Tampere University

Erick Rico Esparza

# COMPARATIVE ANALYSIS OF AQUIFER SUSTAINABILITY IN THE MEXICO VALLEY REGION

**YEB.051 Special Project in Environmental and Energy Engineering**
Supervisor: Chunyan Tang
Intended credits: 8
07.06.2025

# CONTENTS



.

# 1. INTRODUCTION

Water, as a critical natural resource, plays a vital role in the sustainability and development of densely populated regions such as the Mexico Valley. In this area, it is well-known by the population that the aquifers are not only the primary source of water supply for domestic and industrial uses but also a key element in regional planning and environmental management. The continuous pressure from urban expansion and the increasing demand for water have brought to light serious challenges regarding groundwater management, including issues of overexploitation and inefficient allocation[1].

The significance of groundwater in the Mexico Valley is well documented. In urban centers, particularly in Mexico City, the struggle to balance water supply with rising consumption is evident[2]. Infrastructure limitations, coupled with rapid urban growth, have exacerbated water scarcity problems. For example, studies and news articles have highlighted that excessive water extraction has led to declining water tables and potential long-term ecological impacts[2][3]. In addition, recent news coverage has raised concerns about the increasing stress on aquifer systems and the urgent call to “stop the exploitation" of aquifers[4].

The objective of this study is to perform a comparative analysis of the sustainability of the 14 aquifers in the Mexico Valley Region, focusing on the most recent governmental open data from 2023, which provides more complete information to assess the status of aquifers in the Mexico Valley Region[5], using various variables. These include the annual average recharge, natural compromised discharge, the volume of groundwater extraction, conceded/assigned groundwater volume, volume of extraction considered in technical studies, and both positive and negative annual water availability.

By integrating these diverse variables into the analysis, this study aims to provide a comprehensive picture of the current state of groundwater resources and to inform future water management strategies, particularly focusing on evaluating the annual average availability of groundwater in the aquifers of the Mexico Valley Region; determining the extent of aquifer overexploitation and its impact on water availability; examining the discrepancy between the water volume that is legally assigned and the volume that is actually extracted; identifying spatial patterns in both water availability and groundwater extraction across the aquifers; and evaluating whether actual groundwater extraction volumes exceed the natural recharge capacity of the aquifers.

These objectives are supported by the following research questions that will be answered by the end of this study:

1. What is the average annual availability of aquifers in each of the 14 aquifers in Mexico Valley Region, and how many are experiencing water shortages?
2. Is there a direct relationship between overexploitation and the existence of a water deficit, and to what extent does the water allocated correspond to the volume actually extracted?
3. Which areas within the aquifer system face the highest pressure in terms of extraction and resulting deficits; and do the current extraction volumes surpass the natural recharge rates of the aquifers, and what does this indicate about the sustainability of groundwater use?

Groundwater is essential not only for meeting current urban demands but also for ensuring long-term sustainability. The analysis will rely on full datasets provided by a governmental platform[5] and will incorporate a combination of statistical and graphical methods to visualize and interpret the data. This approach will allow for an examination of the sustainability of the aquifer systems, ultimately providing a basis for recommending strategies that balance urban water needs with the preservation of natural resources. So, by focusing on the 2023 data and integrating multiple variables, the study aims to offer clear insights into groundwater sustainability, thus supporting more informed water management decisions in the future.

This report is structured into 5 chapters. While Chapter 1 introduced the topic, research questions, and objectives, Chapter 2 reviews theoretical bases and existing comparative studies, covering definitions, and gaps in research. Chapter 3 explains the methodology, including tools, analytical approach, and study limitations. Chapter 4 presents results and discussion, with evaluations of aquifer availability, extraction effects, spatial patterns, and statistical analysis. Chapter 5 offers conclusions and recommendations, and finally, References and two Appendixes with tables and codes close this report.

# 2. THEORETICAL FOUNDATIONS AND COMPARATIVE ANALYSES IN AQUIFER SUSTAINABILITY STUDIES

This chapter reviews the concepts, theoretical approaches, and comparative studies related to aquifer sustainability. It establishes a conceptual framework with clear definitions, describes the main quantitative and spatial methodologies used in aquifer research, and reviews global and Mexico-focused studies. The chapter concludes by identifying gaps that support the need for an updated comparative analysis for the Mexico Valley.

## 2.1 Conceptual Framework and Definitions

Some concepts are very important for evaluating aquifer sustainability. In this context, an *aquifer* is a naturally occurring, permeable rock or sediment layer that stores and transmits groundwater. *Groundwater*, as part of the hydrological cycle, is *recharged* by *precipitation*, *infiltrates* the *subsurface*, and then it finally *discharges* naturally into streams or wetlands[6]. *Figure 1* presents an overview of the water cycle, where it is possible to visualize these concepts, among others.

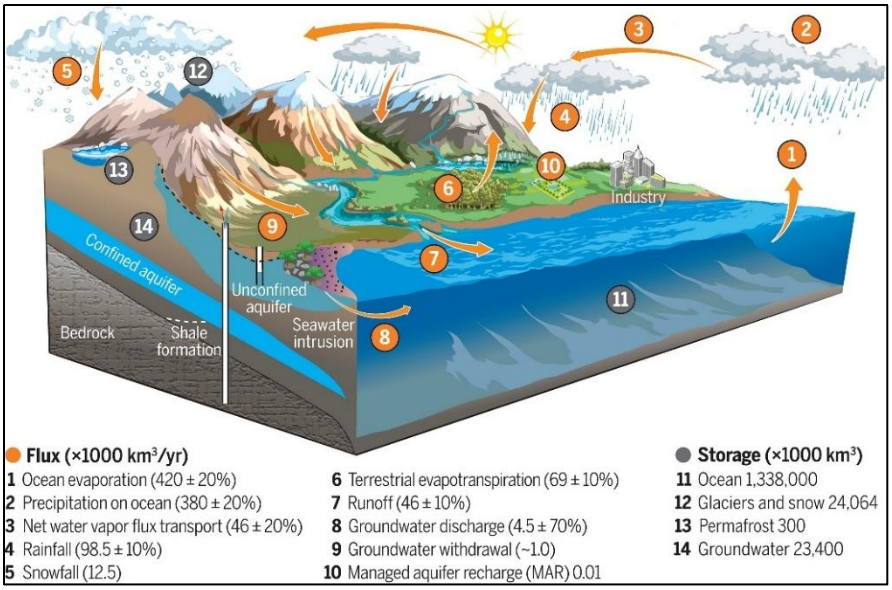


Figure 1. A schematic diagram of the general hydrological cycle with its components, including groundwater discharges and withdrawals[7].

Additionally, there are more basic definitions (*Table 1*) that follow the technical guidelines written by the National Water Commission (CONAGUA—Comisión Nacional del Agua—by its acronym in Spanish)[8][9] and some other specialized literature[10].

Table 1. Key Parameters for Aquifer Management[8][9][10].

| Concept | Abbreviation | Definition | Original Name in Spanish |
|---|---|---|---|
| Average Annual Recharge | R | The volume of water that an aquifer receives, in a specific time interval. | Recarga Total Media Anual |
| Average Annual Availability of Groundwater | DMA | The average annual volume of groundwater available for extraction after subtracting the average annual recharge (R), the natural discharge (DNC) and the volume extracted (VEAS). A positive DMA indicates surplus water, while a negative value reflects a deficit. | Disponibilidad Media Anual de Agua Subterránea |
| Compromised Natural Discharge | DNC | The portion of the aquifer's natural discharge reserved to sustain environmental flow and prevent degradation. | Descarga Natural Comprometida |
| Volume of Groundwater Extraction | VEAS | The total volume of groundwater extracted, which includes registered extractions as well as estimates for informal or unrecorded usage. | Volumen de Extracción de Aguas Subterráneas |
| Concessioned Volume of Groundwater | VCAS | The volume of groundwater that has been legally assigned or granted for use, as recorded by official agencies. | Volumen Concesionado de Aguas Subterráneas |

Moreover, although its reported values are all zero in the dataset, there is another parameter that plays an important role in groundwater assessments, which is introduced in *Table 2*:

Table 2. Additional Parameter for Aquifer Management[5][10].

| Concept | Abbreviation | Definition | Original Name in Spanish |
|---|---|---|---|
| Extraction Volume in Technical Studies | VEXTET | The extraction volume estimated through technical studies, which may include volumes that are actually extracted even if not fully registered. | Volumen de Extracción en Estudios Técnicos |

Understanding these parameters is crucial for comparing aquifer performance, especially when evaluating how much water is legally authorized versus how much is actually being extracted. While VCAS was fully used in this analysis to represent assigned water volumes, VEXTET could not be included due to missing data. Its values were all reported as zero, which prevented a deeper comparison between actual extraction and volumes estimated in technical studies. However, acknowledging the absence of this variable is still important, as it highlights a gap in the available data and points to the need for further technical research. This overall framework sets the stage for the quantitative and spatial evaluations discussed in later sections.

## 2.2 Theoretical and Methodological Approaches in Aquifer Studies

### 2.2.1 Quantitative Assessment of Aquifer Variables

Quantitative assessments of aquifer sustainability rely on long-term monitoring data and statistical methods. Researchers calculate R, DNC, and VEAS to determine the DMA[8], following standards such as *NOM-011-CONAGUA-2015*, which is an official Mexican standard for water conservation, and which in turn establishes the specifications and methods to determine the average annual availability of national waters[11]. Long-term datasets, often spanning decades, are used to compute annual averages that reveal trends and variability in aquifer conditions. This systematic approach helps identify periods of surplus or deficit, informing water management decisions, as seen in different parts of the world and Mexico itself[12][13].

Furthermore, quantitative methods often involve time-series analysis and regression models to correlate aquifer performance with influencing factors such as climate variability, urban growth, and changes in land use[13]. These sorts of methods have allowed researchers to predict future water availability under different scenarios and to assess the impact of anthropogenic pressures (Hernández-Cruz et al., 2022), given that standardized measurement protocols ensure comparability among different aquifer systems and enable policymakers to track progress over time.

Finally, the reliability of these quantitative assessments depends on the quality of monitoring networks. Data from CONAGUA's hydrometric and climatological stations—key features such as groundwater extraction gauges, piezometric-well measurements of water table levels, dedicated recharge monitoring sites, seawater-intrusion sensors, and meteorological records of rainfall and evaporation—provide the backbone for the regional modeling efforts described by Hernández-Cruz et al. (2022). Their work highlights how complete, and standardized datasets not only enable accurate computation of annual water balances but also support evaluation of uncertainties and error margins which is key for decision-making[13]. By ensuring comprehensive spatial coverage and continuous operation of these networks, indicators like DMA can be calculated with confidence and used as actionable guides for sustainable water resource management.

### 2.2.2 Remote Sensing, GIS, and Technical Modeling

Remote sensing and GIS have revolutionized the study of aquifer sustainability by providing spatially explicit data that complement ground-based measurements. Satellite imagery and aerial photographs are used to monitor land cover changes, vegetation indices, and surface water conditions. These spatial data enable researchers to map recharge zones, delineate aquifer boundaries, and assess changes over time[14].

GIS tools allow for the integration of diverse datasets, including rainfall records, topography, and extraction volumes, into comprehensive models. Technical modeling, such as with the SWAT (Soil and Water Assessment Tool) or WRF (Weather Research and Forecasting) models, simulates the impact of climate variability on groundwater recharge and surface runoff[15]. Such models could be critical in regions where extraction is informal or poorly recorded, as they could provide estimates that supplement registered data. These methods enhance spatial resolution and allow for the identification of vulnerable areas within aquifer systems. Moreover, technical models incorporate both historical and current data to generate predictive scenarios, which help evaluate the effects of potential policy changes, climate shifts, and urban expansion on aquifer sustainability (Feng et al., 2024). Hence, the combination of remote sensing, GIS, and hydrological modeling results in a powerful toolkit for assessing and comparing aquifer performance, ultimately supporting informed decision-making at regional and national levels.

## 2.3 Review of Relevant Comparative Studies

### 2.3.1 Global Overview of Aquifer Sustainability Studies

Globally, studies on aquifer sustainability have advanced considerably over the last decades. Integrated water balance models are widely used to evaluate groundwater sustainability across different climatic and socio-economic contexts[16]. Such studies utilize standardized indicators—although called differently—like DMA, DNC, and VEAS to compare aquifer conditions in various regions. These global studies emphasize the role of technological advances in remote sensing and modeling, which have improved the spatial resolution of water resource assessments and allowed for more accurate predictions of aquifer behavior under different environmental pressures[13][14].

Researchers from various regions have applied similar methodologies to examine the impact of climate change, urban expansion, and agricultural practices on groundwater resources, as demonstrated by Palay (2022) and Hernández-Cruz et al. (2022) reveal. These comparative analyses around the world often reveal significant differences in sustainability profiles, which are linked to variations in recharge rates, land use, and extraction practices. To capture the full complexity of aquifer dynamics, the literature emphasizes a multidisciplinary approach that integrates hydrology, remote sensing, and socio-economic analysis[13][15][16]. To bring these global patterns into focus, *Figure 2* presents a world map of aquifer locations, highlighting local and shallow systems, complex hydrological networks, and the major groundwater basins.

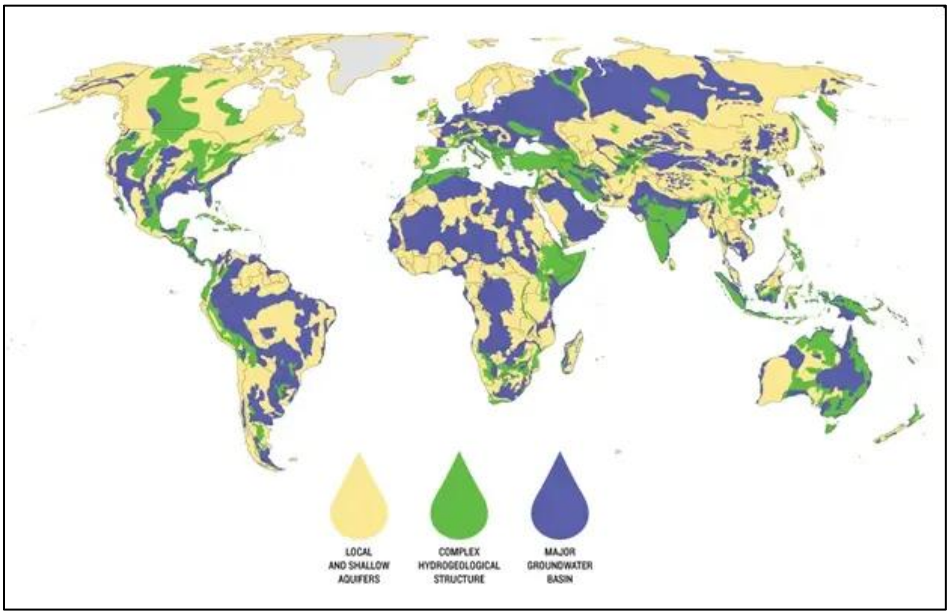


Figure 2. Overview of global aquifer locations: local and shallow aquifers, complex hydrological networks, and principal groundwater basins[16].

Furthermore, these comparative studies highlight challenges common to many countries: overexploitation, reduced recharge, and inadequate governance[13][14][15]. These challenges are often compounded by climate change, making sustainable water management a priority for nations worldwide (Hernández-Cruz et al., 2022; Silva et al., 2022). The methodologies developed and tested in global contexts provide valuable insights and serve as benchmarks for national and regional studies, including those in Mexico. Such international perspectives underscore the importance of continuous monitoring, policy innovation, and technological integration in addressing aquifer sustainability issues[13][17].

Finally, the synthesis of global research provides a solid framework for comparative analyses, offering standardized metrics and methodological approaches that can be adapted to specific regional conditions. As seen, global literature not only informs best practices but also identifies gaps in current knowledge, encouraging further research. These findings serve as a useful reference point for studies—such as this one—focusing on the specific challenges and opportunities present in the Mexico Valley.

### 2.3.2 Overview of Aquifer Sustainability in Mexico

Within Mexico, numerous studies have addressed the sustainability of aquifers, particularly in regions where water stress is most acute (refer to *Figure 3*). CONAGUA's extensive research and annual reports have documented that many aquifers, especially in densely populated and agriculturally intensive areas like the Mexico Valley, face significant overexploitation. These studies often employ the same technical indicators (DMA, DNC, VEAS) to assess water balance and reveal trends over time. The literature shows that some aquifers have been steadily declining due to unsustainable extraction practices, while others remain stable due to effective management measures[8][9][18].

Research in Mexico has also compared aquifer performance between regions, highlighting differences in natural recharge and extraction rates. For example, studies indicate that aquifers in arid regions tend to have lower recharge rates, resulting in higher vulnerability to overexploitation[17]. These regional comparisons are critical for understanding the unique challenges in each area, such as the Mexico Valley's combination of high population density and variable climatic conditions[19].

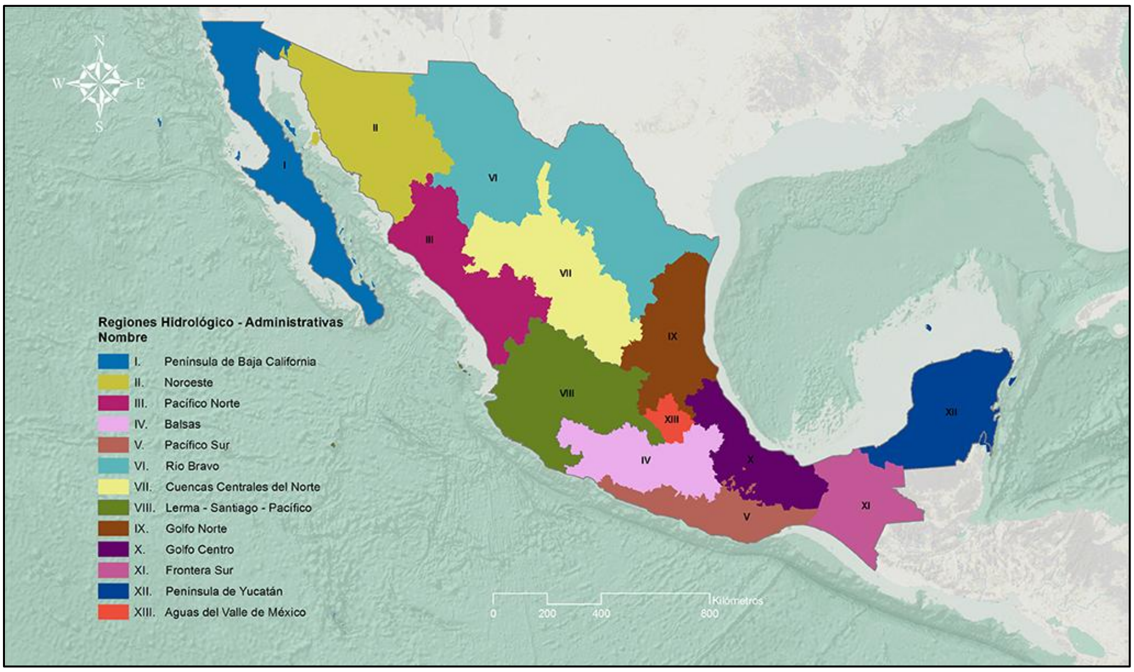


Figure 3. “Regiones Hidrológico-Administrativas de México” – a map of Mexican Hydrological Administrative Regions[20].

Overall, the body of research from national institutions and academic studies provides a strong foundation for evaluating groundwater sustainability in Mexico. Despite the detailed work by CONAGUA[8][9] and other researchers[10][13][17][18][19], the synthesis of these findings indicates that further integrative studies are needed to address current challenges in a comprehensive manner. This regional focus is crucial for developing targeted water management strategies that can effectively address overexploitation and ensure long-term sustainability.

### 2.3.3 Comparative Analyses and Methodological Variations for Mexico

Comparative analyses in Mexico quite often focus on evaluating either a single aquifer across time or and not as frequent on examining multiple aquifer systems using similar quantitative and spatial methodologies[13]. This is shown, for example, by the air quality monitoring systems used in Mexico, which prove that it is possible to build reliable and transparent data frameworks[21]. A similar approach could be applied to groundwater. Researchers, for example, have studied the water balance of the aquifers within the Mexico Valley Region by analyzing and monitoring long-term trends in recharge, natural discharge, and extraction. These studies, performed by Chamizo-Checa et al. (2018), who centers on the *1310 - Valle del Mezquital* aquifer[22], or Neri-Ramírez et al. (2013), who focuses on the *1508 – Cuautitlán-Pachuca* one[23]; or Palacios-Vélez et al. (2016), who compares some of the Mexico Valley Region's aquifers in terms of agriculture sustainability[24], integrate technical data to provide a complete picture of water use, which reveal differences in sustainability among aquifers, influenced by local climate, land use, and population pressures[10]. To give visual perspective to these assessments, *Figure 4* delineates the boundaries and key features of the study area. *Figure 5* then maps the 14 principal aquifers within this region, highlighting their distribution and relative extents.

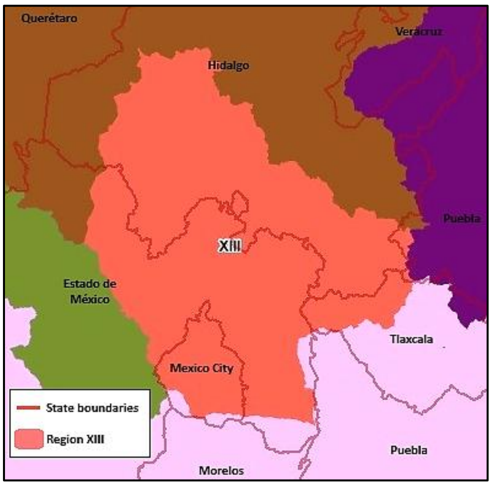


Figure 4. Hydrological Administrative Region XIII: Mexico Valley. Own data adaptation from SEMARNAT[25].

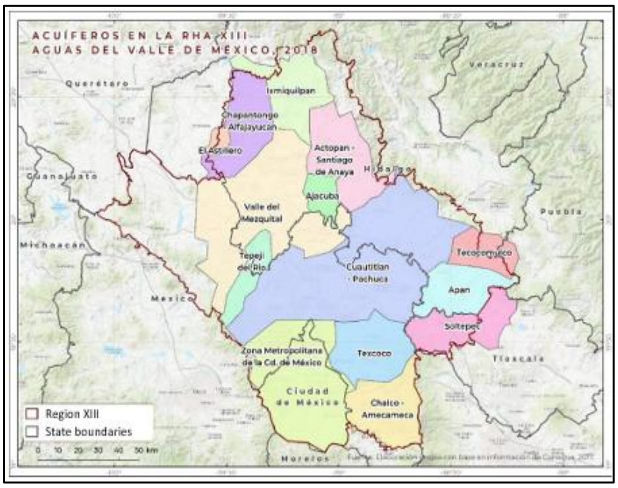


Figure 5. Aquifers (14) in the Hydrological Administrative Region XIII. Own adaptation from CONAGUA[26].

In addition, methodological variations exist among studies due to differences in data availability, monitoring techniques, and modeling approaches[13]. Some studies use high-resolution remote sensing data and advanced technology to provide helpful insights when it comes to water management[14][15], while others rely on historical records provided by CONAGUA[10][22][23][24]. These variations can lead to differing readings of aquifer health, making it essential to standardize measurement techniques for meaningful comparisons. For example, the use of consistent indicators, such as DMA, could allow researchers to compare aquifer sustainability across regions with diverse hydrogeological settings, making use of complete datasets.

Furthermore, comparative studies in Mexico, such as Hernández-Cruz et al. (2022) and Silva et al. (2022) have highlighted the need for integrating socio-economic data—

such as population density or urban expansion—into the assessment of water sustainability[13][17]. Thus, by combining hydrological, technical, and socio-economic analyses, researchers can better understand the drivers of overexploitation. This multifaceted approach provides actionable insights that are crucial for policy development and water management planning in regions like the Mexico Valley one.

### 2.3.4 Gaps and the Need for a Focused Comparative Analysis

Despite extensive research on aquifer sustainability, significant gaps remain in the comparative analysis of the Mexico Valley's aquifers. While numerous studies address individual aquifers or broad regional water balances, there is a lack of integrative research that compares key indicators across the 14 aquifers of this critical region. Such a focused study is essential due to the unique challenges posed by high urban density, rapid population growth, and heterogeneous recharge conditions in the valley.

Furthermore, as seen, current literature often varies in methodological approach, data resolution, and indicator selection, leading to inconsistent assessments of aquifer health. This variability complicates the formulation of comprehensive management strategies and may lead to misinterpretations of sustainability levels. Moreover, while many studies highlight the gap between legally assigned volumes (VCAS) and technical extraction estimates (VEXTET), actual data for the latter parameter is often missing or incomplete, making this comparison difficult to conduct. Acknowledging this limitation, a focused comparative study that uses the available indicators can still provide clearer guidance for decision-makers and highlight the need for improved data collection.

In conclusion, the reviewed literature shows that while extensive research exists on aquifer behavior and water management, a focused, updated comparative analysis of the Mexico Valley aquifers is still needed. An updated and detailed comparative analysis underscores the significance of the current study and ultimately bridges this gap, enabling more accurate identification of at-risk aquifers and informing more effective water management policies. This work—using advanced data processing techniques (including Python-based analyses)—aims to provide a comprehensive assessment of groundwater resources in this strategically important region.

# 3. MATERIALS AND METHODS

This chapter details the methods for analyzing aquifer sustainability in the Mexico Valley Region. It introduces the dataset of 14 aquifers and explains how geographic and hydrological conditions are represented. A data dictionary defines the variables used, while descriptions of the software tools and preprocessing steps illustrate how the data were prepared. The analytical framework is then presented, and potential data gaps are acknowledged to ensure transparency and guide future research.

## 3.1 Dataset Description and Geographic Context

This study relies on a dataset that includes official information about the 14 aquifers located in Mexico Valley Region. As described, these aquifers supply water to a densely populated area, including Mexico City. The dataset is derived from a governmental open-access platform[5] ensuring consistency of the already presented variables.

The region of study covers an area where rapid urbanization and industrial activities have significantly increased the demand for groundwater. The valley's complex topography, as well as its varied geological formations, influence both the recharge capacity of the aquifers and the spatial distribution of extraction points (*Figure 6* illustrates a representative part of the Mexico Valley's topography and hydrographic basin—within the region—showing a one-year hydrological balance that captures the dynamics between precipitation, R, and extraction processes). Hence, the dataset offers a complete overview of each aquifer's location, its administrative limits, and the specific conditions that affect water availability.

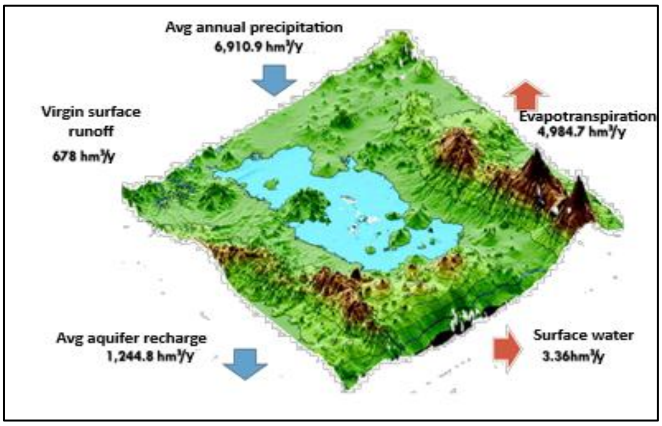


Figure 6. Simplified annual hydrological balance of a segment of the Mexico Valley basin in the 13th Hydrological Administrative Region in 2013[10].

## 3.2 Data Dictionary

The dataset used in this study includes the most relevant variables for assessing aquifer conditions, especially in terms of sustainability and overexploitation status. *Table 3* presents the main fields of the dataset, along with a brief description of each variable. This table is a guide to understanding how the information is structured, what each column means, and how it contributes to the comparative analysis in later sections.

Table 3. Data Dictionary[5][9].

| Column Name | Description | Units | Data Type | Additional Notes |
|---|---|---|---|---|
| clvacuifer | Aquifer key | - | int | Assigned by official sources. |
| acuifero | Aquifer name | - | str | Includes a regional label. |
| clvrha | Hydrological Administrative Region key | - | str | Assigned by official sources. |
| rha | Hydrological Administrative Region | - | int | Mexico has 13 zones, the XIII being "Waters of the Valley of Mexico" in which the Mexico City is located. |
| r | Average Annual Recharge | $hm^3/y$ | float | Represents total recharge into the aquifer, averaged over one year. |
| dnc | Compromised Natural Discharge | $hm^3/y$ | float | Refers to the volume of water naturally discharged that must be maintained for ecological stability. |
| vcas | Concessioned Volume of Groundwater | $hm^3/y$ | float | Officially granted volume of groundwater extraction rights according to local permits. |
| vextet | Extraction Volume in Technical Studies | $hm^3/y$ | int | Represents more refined or updated estimates of actual extraction, based on technical reports. Although in this case all values are zero. |
| veas | Volume of Groundwater Extraction | $hm^3/y$ | float | Actual extraction recorded or estimated from different water users in the region. |
| dmapositiva | Positive Average Annual Availability of Groundwater | $hm^3/y$ | float | Computed from recharge, discharge, and extraction data; indicates surplus (positive) or deficit (negative). |
| dmanegativa | Negative Average Annual Availability of Groundwater | $hm^3/y$ | float | |
| situacion | Aquifer situation | - | str | Indicates whether the aquifer is classified as with availability or not. |
| anio | Year of the underground arrangement layer | - | int | This dataset corresponds to 2023. |
| condicion | Aquifer condition | - | str | Indicates whether the aquifer is classified as overexploited, at risk, or in equilibrium. Derived from the comparison between extraction volumes and recharge. An aquifer is considered overexploited if the extraction/recharge ratio is greater than 1.1, and not overexploited if said ratio is less than 1.1. |

## 3.3 Software and Tools

For this analysis, Python 3 running on Google Collaboratory was used[27], where only three libraries were used for data management and visualization: *pandas* abbreviated as *pd*, *matplotlib.pyplot* as *plt*, and *seaborn* as *sns*, which were useful to effectively handle this dataset. *Figure B1* in *Appendix B* illustrates a screenshot that shows the way the libraries were introduced into the Google Colab notebook.

## 3.4 Data Import, Exploration, and Preprocessing

Now, in this section a description of the process of importing and exploring the dataset is presented. First, it is important to clarify that the data file obtained from the Mexican government open data site, originally named "*Disponibilidad Acuíferos_2023*"[5], included information on all aquifers and their corresponding Hydrological Administrative Region across Mexico. Therefore, the file was manually adjusted in Excel to remove columns and rows that were not needed to perform this specific analysis. After preprocessing, the dataset was filtered to retain only the essential information—specifically, the 14 rows representing each aquifer located within the limits of Mexico Valley Region. Next, the new file, "*Disponibilidad Acuíferos_2023 – Valle de México*", was uploaded to Google Colab using the file uploader feature. Once the file was in the notebook, it was imported into a pandas *DataFrame* using the *pd.read_csv()* function (*Figure B2* in *Appendix B* shows the code used to perform this activity—*df* denotes the *DataFrame* that holds the complete dataset, including the 14 aquifers within the region).

Now, after the abovementioned process, an execution of the command *df.head()* to display the first few rows (see *Figure B3* in *Appendix B*) was implemented. This step was made to confirm that the data had been loaded correctly into the environment and that all columns—like the ones outlined in the Data Dictionary (Section 3.2)—appeared as supposed to. Additionally, to ensure the integrity of the dataset, a checking for missing values was made with the command *df.isnull()*—see *Figure B4* in *Appendix B*. Ultimately, the output obtained verified that the dataset was as expected, with all cells returning "False," which indicated that there were not ‘null’ values.

Finally, after importing and verifying the dataset, the next step was to prepare the data for analysis by addressing language and encoding issues. Specifically, a translation of the Spanish text in the *'condicion'* and *'situacion'* columns into English was performed to improve visualization and comprehension. Additionally, accented characters in the aquifer names were removed to prevent potential encoding errors. This processing resulted in the creation of new columns: *'aquifer'* for the normalized aquifer names, *'condition'* for the translated condition values, and *'situation'* for the translated situation values. *Figure B5* in *Appendix B* shows the Python code used for this.

## 3.5 Analytical Approach

This study applied a structured analytical approach to assess groundwater conditions in the 14 aquifers of Mexico Valley Region. The analysis is divided into several steps that include descriptive statistics, visual exploration, and statistical testing to better understand the dynamics between groundwater availability, extraction, recharge, and overexploitation.

First, to begin the analysis, descriptive statistics such as mean, standard deviation, and average differences were used to describe and summarize the key variables. These included R, DNC, VEAS, and groundwater availability (calculated with R, DNC, and VEAS). These values were used to identify how much water is naturally available in each aquifer and how much is being removed. This step directly supported *research question 1*, which aims to determine the current availability of water in the aquifers and identify cases of water shortage (deficit). Then, average values for each group (overexploited vs. subexploited) were also calculated to understand variations and begin forming hypotheses about the impact of overexploitation, contributing to *research question 2*. Besides, the analysis included a comparison between the officially VCAS and the actual extracted volumes VEAS. By calculating the difference for each aquifer, the study assessed whether extraction aligns with legal allocations. This analysis directly addressed also *research question 2*, revealing that several aquifers, particularly Cuautitlán-Pachuca, exceed their authorized extraction limits, which points to possible gaps in enforcement or monitoring.

Second, to visualize the patterns and trends, different plots were generated throughout the analysis: bar charts to compare recharge, extraction, and committed discharge volumes across aquifers, boxplots to explore distribution differences between overexploited and subexploited aquifers, and scatter plots to examine the spatial relationship between extraction volume and groundwater availability. These visual tools allowed to see which aquifers are under the most pressure, supported the classification of aquifers by condition, and revealed patterns between variables like VEAS and R. For example, in *Section 4.1.5*, instead of analyzing VEAS against VEXT, which was unavailable in the dataset, the analysis focused on comparing VEAS to R. This ratio (VEAS/R) offered a clear indicator of groundwater sustainability and addressed *research question 3*, helping identify which aquifers are extracting beyond their natural limits.

Additionally, although the original dataset did not include geographic coordinates or shapefiles, external interpretation was made by referencing geographic designations in aquifer names. Therefore, tools like *geopandas* or *folium* could have been useful if other specific analyses were needed, but in this case, these libraries were not required. However, governmental online GIS tools were used to help with the understanding of the results, deepen the discussion, and thus come up with conclusions. Hence, by understanding spatial trends, it was possible to focus again on *research question 3* and comprehend which areas face the most pressure.

Finally, to support the visual and descriptive insights, a t-test was performed in the final section (*4.1.6*) of the analysis. This statistical method was used to compare the DMA between overexploited and subexploited aquifers. Although the result was not statistically significant at the standard level ($p \approx 0.084$), the trend was still clear: overexploited aquifers tend to have much lower availability, which supports the idea that excessive extraction directly affects the water balance. The boxplot created alongside the t-test further illustrated this difference. This part of the analysis addressed *research question 2*, which seeks to confirm the relationship between overexploitation and water deficit. While no regression or correlation models were used in this particular study due to the small sample size ($n=14$), future research with more data could apply linear regression or Pearson correlation to further quantify relationships between VEAS, R, and DMA.

This analytical approach was designed to provide a clear, data-driven foundation for evaluating the state of groundwater in Mexico Valley Region. Each method—from basic statistics to hypothesis testing—was selected to match the structure of the dataset and the scope of the research questions. Together, these tools allowed for meaningful insights, even in the absence of complete spatial data or time series, and offered a strong basis for water management recommendations.

## 3.6 Study Limitations

One of the main limitations of this study is the lack of VEXTET. Although this variable was included in the dataset, all values were reported as zero, which made it impossible to assess whether actual groundwater extraction aligns with the volumes estimated in official technical studies. This type of information is important for understanding how planning supports water management, and analyses like VEAS vs. VEXTET could have been performed. Future studies would benefit from updated and complete VEXTET data to improve the accuracy of resource planning assessments. Without this, comparisons between planned and actual extraction volumes remain incomplete, reducing the capacity to identify possible mismatches or inefficiencies in water allocation.

Another limitation is related to the time period of the data. The dataset used in this study corresponds to groundwater availability as reported in 2023. While this information is relatively recent, nearly two years have passed since its publication. Therefore, current groundwater conditions may have slightly changed due to new extraction rates, climate variations, or policy updates. Thus, this analysis provides only a strong snapshot of the situation in 2023, but it cannot reflect the most recent developments. So, to maintain relevance, this kind of analysis should ideally be repeated each year or combined with near real-time monitoring data. Incorporating more dynamic sources, such as satellite-based monitoring or monthly extractions from local agencies, could help enhance temporal precision and policy responsiveness.

A final limitation is the restricted availability of complete historical data. Although data from years such as 2020 and 2018 were reviewed, they showed little variation in terms of aquifer conditions or water availability. For example, aquifer 0901 (*Zona Metropolitana de la Cd. de México—Metropolitan Zone of Mexico City*), which lies directly beneath Mexico City, showed only minor changes over recent years. VEAS was 993.23 $hm^3$/year in 2023 compared to 1 020.03 $hm^3$/year in 2020, and DMA was –480.43 $hm^3$/year versus –507.23 $hm^3$/year. This small variation is consistent across all fourteen aquifers in the Mexico Valley Region, and even the 2018 data show similarly slight differences. Moreover, older datasets from 2015 or 2013 displayed more noticeable differences, particularly in the values of DMA—for instance, there is a deficit of –

591.18 $hm^3$/year for aquifer 0901 in 2015—however, they also contained missing cells and inconsistencies that made them unreliable for comparison.

Because of these restrictions, it was not possible to create a continuous trend graph over time, for instance. If more complete time-series data had been available, a plot of these values could have illustrated long-term trends more clearly. So, as a result, the study focused entirely on the 2023 dataset. While this decision allowed for a more complete analysis, it limited the ability to explore long-term trends or perform time series forecasting. Ideally, future research could aim to reconstruct more consistent historical datasets, which would enable broader temporal analyses and support the detection of gradual changes in water availability across Mexico Valley Region.

In conclusion, although this study faced several challenges—including the absence of technical study extraction data, limited historical records, and reliance on a dataset that may not reflect the most current conditions—it still provides meaningful insights into groundwater management in the region. These limitations emphasize the need for continuous improvement in data collection and the integration of more dynamic monitoring methods in future research. Furthermore, addressing these issues could lead to more comprehensive evaluations of extraction efficiency and resource sustainability. In other words, while the limitations of this study caution against generalizing the findings, they also offer valuable opportunities for future research to refine methodologies and thus enhance the reliability of groundwater assessments.

# 4. RESULTS AND DISCUSSION

This chapter presents the main findings of the groundwater sustainability analysis in the Mexico Valley Region. It begins by assessing the availability and extraction volumes of the 14 aquifers. Similarly, it explores the extent of overexploitation and compares water levels. The analysis also includes a comparison between assigned and extracted volumes, and a statistical test to highlight differences in aquifer conditions, offering insights for sustainable groundwater management. The discussion connects these findings to prior research and highlights important gaps.

## 4.1 Data Analysis and Results

This part shows the results from the file used, and connecting each subsection to one of the five objectives defined in *Chapter 1* of this study, alongside a sixth subsection related to a statistical test in terms of overall DMA.

### 4.1.1 Evaluating the Annual Average Availability of Groundwater in the Aquifers of Region XIII

This section explores the groundwater balance in Mexico Valley Region by analyzing the relationship between the R, the VEAS, and the DNC. Using this relationship, the DMA was calculated for each aquifer to assess their sustainability. A graphical representation was generated to visually compare these three components across all aquifers.

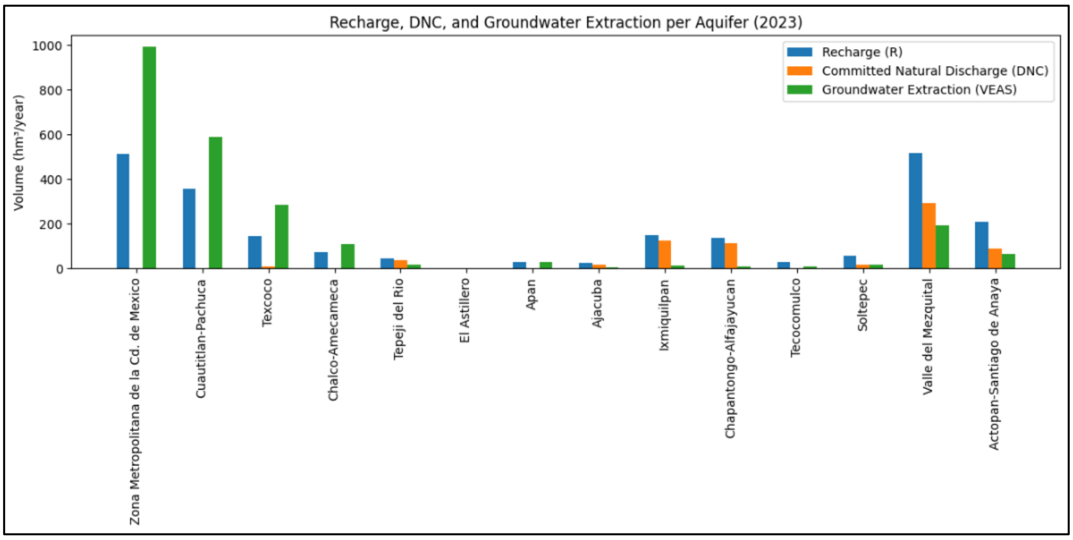


Figure 7. Recharge, DNC, and Groundwater Extraction per Aquifer in 2023.

The results in *Figure 7* and *Table A1* in *Appendix A* clearly show that many aquifers in the region are facing a negative water balance, with extraction volumes surpassing their recharge capacity. The most critical case is the *Metropolitan Zone of Mexico City* aquifer, which exhibits a deficit of over 480 hm³/year. This indicates an alarming level of groundwater overuse that could lead to long-term water insecurity. On the other hand, aquifers like *Ixmiquilpan* and *Ajacuba* have kept a positive balance, showing that sustainable groundwater management is possible within the region. The variation among aquifers focuses the need for local water policies that consider the unique recharge and usage dynamics of each aquifer.

## 4.1.2 Determining the Extent of Aquifer Overexploitation and Its Impact on Water Availability

This section focuses on classifying aquifers according to their level of exploitation and understanding its relationship with water availability. The "condition" variable separates aquifers as either overexploited or “*subexploited”* (unexploited) and is used alongside “negative DMA” data to assess the seriousness of overuse.

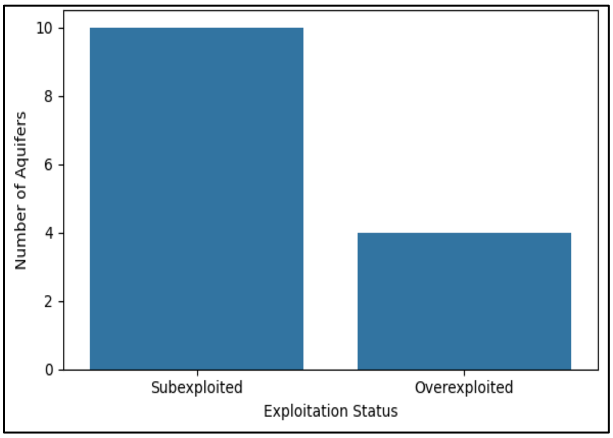


Figure 8. Number of Aquifers by Exploitation Status in 2023.

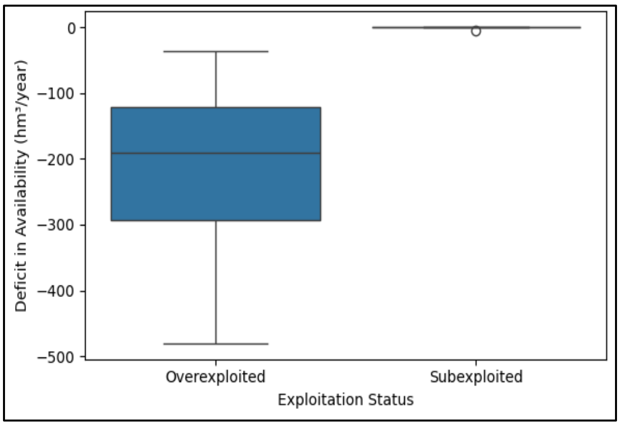


Figure 9. Deficit (DMA-) by Aquifer Condition in 2023.

The data in both *Figure 8* and *9*, and in *Table A2* in *Appendix A* show a clear distinction between the two groups of aquifers. Overexploited aquifers experience significantly higher deficits, averaging over -224 hm³/year, whereas subexploited aquifers mostly remain in positive balance or exhibit minimal deficits. The boxplot clearly reveals a greater variability and severity of deficit in the overexploited group. Despite being a minority in number, overexploited aquifers represent the most critical areas in terms of sustainability. These findings emphasize the need to treat these aquifers as high-priority zones for intervention, not only to reduce current extraction but also to ensure the long-term security of groundwater in the region.

### 4.1.3 Assigned vs. Extracted Water: A Discrepancy Analysis

This section investigates whether aquifers are extracting more groundwater than they are legally permitted to. By comparing the VCAS with the VEAS, it is possible to detect inconsistencies that may signal regulatory gaps or overexploitation beyond legal limits. These discrepancies provide critical insight into the management and governance of groundwater rights in Mexico Valley Region.

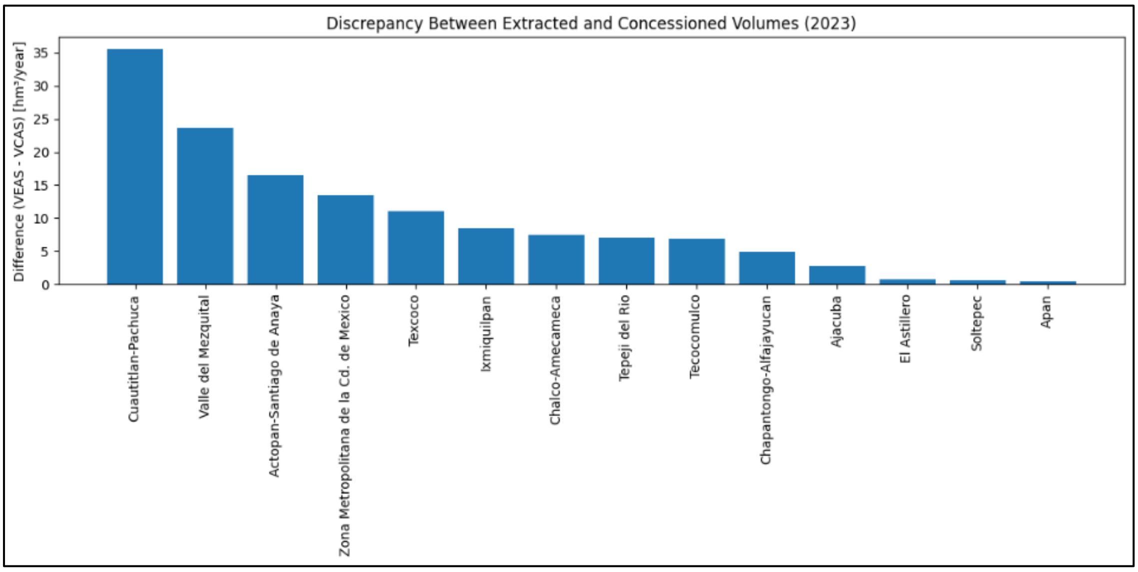


Figure 10. Discrepancy Between Extracted and Concessioned Volumes in 2023.

The results in *Figure 10* and *Table A3* in *Appendix A* reveal that several aquifers are extracting more water than what has been legally assigned to them. Notably, *Cuautitlán-Pachuca* one leads this trend with a discrepancy of over 35 hm³/year, followed by *Valle* *del Mezquital* and the *Metropolitan Zone* ones. This suggests that, despite formal regulations, enforcement may be lacking or actual use may not be properly reported. While some aquifers operate within or below their assigned limits, the consistent pattern of over-extraction in major aquifers highlights the need for closer scrutiny. Bridging the gap between legal allocation and real-world usage is vital for ensuring sustainable groundwater management.

### 4.1.4 Spatial Patterns in Both Water Availability and Groundwater Extraction

This section explores spatial patterns among the aquifers in the Mexico Valley Region by comparing groundwater extraction rates with their calculated availability. A scatter plot was created to identify whether aquifers with high extraction are more likely to show negative water balances. This analysis helps reveal areas under the most pressure and provides clues about the spatial distribution of overexploitation.

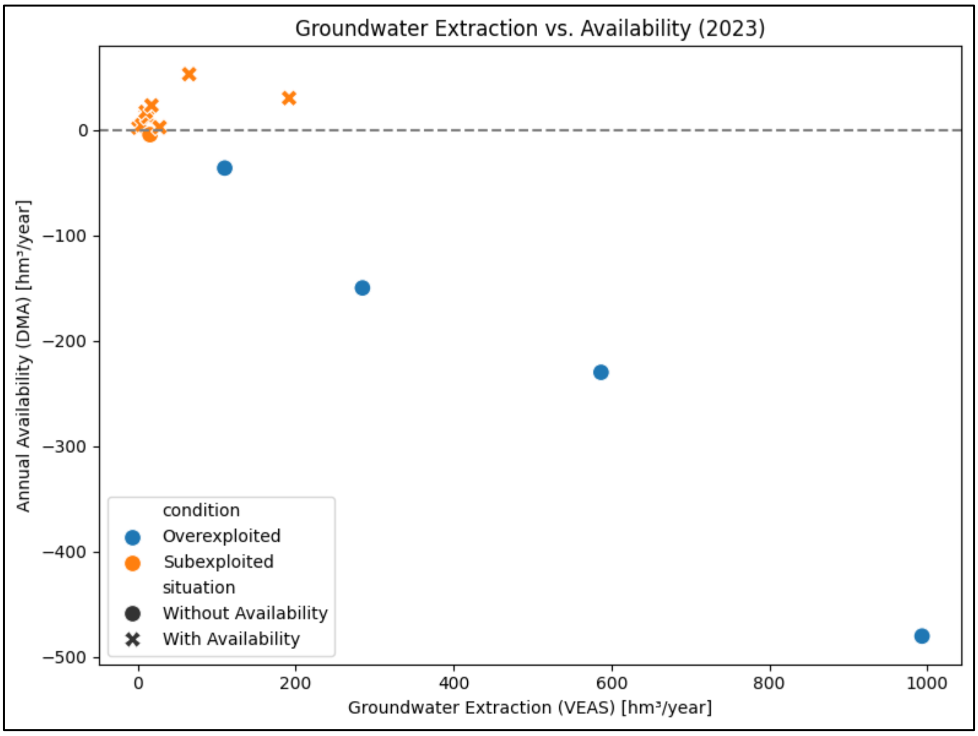


Figure 11. Groundwater Extraction vs. Availability in 2023.

The scatter plot above (*Figure 11*) and *Table 4* below illustrate a concerning trend: aquifers with the highest extraction rates, such as the *Metropolitan Zone* and *Cuautitlán-*

*Pachuca* ones, have the most negative water availability. Interestingly, some aquifers labeled as subexploited, like *Tepeji del Río* one, also show a negative balance, suggesting that labels may not always reflect real conditions. This mismatch implies a need to update classification methods or improve monitoring accuracy. In contrast, aquifers with moderate extraction typically maintain a more sustainable profile. These insights suggest that focusing mitigation efforts on specific high-pressure aquifers could significantly improve regional water security.

Table 4. Spatial Patterns in Aquifer Status.

| aquifer | veas | dma_calculated | condition | situation |
|---|---|---|---|---|
| Zona Metropolitana de la Cd. de Mexico | 993.22991 | -480.42991 | Overexploited | Without Availability |
| Cuautitlan-Pachuca | 586.64264 | -229.94264 | Overexploited | Without Availability |
| Texcoco | 284.50512 | -149.80512 | Overexploited | Without Availability |
| Chalco-Amecameca | 109.99354 | -35.993539 | Overexploited | Without Availability |
| Tepeji del Rio | 15.354466 | -4.254466 | Subexploited | Without Availability |
| El Astillero | 0.724197 | 2.375803 | Subexploited | With Availability |
| Apan | 27.543609 | 2.756391 | Subexploited | With Availability |
| Ajacuba | 5.187226 | 4.912774 | Subexploited | With Availability |
| Ixmiquilpan | 12.79055 | 12.70945 | Subexploited | With Availability |
| Chapantongo-Alfajayucan | 10.346407 | 13.653593 | Subexploited | With Availability |
| Tecocomulco | 9.955644 | 17.344356 | Subexploited | With Availability |
| Soltepec | 17.273231 | 23.526769 | Subexploited | With Availability |
| Valle del Mezquital | 191.53726 | 30.462743 | Subexploited | With Availability |
| Actopan-Santiago de Anaya | 64.955907 | 53.144093 | Subexploited | With Availability |

### 4.1.5 Assessing the Relationship Between Groundwater Extraction and Recharge Capacity

This section explores the sustainability of groundwater use by comparing the actual VEAS with the aquifers' R. The ratio of VEAS to R provides a clear indicator of whether groundwater usage is within ecological limits. A ratio above 1 implies that an aquifer is being depleted faster than it can naturally replenish, signaling unsustainable practices.

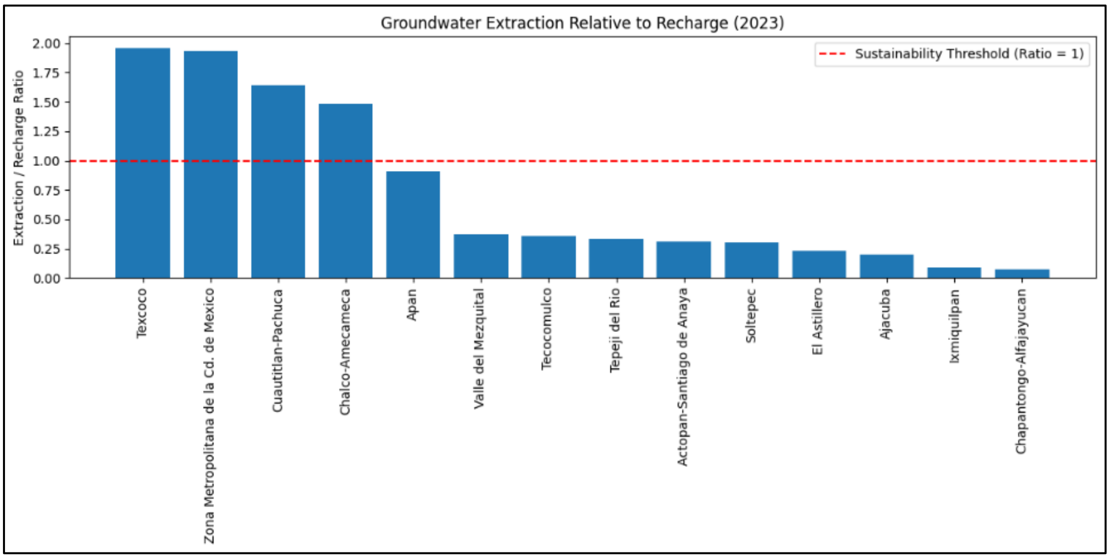


Figure 12. Groundwater Extraction Relative to Recharge in 2023.

The analysis reveals that a majority of aquifers in Mexico Valley Region operate above the sustainability threshold. From *Figure 12* and *Table A4* in *Appendix A*, it is possible to see some interesting facts: *Texcoco* and the *Metropolitan Zone* aquifers extract nearly twice their natural recharge volumes, which raises major concerns about long-term water availability. These high ratios are also consistent with their classification as overexploited and lacking availability. Aquifers closer to or below a 1.0 ratio tend to have more sustainable conditions. This ratio offers a straightforward and powerful metric for water management planning. Integrating it into policy and monitoring frameworks could help prioritize critical areas and ensure that future extraction aligns with ecological boundaries.

### 4.1.6 Statistical Analysis of Groundwater Availability Differences Between Aquifer Conditions

To assess whether groundwater availability (DMA) significantly differs between overexploited and subexploited aquifers, a two-sample t-test was performed. This statistical test compares the mean DMA values of the two groups to identify whether any observed differences are likely due to chance or represent a real underlying pattern. To support this analysis visually, a boxplot (*Figure 13*) was also created to illustrate the distribution of DMA values across both conditions, providing an intuitive comparison of their central tendencies and variability.

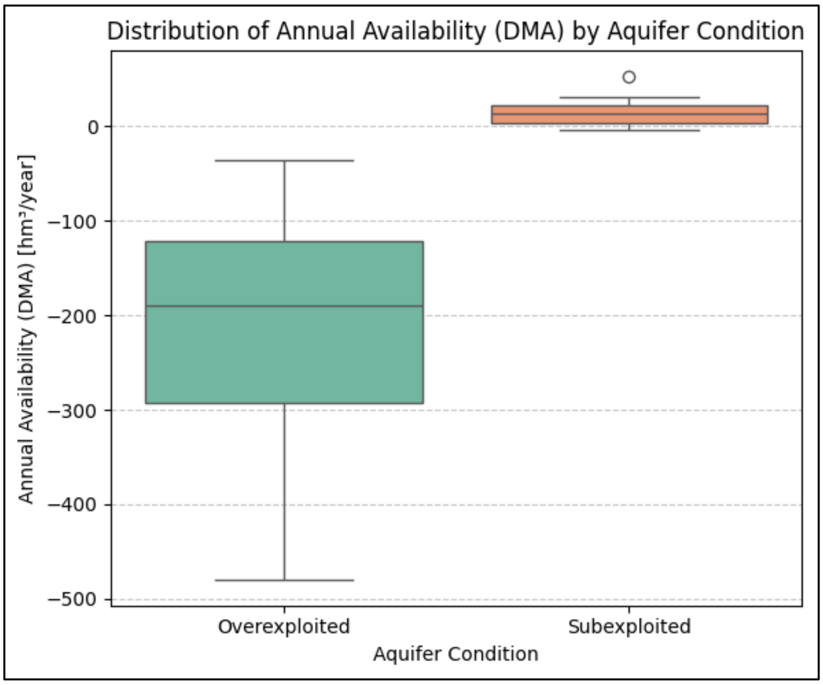


Figure 13. Distribution of Annual Availability (DMA) by Aquifer Condition.

Similarly to the results in *Section 4.1.2*, this boxplot and *Tables A5* and *A6* in *Appendix A* reveal a clear visual distinction between overexploited and subexploited aquifers: the overexploited group tends to have consistently negative DMA values, while subexploited aquifers show a wider range and generally positive balances. This visual trend is supported by the t-test results, which suggest that although overexploited aquifers have a more negative average DMA than subexploited ones, the difference does not reach conventional statistical significance ($p \approx 0.084$). However, considering the small sample size of only 14 aquifers, this result still points to a meaningful pattern that could become significant with a larger dataset, such as one including more years or different regions. Together, the statistical and visual analyses strengthen the broader conclusion that overexploitation is linked to reduced groundwater availability and reinforce the importance of continued monitoring and modeling to support effective water resource management.

## 4.2 Discussion

The five objectives and three research questions outlined this study, and the results speak by themselves. DMA (Objective 1, RQ 1) is positive in nine of the fourteen aquifers, but the remaining five show negative values, with *Mexico City* aquifer alone posting –480 hm³/yr. These figures support broader concerns that climate variability is reducing already limited water supplies. Recent Mexican studies also explain that higher temperatures and longer dry seasons are increasing the pressure on groundwater reserves, especially in arid central basins[17]. Overexploitation, of course, magnifies those shortages. Although the t-test returned $p \approx 0.08$, the median DMA of overexploited aquifers is clearly lower than that of subexploited ones, reflecting studies that link excessive groundwater use to weak coordination between authorities and low public participation in metropolitan areas[6].

Discrepancies between the volume legally assigned and the volume actually extracted (Objective 3, RQ 2) emerge as a structural weakness of the region's water system. *Cuautitlán-Pachuca* and two other aquifers exceed their concessions by up to 36 hm³/yr, a finding that reflects national reviews, which explain that missing or inaccurate field data is one of the main reasons why it is hard to connect policy and practice effectively[13]. Also, the inability to use VEXTET (all zeros) illustrates those data gaps: without reliable field monitoring, regulators cannot correct illegal or under-reported pumping, for example. Additionally, spatial analysis (Objective 4, RQ 3) reinforces this point. High VEAS values group in urban–industrial corridors where fewer monitoring stations are now operating. This is the same issue mentioned by Perevochtchikova (2009) in Mexico City's environmental monitoring report, which shows that many stations have been closed since the 1980s decade, reducing data coverage and weakening trust in the system[21].

Maybe the most severe warning arises from the extraction-to-recharge ratio (Objective 5, RQ 3). Four aquifers already pump more than they naturally gain in an average year, going against sustainability principles. Future climate studies estimate that groundwater recharge could drop by about nine percent by the middle of the century, suggesting that systems which are now close to balance could move into permanent deficit if groundwater extraction is not controlled[17]. Together, these patterns show that a seemingly positive regional balance is hiding serious local problems, weak institutions, and increasing climate risks. The patterns also validate concerns found in the literature, related to the fact that data shortage and weak stakeholder interest hold back adaptive management; hence, when hydrological data is limited and spread out, even advanced tools cannot support good decision-making, as pointed out by Hernández-Cruz et al. (2022).

Another important point comes from looking at governance and monitoring problems through the lens of climate change. Research in this area explains that reducing the gap between authorized and actual extraction depends on better data, more involvement from citizens, and stronger institutions that work at the basin level[6]. The results back that feeling: where monitoring is the poorest and institutional coordination weak, discrepancies—and deficits—are the largest. Conversely, national progress in air-quality monitoring, for instance, shows that well-connected monitoring networks can be created when conditions are favorable[21]. This insight reinforces the urgent need to coordinate policies with scientific observation, focusing efforts on both climate impact assessments and monitoring evaluations to ensure better feedback between data producers, regulators, and end users.

In sum, the analysis shows that the Valley's groundwater system faces problems like uneven availability, concentrated overuse, weak legal alignment, and limited monitoring. The progress made in air quality tracking shows that Mexico can develop strong and transparent data systems when political will, legal support, and funding come together. Reaching the same level of progress for groundwater is less about technical limits and more about the need for better coordination and leadership.

# 5. CONCLUSIONS AND RECOMMENDATIONS

This study assessed groundwater sustainability in the Mexico Valley Region using 2023 aquifer-level data. By analyzing recharge, discharge, extraction, availability, and legal allocation across 14 aquifers, it addressed all five research questions and met the study's objectives. The findings highlight the need for better data monitoring, policy alignment with real extraction patterns, and sustainable water management.

First, while nine aquifers show a positive average annual availability (DMA), five—including the Metropolitan Zone—operate with clear deficits. This confirms that regional averages can mask local shortages. Overexploited aquifers consistently show more negative DMA values, a trend supported by visual and statistical analysis, though the small sample size limited statistical significance. Second, comparing legal allocations (VCAS) with actual extractions (VEAS) revealed major mismatches. Aquifers like Cuautitlán-Pachuca and Valle del Mezquital exceed their legal limits, pointing to weak enforcement and limited monitoring. The absence of data for technical extraction estimates (VEXTET) further limits planning and oversight. Spatial analysis showed that aquifers under the most pressure are in urban and industrial areas, such as the Metropolitan Zone and Texcoco. Notably, some aquifers labeled as 'subexploited', like Tepeji del Río, still show negative balances, suggesting a need to revise classification methods.

Finally, the study confirms that sustainability challenges are not only hydrological but also institutional. Data gaps, weak monitoring, and fragmented governance—especially in high-deficit areas—undermine effective water management. These issues mirror national concerns about declining monitoring capacity and delayed reporting. Despite these challenges, progress is possible. Improving data collection, aligning legal concessions with actual conditions, investing in monitoring, and promoting efficient use can lead to more sustainable outcomes. Mexico's success with air-quality monitoring shows that integrated, transparent systems are achievable. Applying similar strategies to groundwater could help secure water for future generations in the Mexico Valley and beyond.

# REFERENCES


[1] García C., B. (2022). Indicadores de agua limpia y saneamiento en la Ciudad de México y el reto ante un desarrollo sostenible de los recursos hídricos. *Revista Internacional de Estadística y Geografía, 13*(3), 108-121. Available at: https://rde.inegi.org.mx/index.php/2022/10/10/indicadores-de-agua-limpia-y-saneamiento-en-la-ciudad-de-mexico-y-el-reto-ante-un-desarrollo-sostenible-de-los-recursos-hidricos/

[2] Guerrero, T., Rives C. Rodríguez A., & Saldívar Y. (2009). El agua en la Ciudad de México. *Ciencias* 94, April-June, 16-23. Available at: https://www.revistacienciasunam.com/es/43-revistas/revista-ciencias-94/203-el-agua-en-la-ciudad-de-mexico.html

[3] Vilella, P. (2024, April 26). Sequía, sobreexplotación de los acuíferos y problemas de gestión: el Valle de México vive al borde del colapso hídrico. *El Diario.* Available at: https://www.eldiario.es/internacional/sequia-sobreexplotacion-acuiferos-problemas-gestion-valle-mexico-vive-borde-colapso-hidrico_1_11283058.html

[4] López, G. (2024, October 12). Llaman a frenar la explotación de acuíferos en el Valle de México. *El Sol de México.* Available at: https://oem.com.mx/elsoldemexico/metropoli/llaman-a-frenar-la-explotacion-de-acuiferos-en-el-valle-de-mexico-13112601

[5] CONAGUA-Sistema Nacional de Información del Agua (n.d.). *Acuíferos.* Situación de los Recursos Hídricos. Retrieved from: https://sinav30.conagua.gob.mx:8080/SINA/?opcion=acuiferos

[6] Torres B., L. (2017, February). La gestión del agua potable en la Ciudad de México. Los retos hídricos de la CDMX: Gobernanza y sustentabilidad. *Instituto Nacional de Administración Pública.* Available at: http://aldf.gob.mx/archivo-027a57875ea54db65fb86646226b9611.pdf

[7] Xingxing K., Liu J., Scanlon B. R., Jiao J. J., Jasechko S., Lancia M., Biskaborn B. K., Wada Y., Li H., Zeng Z., Yao Y., Gleeson T., Nicot J. P., Luo X., Zou Y. & Zheng C. (2024). The changing nature of groundwater in the global water cycle. *Science*, *383*(6686). https://doi.org/10.1126/science.adf0630

[8] CONAGUA / Disponibilidad por Acuíferos. (n.d.). Available at: https://sigagis.conagua.gob.mx/gas1/sections/Disponibilidad_Acuiferos.html

[9] Comisión Nacional del Agua. (2024, March). *Estadísticas del Agua en México 2023.* Available at: https://sinav30.conagua.gob.mx:8080/Descargas/pdf/EAM2023_f.pdf

[10] Lanuza G., A. (2017, November). *Análisis del impacto de la industria de bebidas no alcohólicas en los recursos hídricos en la Cuenca del Valle de México*. Available at: https://zaloamati.azc.uam.mx/server/api/core/bitstreams/34a97ac5-2ca8-4d9d-b3b3-08024b82b097/content

[11] Gobierno de México. (2015). NORMA Oficial Mexicana NOM-011-CONAGUA-2015, Conservación del recurso agua-Que establece las especificaciones y el método para determinar la disponibilidad media anual de las aguas nacionales. *Diario Oficial de la Federación. Available at:* https://www.dof.gob.mx/nota_detalle.php?codigo=5387027&fecha=27/03/2015#gsc.tab=0

[12] Stoll, S., Franssen, H. J. H., Barthel, R., & Kinzelbach, W. (2011). What can we learn from long-term groundwater data to improve climate change impact studies? *Hydrology and Earth System Sciences*, *15*(12), 3861–3875. https://doi.org/10.5194/hess-15-3861-2011

[13] Hernández-Cruz A., Sandoval-Solís S., & Mendoza-Espinosa L. G. (2022, October). An overview of modeling efforts of water resources in Mexico: Challenges and opportunities. *Environmental Science & Policy, 136,* 510-519. https://doi.org/10.1016/j.envsci.2022.07.005

[14] Mazari-Hiriart M., Tapia-Palacios M. A., Zarco-Arista A. E. & Espinosa-García A. C. (2019). Challenges and Opportunities on Urban Water Quality in Mexico City. *Frontiers in Environmental Science,* 7(169). https://doi.org/10.3389/fenvs.2019.00169

[15] Feng Z. K., Zhang J. S. & Niu W. J. (2024). A state-of-the-art review of long short-term memory models with applications in hydrology and water resources. *Applied Soft Computing,* 167(B), 112352. https://doi.org/10.1016/j.asoc.2024.112352

[16] Palay, I. (2022, March 21). *How groundwater can help us build a more water-secure world - IWMI Blog*. International Water Management Institute (IWMI). Available at: https://www.iwmi.org/news/how-groundwater-can-help-us-build-a-more-water-secure-world/

[17] Silva, J. A., Monroy Becerril, D. M. & Martínez Díaz, E. (2023). Effects of climate change on water resources in Mexico. *Management of Environmental Quality*, *34*(2), 408-427. https://doi.org/10.1108/MEQ-03-2022-0081

[18] Ávila G., P. (2008). Vulnerabilidad socioambiental, seguridad hídrica y escenarios de crisis por el agua en México. *Ciencias*, *90*, April- June, 46-57. Available at: https://www.redalyc.org/pdf/644/64411395008.pdf

[19] Pérez, I. (2020, April 27). Acuíferos, nuestra valiosa fuente de agua. *Ciencia UNAM*. Available at: https://ciencia.unam.mx/leer/988/acuiferos-una-valiosa-fuente-de-agua-

[20] *SEMARNAT - Atlas / Agua*. (n.d.). https://gisviewer.semarnat.gob.mx/aplicaciones/atlas2015/agua_RHA.html

[21] Perevochtchikova, M. (2009). La situación actual del sistema de monitoreo ambiental en la Zona Metropolitana de la Ciudad de México. *Estudios Demográficos y Urbanos*, *24*(3), 513-547. https://doi.org/10.24201/edu.v24i3.1327

[22] Chamizo-Checa S., Otazo-Sánchez E., Gordillo-Martínez A. J., González- Ramírez, C. A., Suárez-Sánchez J. & Muñoz-Nava H. (2018). El cambio climático y la disponibilidad agua en sub-cuencas del Valle del Mezquital, México. *Revista Iberoamericana de Ciencias*, *5*(5), 40-51. Available at: https://www.reibci.org/publicados/2018/oct/3000102.pdf

[23] Neri-Ramírez, E., Rubiños-Panta, J. E., Palacios-Velez, O. L., Oropeza-Mota, J. L., Flores-Magdaleno, H., & Ocampo-Fletes, I. (2013). EVALUACIÓN DE LA SUSTENTABILIDAD DEL ACUÍFERO CUAUTITLÁN-PACHUCA MEDIANTE EL USO DE LA METODOLOGIA MESMIS. *Revista Chapingo Serie Ciencias Forestales Y Del Ambiente*, *19*(2), 273–285. https://doi.org/10.5154/r.rchs-cfa.2011.11.086

[24] Palacios-Vélez, O. L. & Escobar-Villagrán, B. S. (2016). La sustentabilidad de la agricultura de riego ante la sobreexplotación de acuíferos. *Tecnología y ciencias del agua, 7*(2), 5-16. Available at: https://www.scielo.org.mx/scielo.php?pid=S2007-24222016000200005&script=sci_arttext

[25] SEMARNAT (n.d.). *ESDIG*. https://gisviewer.semarnat.gob.mx/geointegrador2Beta/index.html

[26] CONAGUA. (n.d.). *Programa Hídrico Regional 2021-2024 - Región Hidrológico-Administrativa XIII Aguas del Valle de México*. Gobierno de México. Available at: https://files.conagua.gob.mx/conagua/generico/PNH/PHR_2021-2024_RHA_XIII_OCAVM.pdf

[27] Google Colab. (n.d.). Available at: https://colab.research.google.com/

# APPENDIX A: RESULTS IN TABLE FORMAT

Table A1. Groundwater Availability Analysis.

| aquifer | r | dnc | veas | dma_calculated | dmapositiva | dmanegativa |
|---|---|---|---|---|---|---|
| Zona Metro-politana de la Cd. de Mexico | 512.8 | 0 | 993.22991 | -480.42991 | 0 | -480.42991 |
| Cuautitlan-Pachuca | 356.7 | 0 | 586.64264 | -229.94264 | 0 | -229.94264 |
| Texcoco | 145.1 | 10.4 | 284.50512 | -149.80512 | 0 | -149.80512 |
| Chalco-Ame-cameca | 74 | 0 | 109.99354 | -35.993539 | 0 | -35.993539 |
| Tepeji del Rio | 46.3 | 35.2 | 15.354466 | -4.254466 | 0 | -4.254466 |
| El Astillero | 3.1 | 0 | 0.724197 | 2.375803 | 2.375803 | 0 |
| Apan | 30.3 | 0 | 27.543609 | 2.756391 | 2.756391 | 0 |
| Ajacuba | 25.7 | 15.6 | 5.187226 | 4.912774 | 4.912774 | 0 |
| Ixmiquilpan | 150.1 | 124.6 | 12.79055 | 12.70945 | 12.70945 | 0 |
| Chapan-tongo-Alfaja-yucan | 136.9 | 112.9 | 10.346407 | 13.653593 | 13.653593 | 0 |
| Tecocomulco | 27.8 | 0.5 | 9.955644 | 17.344356 | 17.344356 | 0 |
| Soltepec | 57 | 16.2 | 17.273231 | 23.526769 | 23.526769 | 0 |
| Valle del Mez-quital | 515 | 293 | 191.53726 | 30.462743 | 30.462743 | 0 |
| Actopan-San-tiago de Anaya | 208.1 | 90 | 64.955907 | 53.144093 | 53.144093 | 0 |

Table A2. Average Availability by Condition.

| condition | dma_calculated | dmanegativa | number of aquifers |
|---|---|---|---|
| Overexploited | -224.0428 | -224.0428 | 4 |
| Subexploited | 15.663151 | -0.425447 | 10 |

Table A3. Discrepancy Between VEAS and VCAS.

| aquifer | veas | vcas | extraction_vs_concession |
|---|---|---|---|
| Cuautitlan-Pachuca | 586.64264 | 551.07594 | 35.566701 |
| Valle del Mezquital | 191.53726 | 167.95202 | 23.585233 |
| Actopan-Santiago de Anaya | 64.955907 | 48.492584 | 16.463323 |
| Zona Metropolitana de la Cd. de Mexico | 993.22991 | 979.7622 | 13.467718 |
| Texcoco | 284.50512 | 273.40471 | 11.100417 |
| Ixmiquilpan | 12.79055 | 4.270279 | 8.520271 |
| Chalco-Amecameca | 109.99354 | 102.57045 | 7.42309 |
| Tepeji del Rio | 15.354466 | 8.335828 | 7.018638 |
| Tecocomulco | 9.955644 | 3.109183 | 6.846461 |
| Chapantongo-Alfajayucan | 10.346407 | 5.443209 | 4.903198 |
| Ajacuba | 5.187226 | 2.459013 | 2.728213 |
| El Astillero | 0.724197 | 0.038982 | 0.685215 |
| Soltepec | 17.273231 | 16.717709 | 0.555522 |
| Apan | 27.543609 | 27.053197 | 0.490412 |

Table A4. Extraction vs. Recharge Ratio.

| aquifer | veas | r | veas_vs_r_ratio | condition | situation |
|---|---|---|---|---|---|
| Texcoco | 284.50512 | 145.1 | 1.960752 | Overexploited | Without Availability |
| Zona Metropolitana de la Cd. de Mexico | 993.22991 | 512.8 | 1.936876 | Overexploited | Without Availability |
| Cuautitlan-Pachuca | 586.64264 | 356.7 | 1.644639 | Overexploited | Without Availability |
| Chalco-Amecameca | 109.99354 | 74 | 1.486399 | Overexploited | Without Availability |
| Apan | 27.543609 | 30.3 | 0.90903 | Subexploited | With Availability |
| Valle del Mezquital | 191.53726 | 515 | 0.371917 | Subexploited | With Availability |
| Tecocomulco | 9.955644 | 27.8 | 0.358117 | Subexploited | With Availability |
| Tepeji del Rio | 15.354466 | 46.3 | 0.33163 | Subexploited | Without Availability |
| Actopan-Santiago de Anaya | 64.955907 | 208.1 | 0.312138 | Subexploited | With Availability |
| Soltepec | 17.273231 | 57 | 0.303039 | Subexploited | With Availability |
| El Astillero | 0.724197 | 3.1 | 0.233612 | Subexploited | With Availability |
| Ajacuba | 5.187226 | 25.7 | 0.201838 | Subexploited | With Availability |
| Ixmiquilpan | 12.79055 | 150.1 | 0.085214 | Subexploited | With Availability |
| Chapantongo-Alfajayucan | 10.346407 | 136.9 | 0.075576 | Subexploited | With Availability |

Table A5. T-test results.

| Test Type | T-statistic | P-value | Significant? |
| --- | --- | --- | --- |
| t-test | 2.53869433849257 | 0.0842451635637601 | No (p > 0.05) |

Table A6. T-Test Summary: DMA by Aquifer Condition.

| Group | Mean DMA | Standard Deviation |
| --- | --- | --- |
| Subexploited | 15.663151 | 16.85459 |
| Overexploited | -224.0428 | 188.54082 |

# APPENDIX B: PYTHON CODES AND OUTPUTS

Figure B1. Libraries importing code snippet.

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Figure B2. File import into Google Colab.

```python
from google.colab import files

df = pd.read_csv("Disponibilidad Acuíferos_2023 - Valle de México.csv")
```

Figure B3. *df.head()* output.

```python
df.head()
```

| | clvacuifero | acuifero | clvrha | rha | r | dnc | vcas | vextet | veas | dmapositiva | dmanegativa | situacion | anio | condicion |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 901 | Zona Metropolitana de la Cd. de México | XIII | Aguas del Valle de México | 512.8 | 0.0 | 979.762196 | 0 | 993.229914 | 0.000000 | -480.429914 | Sin Disponibilidad | 2023 | SOBREEXPLOTADO |
| 1 | 1308 | El Astillero | XIII | Aguas del Valle de México | 3.1 | 0.0 | 0.038982 | 0 | 0.724197 | 2.375803 | 0.000000 | Con Disponibilidad | 2023 | SUBEXPLOTADO |
| 2 | 1309 | Chapantongo-Alfajayucan | XIII | Aguas del Valle de México | 136.9 | 112.9 | 5.443209 | 0 | 10.346407 | 13.653593 | 0.000000 | Con Disponibilidad | 2023 | SUBEXPLOTADO |
| 3 | 1310 | Valle del Mezquital | XIII | Aguas del Valle de México | 515.0 | 293.0 | 167.952024 | 0 | 191.537257 | 30.462743 | 0.000000 | Con Disponibilidad | 2023 | SUBEXPLOTADO |
| 4 | 1311 | Ajacuba | XIII | Aguas del Valle de México | 25.7 | 15.6 | 2.459013 | 0 | 5.187226 | 4.912774 | 0.000000 | Con Disponibilidad | 2023 | SUBEXPLOTADO |

Figure B4. *df.isnull()* output.

```
df.isnull()
```

| | clvacuifero | acuifero | clvrha | rha | r | dnc | vcas | vextet | veas | dmapositiva | dmanegativa | situacion | anio | condicion |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **1** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **2** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **3** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **4** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **5** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **6** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **7** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **8** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **9** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **10** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **11** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **12** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |
| **13** | False | False | False | False | False | False | False | False | False | False | False | False | False | False |

Figure B5. Python code for translations.

```python
# Creating column without accents for aquifer names
df['aquifer'] = df['acuifero'].str.replace('á', 'a').str.replace('é', 'e') \
                              .str.replace('í', 'i').str.replace('ó', 'o') \
                              .str.replace('ú', 'u').str.replace('ñ', 'n') \
                              .str.replace('Á', 'A').str.replace('É', 'E') \
                              .str.replace('Í', 'I').str.replace('Ó', 'O') \
                              .str.replace('Ú', 'U').str.replace('Ñ', 'N')

# Translating values from the "condicion" column
df['condition'] = df['condicion'].map({
    'SOBREEXPLOTADO': 'Overexploited',
    'SUBEXPLOTADO': 'Subexploited'
})

# Translating values from the "situacion" column
df['situation'] = df['situacion'].map({
    'Con Disponibilidad': 'With Availability',
    'Sin Disponibilidad': 'Without Availability'
})
```

Figure B6. Python code for *Section 4.1.1*

```
# Calculating average annual availability (DMA)
df['dma_calculated'] = df['r'] - df['dnc'] - df['veas']

# Sorting to graph
df_sorted = df.sort_values(by='dma_calculated')

# Bar chart
plt.figure(figsize=(12, 6))
bar_width = 0.2
x = range(len(df_sorted))

plt.bar(x, df_sorted['r'], width=bar_width, label='Recharge (R)')
plt.bar([p + bar_width for p in x], df_sorted['dnc'], width=bar_width, label='Committed Natural Discharge (DNC)')
plt.bar([p + bar_width*2 for p in x], df_sorted['veas'], width=bar_width, label='Groundwater Extraction (VEAS)')

plt.xticks([p + bar_width for p in x], df_sorted['aquifer'], rotation=90)
plt.ylabel('Volume (hm³/year)')
plt.title('Recharge, DNC, and Groundwater Extraction per Aquifer (2023)')
plt.legend()
plt.tight_layout()
plt.show()

# Showing results table
df_sorted[['aquifer', 'r', 'dnc', 'veas', 'dma_calculated', 'dmapositiva', 'dmanegativa']]
```

Figure B7. Python code for *Section 4.1.2*

```
# Calculating average annual availability
df['dma_calculated'] = df['r'] - df['dnc'] - df['veas']

# Number of aquifers by condition
condition_counts = df['condition'].value_counts().reset_index()
condition_counts.columns = ['Condition', 'Number of Aquifers']

# Tally chart
plt.figure(figsize=(6, 4))
sns.barplot(x='Condition', y='Number of Aquifers', data=condition_counts)
plt.title('Number of Aquifers by Exploitation Status (2023)')
plt.ylabel('Number of Aquifers')
plt.xlabel('Exploitation Status')
plt.tight_layout()
plt.show()

# Box plot to see distribution of the deficit (DMA-)
plt.figure(figsize=(6, 4))
sns.boxplot(data=df, x='condition', y='dmanegativa')
plt.title('Deficit (DMA-) by Aquifer Condition (2023)')
plt.ylabel('Deficit in Availability (hm³/year)')
plt.xlabel('Exploitation Status')
plt.tight_layout()
plt.show()

# Average deficits by condition
df.groupby('condition')[['dma_calculated', 'dmanegativa']].mean().reset_index()
```

Figure B8. Python code for *Section 4.1.3*

```
# Calculating the difference between the extracted and granted volume
df['extraction_vs_concession'] = df['veas'] - df['vcas']

# Sorting by difference
df_sorted = df.sort_values(by='extraction_vs_concession', ascending=False)

# Bar chart
plt.figure(figsize=(12, 6))
plt.bar(df_sorted['aquifer'], df_sorted['extraction_vs_concession'])
plt.xticks(rotation=90)
plt.ylabel('Difference (VEAS - VCAS) [hm³/year]')
plt.title('Discrepancy Between Extracted and Concessioned Volumes (2023)')
plt.tight_layout()
plt.show()

# Showing table with discrepancies
df_sorted[['aquifer', 'veas', 'vcas', 'extraction_vs_concession']]
```

Figure B9. Python code for *Section 4.1.4*

```
# Calculating DMA
df['dma_calculated'] = df['r'] - df['dnc'] - df['veas']

# Scatter plot: VEAS vs DMA
plt.figure(figsize=(8, 6))
sns.scatterplot(data=df, x='veas', y='dma_calculated', hue='condition', style='situation', s=100)
plt.axhline(0, color='gray', linestyle='--')
plt.title('Groundwater Extraction vs. Availability (2023)')
plt.xlabel('Groundwater Extraction (VEAS) [hm³/year]')
plt.ylabel('Annual Availability (DMA) [hm³/year]')
plt.tight_layout()
plt.show()

# Table to compare status and overexploitation
df[['aquifer', 'veas', 'dma_calculated', 'condition', 'situation']].sort_values(by='dma_calculated')
```

Figure B10. Python code for *Section 4.1.5*

```
# Calculating the VEAS / R ratio
df['veas_vs_r_ratio'] = df['veas'] / df['r']

# Sorting for display
df_sorted = df.sort_values(by='veas_vs_r_ratio', ascending=False)

# Bar chart
plt.figure(figsize=(12, 6))
plt.bar(df_sorted['aquifer'], df_sorted['veas_vs_r_ratio'])
plt.xticks(rotation=90)
plt.ylabel('Extraction / Recharge Ratio')
plt.title('Groundwater Extraction Relative to Recharge (2023)')
plt.axhline(1, color='red', linestyle='--', label='Sustainability Threshold (Ratio = 1)')
plt.legend()
plt.tight_layout()
plt.show()

# Summary table
df_sorted[['aquifer', 'veas', 'r', 'veas_vs_r_ratio', 'condition', 'situation']]
```

Figure B11. Python code for *Section 4.1.6*

```
from scipy.stats import ttest_ind

# Calculating DMA
df['dma_calculated'] = df['r'] - df['dnc'] - df['veas']

# Separating into two groups
subexploited = df[df['condicion'] == 'SUBEXPLOTADO']['dma_calculated']
overexploited = df[df['condicion'] == 'SOBREEXPLOTADO']['dma_calculated']

# T-test for difference of means
t_stat, p_val = ttest_ind(subexploited, overexploited, equal_var=False)

# Summary results
summary = pd.DataFrame({
    'Group': ['Subexploited','Overexploited'],
    'Mean DMA': [subexploited.mean(), overexploited.mean()],
    'Standard Deviation': [subexploited.std(), overexploited.std()]
})

print(summary)
print("\nT-statistic:", t_stat)
print("P-value:", p_val)

# Creating the boxplot
plt.figure(figsize=(6, 5))
sns.boxplot(data=df, x='condition', y='dma_calculated', palette='Set2')

# Adding titles and tags
plt.title('Distribution of Annual Availability (DMA) by Aquifer Condition')
plt.ylabel('Annual Availability (DMA) [hm³/year]')
plt.xlabel('Aquifer Condition')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```