

Relación entre las explotaciones ganaderas y la contaminación por nitratos de los acuíferos en Cataluña.

RESUMEN.

El estudio tiene como objetivo determinar si existe una relación entre la concentración de explotaciones ganaderas intensivas y la contaminación de acuíferos por nitratos. Para ello, se han cruzado datos del registro de explotaciones ganaderas de la Generalitat de Catalunya, análisis de aguas de acuíferos en riesgo de contaminación por nitratos realizados por la ACA, y la superficie de las comarcas.

La metodología incluye la depuración de datos utilizando SQL, y un análisis estadístico con Python para examinar la distribución de las variables de estudio. La información se integra en Power BI para obtener visualizaciones de los resultados. Estas fases se desarrollan de manera simultánea.

Las variables analizadas son la concentración de URM, la concentración de nitratos de las deyecciones ganaderas, y los niveles de nitratos en aguas subterráneas. Se parte de la hipótesis de que una mayor concentración de deyecciones ganaderas afecta negativamente la calidad de las aguas subterráneas.

El análisis muestra que las variables no siguen una distribución normal, y que sí existe una correlación positiva fuerte entre la concentración de URM y los niveles de nitrógeno de las deyecciones. Sin embargo, la correlación entre estas variables y la contaminación de los acuíferos tiene un poco más débil de lo esperado en sentido positiva.

En conclusión del resultado de este estudio es que hay otros factores que influyen en la contaminación por nitratos, como es la alta densidad de cultivos en invernaderos, o la presencia de residuos industriales.

Para mejorar el presente estudio se propone realizar un análisis comarcal individual. También propondría realizar una búsqueda de más datos de análisis de nitratos en acuíferos.

INTRODUCCIÓN.

La actividad ganadera en Cataluña es una parte fundamental de la economía rural, y se caracteriza por su diversidad de especies, y modernización. Además la región es la principal productora de carne de cerdo en España, representando aproximadamente el 55% de la producción nacional

Una de las ventajas es su alta modernización por la implementación de nuevas tecnologías en el manejo de las granjas, la alimentación animal y el control sanitario, que ha favorecido el aumento de la productividad, y la mejora de la calidad de los productos.

Un inconveniente, es que genera problemas medioambientales, como la contaminación del agua y del suelo por nitratos. Los nitratos son comúnmente encontrados en la naturaleza, especialmente en suelos. Son componentes clave de muchos fertilizantes, proporcionando nitrógeno disponible para las plantas, ya que es un nutriente fundamental para el crecimiento y desarrollo de las plantas.

Los nitratos son también el principal residuo de las granjas, generado por las deyecciones ganaderas. Estos residuos son estiércol y purines. Se suelen extender en los campos de cultivo como fertilizantes. Un exceso de sus aportaciones es perjudicial y pueden contaminar los ríos y acuíferos, donde llegan por escorrentía superficial y lixiviación.

Los ríos y acuíferos son muy importantes para la vida, ya que son las principales fuentes de agua dulce en nuestro planeta, necesarios para la vida. Su disponibilidad se ve amenazada por esta contaminación y el cambio climático.

Además por el cambio climático, están sucediendo episodios importantes de sequía que provocan un aumento de la concentración de nitratos en las aguas subterráneas, que se alternan con períodos de lluvias intensas y torrenciales que provocando escorrentía superficial.

Por todo ello, la gestión sostenible del agua debe ser una prioridad para garantizar el acceso a agua de calidad para las generaciones presentes y futuras. Para abordar estas problemáticas se han implementado normativas medioambientales, y se promueven prácticas de producción sostenible y ecológica para minimizar el impacto.

El objetivo de este proyecto es evaluar la influencia de la concentración de granjas ganaderas en la contaminación de las aguas subterráneas por nitratos en Catalunya. A través de un análisis de datos, se pretende identificar patrones y tendencias en la contaminación.

METODOLOGIA.

Para llevar a cabo el presente estudio, los datos necesarios se han obtenido de diferentes fuentes. En primer lugar, se descargaron los datos del registro de explotaciones ganaderas de la web de Dades Obertes de Catalunya (DACC, 2024).

Después se descargan los datos de registro de analíticas de aguas subterráneas de zonas vulnerables de ser contaminadas por nitratos, realizado por el ACA (Agència Catalana de l'Aigua, 2019). Además, se obtienen las superficies en km2 de las diferentes comarcas de Cataluña (Idescat, 2022).

Primero de todo, damos forma a la información del registro de explotaciones ganaderas en SQL hasta obtener el modelo en estrella para facilitar las consultas y las visualizaciones.

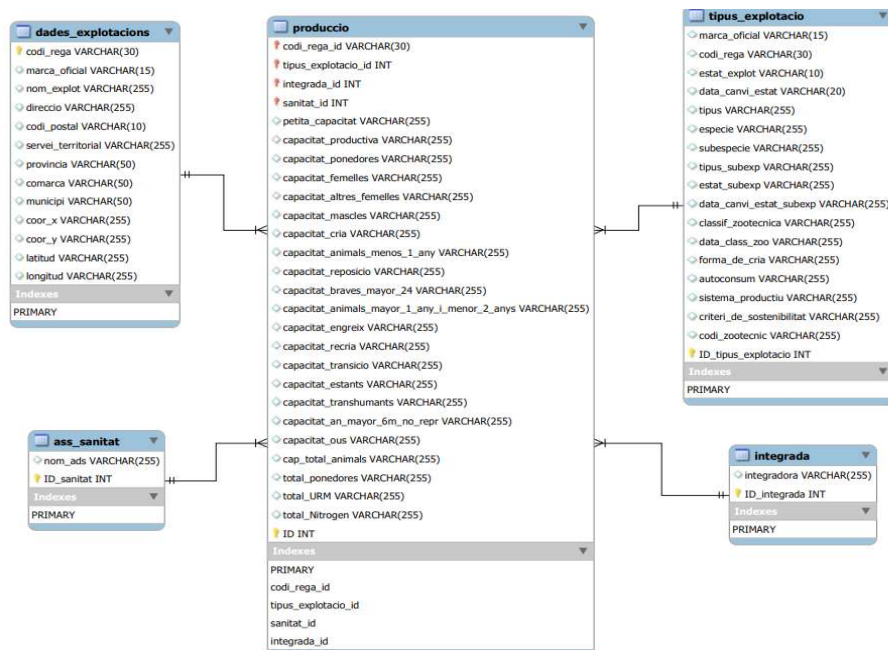


Imagen1. Modelo en estrella de la base de datos de explotaciones ganaderas de Catalunya.

Las variables que nos interesan para el estudio, se visualizarán a través de Power BI para apreciar como es su distribución geográfica. Para ello, previamente se unifican todos los datos, integrando los nitratos de los acuíferos y la superficie de las comarcas con los datos de las explotaciones ganaderas. El resultado del análisis esperamos corroborar la afirmación, ya conocida, que a mayor concentración de granjas intensivas mayor es la contaminación por nitratos de los acuíferos.

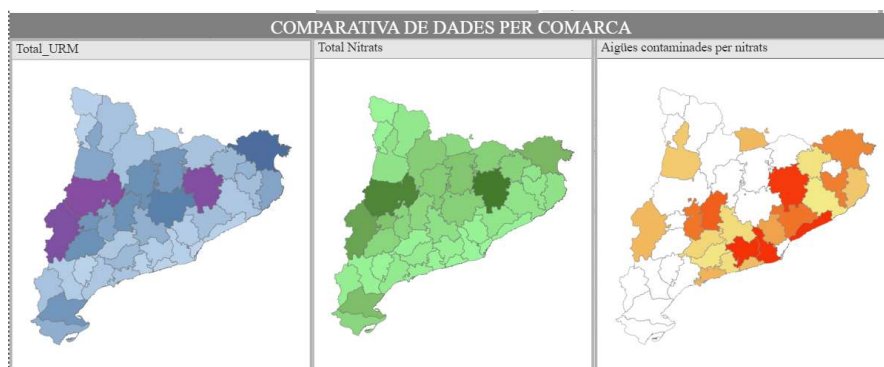


Imagen2. Distribuciones en el territorio de los datos de carga ganadera, residuos producidos en las granjas como nitratos y los niveles de nitratos en aguas subterráneas.

Desde los mapas personalizados de Power BI, hemos generado los 3 mapas de la *Imagen2*. En el que visualizamos la distribución de la carga ganadera en el territorio (azul), concentración nitratos deyecciones (verde) y nitratos en aguas subterráneas (naranja). Del análisis por comarcas podemos apreciar que hay comarcas con alta carga ganadera y altos nitratos en aguas subterráneas, como puede ser la comarca de Osona. También, hay comarcas donde las concentraciones son opuestas, como puede ser la comarca Maresme

Después de la prospección de los datos, pasamos a realizar el análisis estadístico. Éste puede ser numérico y/o gráfico. En este punto, nos preguntamos si nuestras variables siguen una distribución normal para seleccionar los métodos más apropiados de análisis. Entonces, realizamos el Test de Shapiro-Wilk ([Machine Learning Mastery, s.f.](#)) que tiene como hipótesis nula: la muestra sigue una distribución gaussiana, e hipótesis alternativa: la muestra no sigue una distribución normal, y asume que las observaciones de cada muestra son independientes y se distribuyen de manera idéntica.

Sabiendo que nuestra distribución no es gaussiana, decidimos aplicar el Test de Correlación de Spearman, el cual comprueba si dos muestras tienen una relación monótona, que tiene como hipótesis nula: las dos muestras son independientes, e hipótesis alternativa: Existe una dependencia entre las muestras., y asume que las observaciones en cada muestra son independientes y se distribuyen de manera idéntica, y por otro lado, podemos demostrar numéricamente que como es la fuerza y la dirección de la correlación, Appinio (2024).

Para que quede más claro y corroborar los datos se hacen visualizaciones de relación, distribución y correlación con los siguientes gráficos: *gráfico de dispersión, un Pairplot: y matriz correlación de Spearman.*

RESULTADOS.

A continuación, se tabulan los datos obtenidos del test de Normalidad de Saphiro-Wilk, y del test de correlación de Spearman.

Tabla1. Resultados del Test de Normalidad de Shapiro-Wilk.

VARIABLES	Estadístico	p-Valor	Resultado
Concentración URM	0,846	0,003	Probably not Gaussian
Concentración Nitrógeno	0,889	0,018	Probably not Gaussian
Valor mg/l NO3	0,728	0,000	Probably not Gaussia

Tabla2. Resultados del Test de correlación de Spearman.

CORRELACION	Estadístico	p-Valor	Resultado
Concentración URM - Concentración Nitrógeno	0,912	0,000	Probably dependent
Concentración Nitrógeno - Valor mg/l NO3	0,020	0,931	Probably independent

El resultado del test de Shapiro-Wilk es que las muestras no siguen una distribución gaussiana.Por eso, aplicamos el Test de Correlación de Spearman. Los resultados del mismo son que las variables de concentración tienen un p-Valor de cero. Esto significa que son dos variables dependientes. Para las variables concentración de Nitrógeno y Valor mg/l NO3, el valor de p-Valor es de 0.931. Esto significa que son variables independientes (*Tabla2*). Si observamos la Imagen 5, donde se gráfica un pairplot y una matriz de correlación, observamos que la representación concuerda con estos resultados.

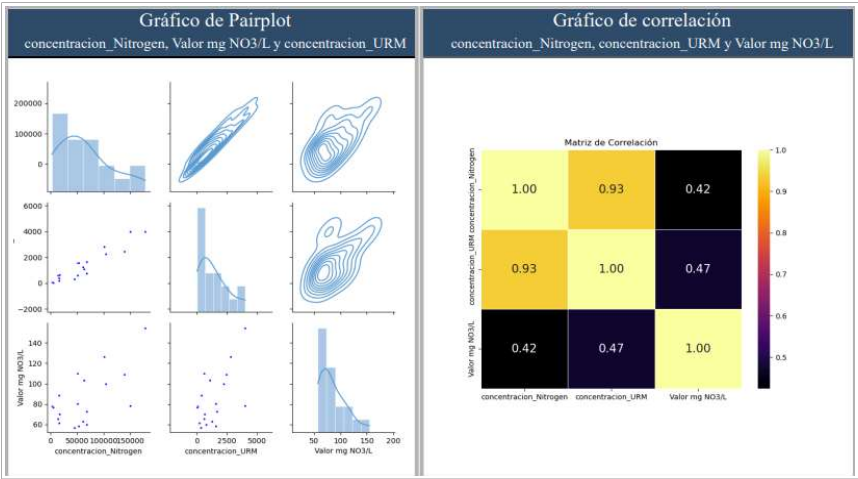


Imagen 3. Izquierda. Gráfico Pairplot. Derecha Matriz de correlación de todas nuestras variables

DISCUSIÓN.

Según el test de normalidad de Shapiro-Wilk, las variables estudiadas no siguen una distribución normal. Partiendo de esta información, aplicamos métodos estadísticos no paramétricos, como es el Test de correlación de Spearman, el cuál nos determina que las variables referentes a las granjas tienen una fuerte relación en sentido positivo. Lo cual es lógico, ya que más animales generan más desechos. En relación a la correlación entre la concentración de URM y el valor de NO₃ nos determina una relación más débil en sentido positivo.

En la *imagen3*, podemos ver cómo los valores de NO₃ muy altos se dispersan, por lo que podemos decir que hay pocas comarcas con altas concentraciones de Nitrógeno y de contaminación de NO₃. Aunque se puede observar que la mayor concentración de NO₃, se encuentra por debajo del umbral de los 90mg/l, lo que significa que existe una tendencia de cúmulo en valores de baja concentración de nitrógeno. Esto también lo podemos notar en el histograma de la imagen donde se aprecia que las variables de estudios presentan un sesgo a la izquierda.

Si analizamos en la misma Imagen3, el resultado de la matriz de correlación nos corrobora lo que hemos visto hasta ahora. Las variables concentración de URM y de Nitrógeno presentan una fuerte relación en sentido positivo, con un valor de 0.91, como ya se ha comentado. Y la relación entre estas dos variables y la variable objeto de estudio, Valor NO₃, presentan una relación un poco más débil (0.42 - 0.47), eso sí en sentido positivo.

Inicialmente partimos de la premisa que existe una fuerte influencia entre la concentración de carga ganadera, y que mayor contaminación por nitratos de las aguas residuales. Con el estudio realizado, esta relación está presente con una relación menor a la esperada.

Este estudio es el inicio del análisis del origen de la contaminación de aguas subterráneas por nitratos, que se podría mejorar en gran medida haciendo un análisis individual por comarca o municipio, ampliando las bases de datos de analíticas de nitratos en aguas subterráneas, con la idea de cubrir todo el área de estudio. Interesante, intentar identificar otros factores que afectan a la contaminación aparte de la carga ganadera, como son concentración de explotaciones agrícolas y la industria.

CONCLUSIONES.

La conclusión del estudio realizado nos dice que elevadas concentración de ganadera intensiva ejercen una influencia media en la contaminación de las aguas subterráneas. Esto puede que ser debido a la influencia de otros factores que provocan contaminación, como son la alta concentración de explotaciones agrícolas (invernaderos de flores o huertas) y la industria.

Por otro lado, para ampliar el estudio y mejorarlo sería interesante ampliar las bases de datos de NO₃ en aguas subterráneas, y como ya hemos comentado, relacionar los datos por municipio en lugar de comarca.

BIBLIOGRAFIA.

DACC. (2024, 8 de mayo). *Registre d'explotacions ramaderes*. Generalitat de Catalunya.
<https://agricultura.gencat.cat/ca/serveis/registres-oficials/ramaderia-sanitat-animal/registre-explotacions-ramaderes>

Agència Catalana de l'Aigua. (2019). *Filtre de dades*. Generalitat de Catalunya.
<https://aplicaciones.aca.gencat.cat/sdim21/filtre.do>

Idescat. (2022). *Indicadors*. <https://www.idescat.cat/indicadors/>

Geopandas. (s.f.). *User guide: mapping*. https://geopandas.org/en/stable/docs/user_guide/mapping.html

Machine Learning Mastery. (s.f.). *Statistical hypothesis tests in Python cheat sheet*.
<https://machinelearningmastery.com/statistical-hypothesis-tests-in-python-cheat-sheet/>

Appinio. (2024, 6 de mayo). *Análisis de correlación: ¿Qué es? Definición, ejemplos*.
<https://appinio.com/es/blog/investigacion-de-mercados/analisis.correlacion>