considerados jurídicamente como derechos y en el marco nacional e internacional asumen la forma de propiedad intelectual, bajo sus diversas modalidades: patente, marca, modelo de réplica, dibujo industrial (Carvajal, 2002).

La población está conformada por los apiarios, que son el foco de atención de la investigación. Así mismo se hicieron pruebas mediante la movilidad de los apiarios, esto con base la analítica de la información recabada previamente. La muestra contemplada para la investigación fueron los apiarios del estado de Jalisco, ubicados en la zona delimitada en el capítulo I de este documento, pertenecientes a la empresa Frutimielmex.

El procedimiento a seguir para llevar a cabo la investigación a buen término se basó en lo siguiente:

- Investigación previa. En la que se recabó la información de las coordenadas de ubicación de los apiarios de la empresa caso de estudio; además de los tipos de vegetación alrededor de los apiarios, cuerpos de agua, zonas urbanas, cultivos de agricultura protegida (invernaderos, malla sombra y macro túneles), agricultura de riego (cultivo de aguacate) y agricultura de temporal de ciclo anual, para identificar los factores de riesgo para las abejas, de tal manera que se puedan generar mapas que reflejen toda esta información.
- Análisis. Se investigaron las características de las diferentes técnicas existentes para análisis y minería de datos, para definir la que más se adecuó al proyecto: árboles de decisión.

- **Diseño.** Se diseñaron las interfaces del software en el que se obtienen la predicción de los niveles de riesgo para las abejas, basándose en las coordenadas y los factores de riesgos definidos previamente.
- **Desarrollo.** Se implementó el software donde se aplicó la técnica de árboles de decisión, para obtener, mostrar y contrastar los niveles de riesgo para la posible reubicación física de los apiarios.
- **Prueba.** Se puso a prueba el software para procesar los resultados obtenidos y contrastarlos con las distintas ubicaciones.
- **Documentación.** Se elaboró un informe final del proyecto, el impacto económico y algunos artículos científicos en los que se difundan los resultados obtenidos.

V. RESULTADOS

Derivado de las pruebas realizadas, se obtuvieron los siguientes resultados:

En la fig. 4 se muestra la gráfica correspondiente, donde los datos del eje vertical presentan los porcentajes y los datos en el eje horizontal representan cada una de las zonas de trabajo. En la gráfica se aprecian numeradas las zonas del 1 al 28, excepto la zona 25, dado que esta corresponde en su totalidad al estado de Colima, y no es objeto de estudio de esta investigación, por lo tanto, se trabajó con 27 zonas correspondientes al estado de Jalisco. Posteriormente, se realizó una comparativa de cada una de las 27 zonas efectivas de trabajo de 576 kilómetros cuadrados, en la cual se expresa el porcentaje que existe de hectáreas de cultivos de agricultura protegida (invernaderos, malla sombra y macro túneles), agricultura de riego (cultivo de aguacate) y agricultura de temporal de ciclo anual, en relación con el total de la zona.

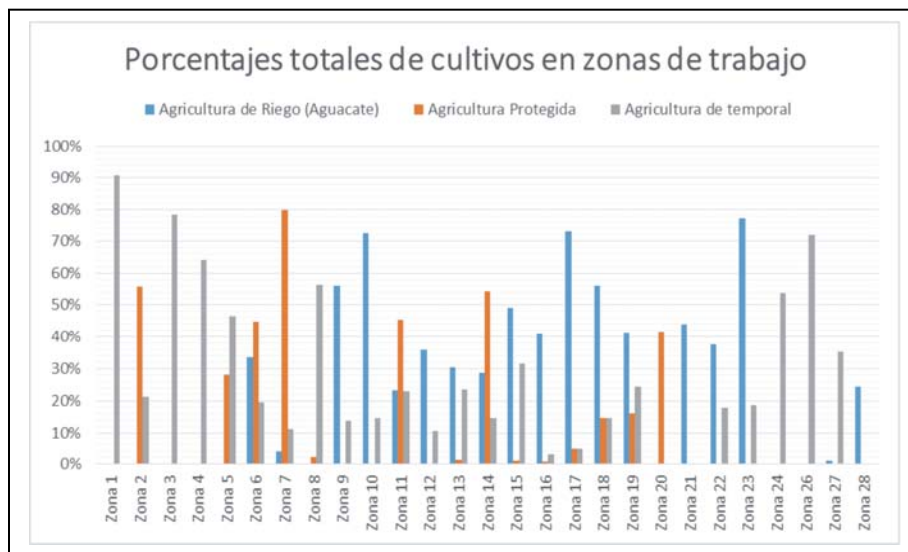


Fig. 4. Comparativa de zonas de trabajo con porcentajes totales.

De acuerdo a la ponderación dada a cada uno de los cultivos considerados como factores de riesgo descritos en el apartado I. INTRODUCCION, se generó la gráfica para visualizar de manera clara los diferentes porcentajes de cultivo. En la fig. 5 se muestra la gráfica correspondiente, donde los datos del eje vertical representan el porcentaje ponderado que tiende a 40% y los datos en el eje horizontal cada una de las 27 zonas de trabajo.

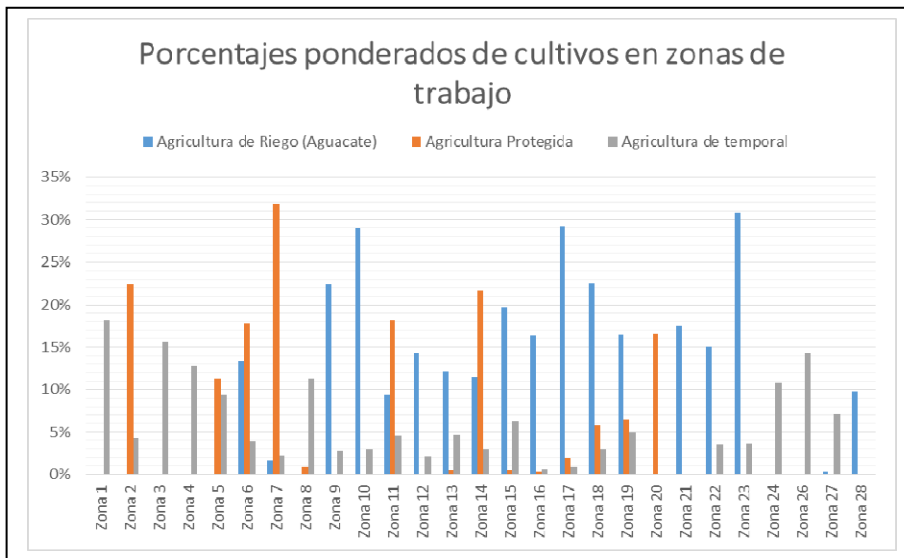


Fig. 5. Comparativa de zonas de trabajo con porcentajes ponderados.

Así mismo, se graficaron los niveles de riesgo identificados de forma general en cada una de las zonas, esto derivado únicamente del análisis de datos, tomando como referencia los porcentajes de la Tabla I (presentada en la sección de Introducción), sin tomar en cuenta los resultados obtenidos del árbol de decisión. En la fig. 6 se muestra la gráfica, donde los datos del eje vertical representan el nivel de riesgo de la zona, de tal manera que 1 es riesgo bajo, 2 riesgo medio y 3 riesgo alto; y los datos en el eje horizontal representan cada una de las 27 zonas de trabajo.

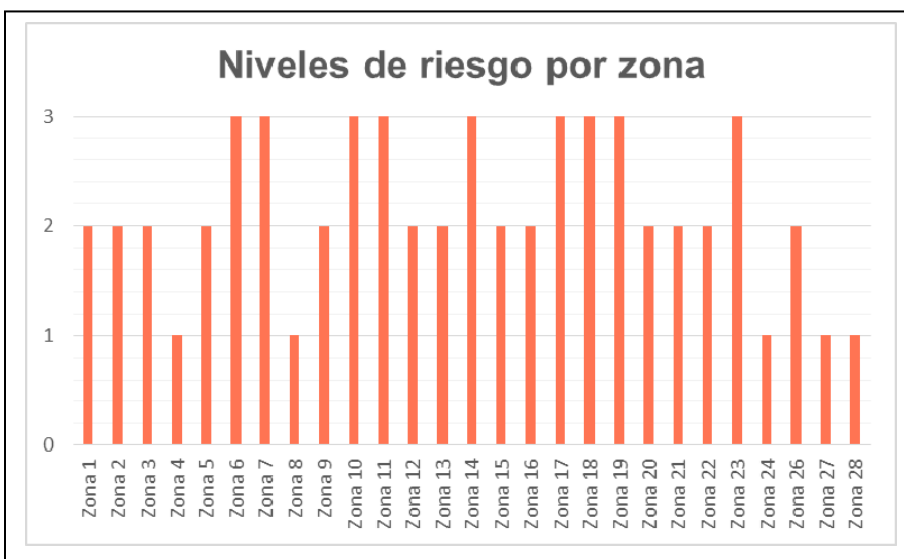


Fig. 6. Niveles de riesgo por zona.

De manera específica, y donde entró en juego el papel del árbol de decisión, fue mediante el análisis de los resultados de las predicciones, tomando como referencia 1 zona de trabajo en donde el nivel de riesgo en general es considerado como Alto, y para el cual se analizaron 6 puntos de coordenadas aleatorias ubicadas dentro de la zona, obteniendo resultados derivados de la predicción y que están ciertamente influenciados por el nivel de riesgo de las zonas vecinas, respecto a la zona analizada. La zona elegida fue la 23; en la fig. 7 se muestran dicha zona:

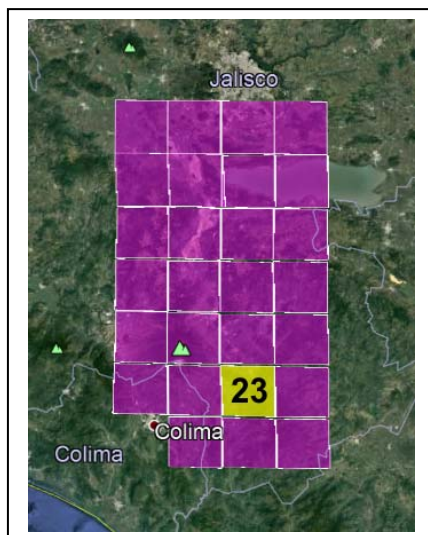


Fig. 7. Zona elegida para el análisis de riesgo alto.

Tomando como referencia la figura anterior, a continuación se muestran los datos obtenidos para la zona analizada (marcada con el número 23), donde por una parte en la tabla II se muestran las coordenadas de cada uno de los 6 puntos, así como el nivel de riesgo obtenido de la predicción del árbol de decisión. Y por otra en la fig. 8 se muestra la gráfica correspondiente al análisis, en la cual los datos del eje vertical representan el nivel de riesgo de la zona, donde 1 es riesgo bajo, 2 riesgo medio y 3 riesgo alto; y los datos en el eje horizontal representan cada una de los 6 puntos analizados.

Tabla II. Resultado del análisis de riesgo de la zona 23

Zona 23 - Riesgo Alto			
Punto	Latitud	Longitud	Riesgo
1	19.477226	-103.421627	Alto
2	19.303934	-103.408282	Alto
3	19.296921	-103.320519	Alto
4	19.383061	-103.224217	Bajo
5	19.475717	-103.226251	Medio
6	19.481283	-103.314895	Alto

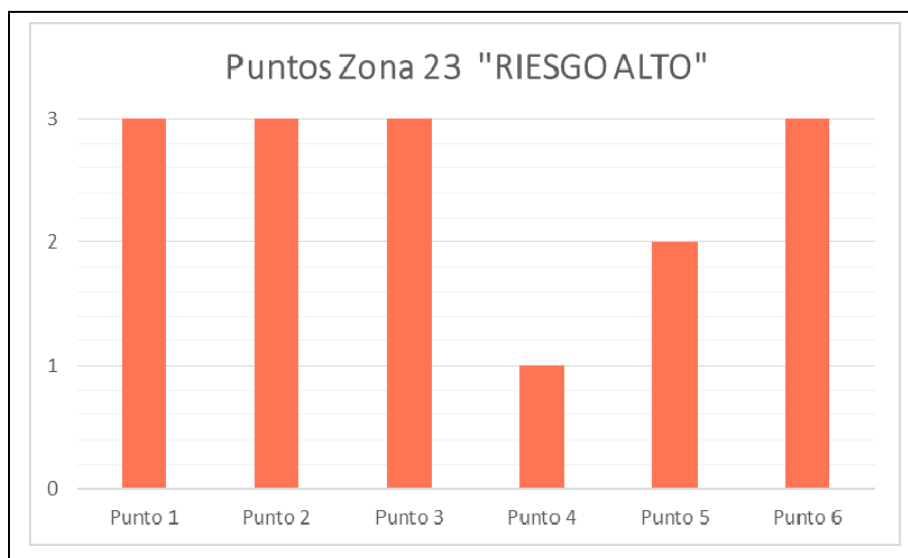


Fig. 8. Gráfica de puntos zona 19.

VI. CONCLUSIONES

Mediante la predicción computacional a través de árboles de decisión, la cual es una técnica de minería de datos (descrita en la sección de FUNDAMENTO TEORICO), se pueden comprobar diferentes niveles de riesgo para las abejas, considerando como factores los cultivos de agricultura protegida, agricultura de riego (cultivo de aguacate) y agricultura de temporal, los cuales se tomaron como tal debido a dos razones: la primera debido al uso desmesurado de agroquímicos y sus derivados en este tipo de agricultura, lo cual está sustentado en investigaciones previas; y la segunda, dada la información por parte de la empresa Frutimielmex, que con base a su experiencia en la producción apícola, identifican como factores de riesgo para las abejas los tipos de agricultura antes mencionados.

Dada la investigación realizada, pruebas hechas y los resultados obtenidos, se puede concluir que los niveles de riesgo en determinada zona, ya sea considerada como riesgo alto, medio o bajo, pueden cambiar según el punto elegido dentro de la misma zona; esto debido a la influencia que tienen los niveles de riesgo de las zonas vecinas, y por los procedimientos internos que realiza el árbol de decisión, al tomar como parámetro de entrada las coordenadas de cierto punto, procesarlas y emitir como resultado la predicción del nivel de riesgo.

Es importante mencionar que como trabajo futuro, se pueden utilizar las bases obtenidas en esta investigación, con la herramienta de árboles de decisión, para generar predicciones en otras áreas de trabajo establecidas previamente, teniendo como base el análisis de la cantidad de hectáreas de los cultivos determinados como factores de riesgo.

REFERENCIAS

- Blacquièrè, T., Smagghe, G., van Gestel, C., y Mommaerts, V. (2012). Neonicotinoids in bees: a review on concentrations, side-effects and risk assessment. *Ecotoxicology*, 21(4), 973-992. doi: 10.1007/s10646-012-0863-x
- Díaz, A. (2016). Jalisco enmielado. Mural. Consultado el 3 Mayo 2018, de: <https://www.mural.com/aplicacioneslibre/articulo/default.aspx?id=783936&md5=1e3b31496020bb9de80f21226c11074e&ta=0dfdbac11765226904c16cb9ad1b2efe>
- El Informador. (2016). Jalisco, tercer lugar nacional como productor de miel. Consultado el 2 de Mayo 2018, de: <http://www.informador.com.mx/economia/2016/679301/6/jaliscotercer-lugar-nacional-como-productor-de-miel.htm>
- Franco-Arcega, A., Carrasco-Ochoa, J., Sánchez-Díaz, G., y Martínez-Trinidad, J. (2013). Decision Tree based Classifiers for Large Datasets. Consultado el 3 de Mayo 2018, de: <http://www.cys.cic.ipn.mx/ojs/index.php/CyS/article/view/1485/1495>
- Matsudo, N. L. (2001). Árboles de decisión, una técnica de Data Mining, desde una perspectiva informática y estadística (Tesis Licenciatura). Consultado el 3 de Mayo 2018, de: <https://www.dc.uba.ar/academica/tesis-de-licenciatura/2001/>
- Mitchell, E., Mulhauser, B., Mulo, M., Mutabazi, A., Glauser, G., y Aebi, A. (2017). A worldwide survey of neonicotinoids in honey. *Science*, 358(6359), 109-111. doi: 10.1126/science.aan3684
- Ran, C., y Wang, X. (2015). Building a Decision Tree Model for Campus Information Score Based on the Algorithm C5.0. *Applied Mechanics And Materials*, (719-720), 805-811. doi: 10.4028/www.scientific.net/amm.719-720.805
- R Foundation. (2018). The R Project for Statistical Computing. Consultado el 30 Abril 2018, de: <https://www.r-project.org/>
- Secretaría de Desarrollo Rural [SEDER]. (2014). Jalisco: Gigante Agroalimentario, Edición 2014. Consultado el 2 de Mayo 2018, de: <http://seder.jalisco.gob.mx/content/jalisco-gigante-agroalimentario>
- Song, Y. Y., y Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*, 27. 130–135. doi: 10.11919/j.issn.1002-0829.215044
- Tecnocustomer.powweb.com. (2016). Consultado el 3 Mayo 2018, de: <http://tecnocustomer.powweb.com/apitecnic/index.php>
- Weka 3. (2018). Data Mining with Open Source Machine Learning Software in Java Cs.waikato.ac.nz. Consultado el 30 de Abril 2018, de: <https://www.cs.waikato.ac.nz/ml/weka/>