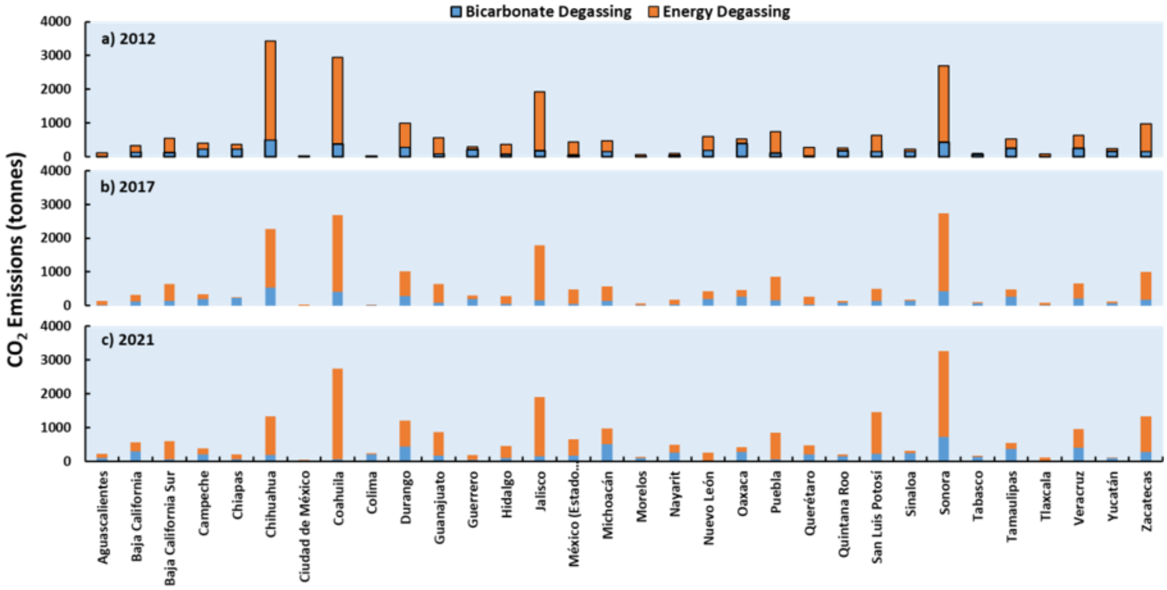
Supplementary Items for:

Hidden carbon fluxes from groundwater use in Mexico using satellite storage change and bicarbonate degassing

Supplementary Item S1. Geographic distribution of Mexico's 32 federal entities.

SF 1. Geographic distribution of Mexico's 32 federal entities. State boundaries are delineated, and major adjacent water bodies— including the North Pacific Ocean, Gulf of Mexico, Bahia de Campeche, and Caribbean Sea—are labelled. This map provides spatial context for state-level analyses of environmental and hydrological variables examined in this study. International borders with the United States, Honduras, and Nicaragua are shown for reference.

Supplementary Item S2. State-wise stacked CO2 emissions



SF 2. State-wise stacked CO_2 emissions (tonnes) from bicarbonate degassing, and pumping energy CO_2 equivalent for 2012 (a), 2017 (b), and 2021 (c). Total emissions rise sharply over the decade, primarily driven by Chihuahua, Sonora, Zacatecas, Guanajuato, and San Luis Potosí, where intensive irrigation and groundwater depletion lead to both chemical and energy-related fluxes. In contrast, Veracruz, Tabasco, Campeche, Yucatán, and Chiapas maintain low emission magnitudes due to high recharge and minimal dependence on pumped irrigation. Stacked emission profiles are computed using state-level pumping depth distributions derived from the Profundidad al Nivel Estático (PNE) dataset for the years 2012, 2017 and 2021 and using the degassing fraction of 0.35.

Supplementary Item S3. GRACE RL06M uncertainty characterization and propagation

Groundwater storage anomalies were derived from the JPL GRACE/GRACE-FO RL06M mascon solution. The RL06M product retains the equal-area 3° spherical cap mascon architecture and post-processing framework (including Coastline Resolution Improvement filtering and hydrological gain factors) described by Wiese et al. (2016). Although RL06M incorporates updated background models, calibration and dealiasing relative to RL05M, the structure of measurement and leakage uncertainty remains consistent with the mascon error characterization framework documented by Wiese et al. (2016).

Components of GRACE Uncertainty

Total uncertainty in mascon-derived terrestrial water storage arises from two primary components:

1. Measurement error (Em), representing instrument noise and processing uncertainty derived from the formal covariance matrix.
2. Leakage error (El), arising from spatial smoothing associated with mascon parameterization and partial mascon coverage of region boundaries.

Wiese et al. (2016) quantified these terms for 52 hydrological basins using synthetic simulations. For basins comparable in size to the Sacramento–San Joaquin basin (~160,000 km^2), they report:

Measurement error (E_m) = 13.7 mm equivalent water height (EWH)

Residual leakage error after application of the CRI filter and gain factors = 4.5 mm EWH.

The total root-mean-square (RMS) uncertainty is therefore:

$$\text{Total uncertainty} = \sqrt{(Em^2 + El^2)}$$

$$= \sqrt{(13.7^2 + 4.5^2)}$$

$$= \sqrt{(187.69 + 20.25)}$$

$$= \sqrt{207.94}$$

$$\approx 14.4 \text{ mm EWH}$$

Scale Dependence of Uncertainty

Wiese et al. (2016) demonstrate that uncertainty depends strongly on region size relative to the native mascon footprint (~3°; ~90,000 km^2 typical area). As spatial scale decreases, leakage error increases relative to measurement error. Based on their synthetic basin analysis, uncertainty regimes can be classified as follows:

- Large regions (>160,000 km^2): measurement-dominated uncertainty; total RMS ≈ 14–16 mm EWH.
- Intermediate regions (~100,000–150,000 km^2): moderate leakage sensitivity; total RMS ≈ 15–18 mm EWH.
- Small regions (<50,000 km^2): leakage may approach or exceed measurement error; total RMS ≈ 18–25 mm EWH unless aggregation is applied.

Because most individual aquifers are substantially smaller than the mascon footprint, groundwater storage anomalies were evaluated at state-aggregated scales to ensure stability and remain within the reliable detection envelope of the RL06M mascon framework.

Representative Uncertainty for Mexican Aquifer–State Systems

State (Representative Aquifer)	Effective Aggregation Area (km²)	Scale Category	Measurement Error (mm EWH)	Residual Leakage Error (mm EWH)	Total RMS Uncertainty (mm EWH)
Sonora (Costa de Hermosillo aquifer aggregated to state)	~179,000	Large	~13–15	~4–6	~14–16
Chihuahua (Chihuahua–Sacramento Valley aquifers aggregated to state)	~247,000	Large	~13–15	~4–6	~14–16
Yucatán (Península de Yucatán aquifer system aggregated regionally)	~124,000	Intermediate	~13–14	~5–8	~15–18
Guanajuato (Valle de León aquifer aggregated to state)	~30,600	Small	>15	8–15	~18–25

For large northern states such as Sonora and Chihuahua, the aggregation area exceeds the 160,000 km^2 benchmark, implying total GRACE uncertainty of approximately 14–16 mm EWH. For intermediate-scale regions such as Yucatán (~124,000 km^2), leakage sensitivity increases modestly, yielding uncertainties of ~15–18 mm EWH. Smaller states such as Guanajuato are more leakage sensitive; therefore, interpretation relies on aggregated spatial evaluation rather than sub- mascon aquifer-scale extraction.

<u>Signal-to-Noise Ratio of Observed Depletion</u>

Signal-to-noise ratios (SNR) were calculated using a conservative uncertainty estimate of 14.5 mm EWH for large state-level regions.

Example: Chihuahua

Observed cumulative depletion ≈ −750 mm

SNR = 750 / 14.5 ≈ 51.7

Example: Moderate depletion (≈ −400 mm)

SNR = 400 / 15 ≈ 26.7

Example: Weak anomaly (≈ −10 mm)

SNR = 10 / 14.5 ≈ 0.69

Depletion hotspots exceeding 400–750 mm therefore exhibit SNR values >25–50, indicating that GRACE measurement and residual leakage uncertainty contribute less to the magnitude of inferred groundwater loss.

Propagation into Emission Estimates

Because CO_2 emissions derived in this study scale linearly with volumetric groundwater depletion, uncertainty in equivalent water height propagates directly into volumetric and emission estimates. For large state-level aggregations, GRACE contributes approximately ±14–16 mm EWH uncertainty. For depletion magnitudes exceeding 300–750 mm, this corresponds to a relative low uncertainty. Consequently, uncertainties associated with pumping lift assumptions, pump efficiency, and energy emission factors exceed the GRACE measurement contribution in the final emission budget.

Supplementary Item S4. Pumping energy and efficiency parameterization

Electrical energy required for groundwater abstraction was calculated as:

$$E = \frac{\rho g V H}{\eta}$$

where ρ is water density, g gravitational acceleration, V extracted groundwater volume, H pumping lift, and η electrical pumping efficiency.

Experimental evaluation of centrifugal irrigation pumps under variable frequency operation demonstrates significant efficiency variability under practical hydraulic conditions. Mexico-specific assessments indicate lower field efficiencies. Direct measurement of irrigation pumping plants in the eastern Valley of Mexico reported efficiencies between 23% and 56%, with rehabilitation recommended below 40%. Evaluation of energy performance in Mexican aqueduct pumping systems further demonstrates persistent operational inefficiencies and measurable energy-saving potential through system optimization. Optimization studies of irrigation pumping networks confirm that real-world performance frequently deviates from nominal design efficiencies due to operational constraints.

Based on this evidence, we adopt a conservative Mexico-relevant efficiency range of 0.30–0.50, with a central value of 0.35.

Supplementary Item S5. Degassing fractions calculations and theory

Carbonate weathering in calcite-bearing aquifers can be represented as:

$$\mathbf{CO_2 + H_2O + CaCO_3 \rightarrow 2HCO^- + Ca^{2+}}$$

Under this reaction pathway, half of the bicarbonate carbon originates from atmospheric/soil CO_2 and half from the carbonate mineral matrix. Upon atmospheric exposure, re-equilibration may proceed as:

$$\mathbf{Ca^{2+} + 2HCO^- \rightarrow CaCO_3 + CO_2 + H_2O}$$

This stoichiometric relationship implies that 50% of bicarbonate carbon represents a conservative upper bound for potential atmospheric CO□ release in carbonate- buffered systems. We have calculated the degassing fractions using PHREEQC (v3.8.6) equilibrium simulations by incorporating measured pH, temperature, and alkalinity ranges from the groundwater dataset. The equilibrium simulations were performed under the following five assumptions:

1. Atmospheric Boundary Condition: Groundwater re-equilibrates with atmospheric CO_2 at:

   $$\mathbf{\log 10\ \mathit{p}CO2 = -3.5}$$

   Thermodynamic Equilibrium: Reactions proceed to equilibrium; kinetic effects are not considered.

2. Carbonate System Control: Carbon speciation and mineral equilibria are governed by the phreeqc.dat thermodynamic database.

3. Closed Water Mass During Degassing: Only dissolved inorganic carbon redistributes between aqueous and gaseous phases; water mass remains constant.

Because equilibrium degassing is controlled by hydrochemical state variables (pH, alkalinity, temperature, ionic strength) rather than sample count, a representative subset of 50 samples was selected to span the complete hydrochemical envelope.

The selected dataset spans:

pH: 6.3 – 9.3
Temperature: 18.1 – 59.8 °C
Calcium: 2.15 – 1333 mg/L
Magnesium: 0.20 – 778.2 mg/L
Sodium: 8.36 – 2905 mg/L
Chloride: 5 – 5371 mg/L
Sulphate: 3.72 – 3470.66 mg/L
Alkalinity (as HCO_3^-): 29 – 950 mg/L

These ranges encompass low and high alkalinity waters, carbonate-dominated and saline waters, and the full observed pH spectrum. Additional samples within this envelope do not expand the thermodynamic limits of degassing. Degassing Fraction Calculation. For each sample:

$$\boldsymbol{f_i} = \frac{\boldsymbol{C_{initial,i}} - \boldsymbol{C_{equilibrium,i}}}{\boldsymbol{C_{initial,i}}}$$

where:

$C_{initial,i}$ = Total dissolved inorganic carbon before equilibration

$C_{equilibrium,i}$ = Total dissolved inorganic carbon after equililibration at $\log p\mathrm{CO_2}$ = −3.5

Sample Calculations

Sample 1

$$\boldsymbol{Cinitial} = \mathbf{5.483 \times 10{-}3} \text{ mol/kg}$$

$$\boldsymbol{Cequilibrium} = \mathbf{4.044 \times 10{-}3} \text{ mol/kg}$$

$$\boldsymbol{f} = \frac{\mathbf{5.483 \times 10^{-3} - 4.044 \times 10^{-3}}}{\mathbf{5.483 \times 10^{-3}}}$$

$$\boldsymbol{f} = \mathbf{0.262}\ (\mathbf{26.2}\%)$$

Sample 2

$$\boldsymbol{Cinitial} = \mathbf{9.145 \times 10{-}3}$$

$$\boldsymbol{Cequilibrium} = \mathbf{6.300 \times 10{-}3}$$

$$\boldsymbol{f} = \frac{\mathbf{9.145 - 6.300}}{\mathbf{9.145}}$$

$$\boldsymbol{f} = \mathbf{0.311}\ (\mathbf{31.1}\%)$$

Across the representative subset:

$\mathbf{0.00} \le \boldsymbol{f} \le \mathbf{0.35}$

Therefore:

$$\boldsymbol{f_{modelled,max}} = \mathbf{0.35} < \mathbf{0.50}$$

The 50% value is retained as a conservative theoretical upper bound based on carbonate stoichiometry, while 0.00–0.35 represents the thermodynamically constrained degassing range for the present aquifer system.

Supplementary Item S6. Pumping lift estimation and uncertainty characterization

Percentile-based lift statistics (Q50–Q90) are used to characterize variability and uncertainty, while a representative value of 100 m is applied in the emission calculations. State-level pumping lift (h) was added from piezometric depth records compiled at the state scale for the period 2010–2024 to show the sensitivity associated with the depth variation. For each state, all available annual PNE [Profundidad al Nivel Estático (PNE; static water level depth, m below ground surface)] observations between 2010 and 2024 were aggregated, and percentile statistics were computed using the Excel-equivalent PERCENTILE.INC formulation to ensure reproducibility and consistency with reported state statistics. Three distribution metrics were evaluated: median (Q50), upper quartile (Q75), and 90th percentile (Q90).

The median (Q50) represents a typical well depth but underestimates the energy burden because pumping energy scales linearly with lift. Conversely, the 90th percentile (Q90) may be disproportionately influenced by extreme deep wells and may overstate representative energy demand at the state scale. The Q75 depth represents the upper-typical operational range, capturing intensively pumped wells without being dominated by statistical outliers. This approach is consistent with energy-intensity analyses in groundwater pumping studies, where conservative but distribution-aware metrics are preferred over fixed-depth assumptions.

To reduce interannual variability while preserving long-term depletion signals, a 5- year rolling mean was applied prior to percentile extraction for the benchmark years 2012, 2017, and 2021. Specifically:

- 2012 depth metrics were computed using 2009–2012 data
- 2017 depth metrics were computed using 2014–2017 data
- 2021 depth metrics were computed using 2018–2021 data

For states with incomplete annual records within a rolling window, percentiles were computed from all available values within that window. Where missing values persisted, imputation was avoided unless explicitly required for sensitivity testing; no artificial depth assumptions were introduced into the core analysis.

Uncertainty treatment for lift

This framework propagates statistically observed variability in well depth into energy and CO_2 calculations. Because pumping energy scales linearly with lift:

$$\boldsymbol{E} \propto \boldsymbol{h}$$

the relative uncertainty in lift directly translates into proportional uncertainty in pumping emissions.
Groundwater pumping energy was calculated as:

$$\boldsymbol{E} = \frac{\boldsymbol{\rho g V H}}{\boldsymbol{\eta}}$$

where:

ρ = water density (1000 kg m^{-3})
g = gravitational acceleration (9.81 m s^{-2})
V = pumped groundwater volume (m^3)
h = lift (m)
η = pumping electrical efficiency (0.30–0.50 range; central 0.35)

References:

1. Wiese, D. N., Landerer, F. W., & Watkins, M. M. (2016). Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. Water Resources Research, 52(9), 7490-7502.
2. Córcoles, J. I., Tarjuelo, J. M. & Moreno, M. A. Methodology to improve pumping station management of on-demand irrigation networks. *Biosyst. Eng.* **144**, 94–110 (2016).
3. Bordeasu, D., Dragan, F., Filip, I. & Tirian, G. O. Estimation of centrifugal pump efficiency at variable frequency for irrigation systems. *Sustainability* **16**, 2024 (2024).
4. Köpp, L. M., Peiter, M. X., Robaina, A. D. & Toescher, C. F. Characterization of pumping stations in irrigated rice crops from Rio Grande do Sul, Brazil. *Eng. Agríc.* **36**, 2016 (2016).
5. Firouzabadi, A. G., Seyedan, S. M., Jovzi, M. & Peyghan, K. Study on the efficiency and energy consumption of electric and diesel pumping stations. *Proc. Inst. Civ. Eng. Energy* **175**, 2022 (2022).
6. Mauricio, C. G., Bonifacio, G. P. & Alejandra, G. P. Y. Irrigation pumping plant efficiency in wells in eastern Valley of Mexico. *ASABE Annu. Int. Meet.* (2014).
7. Euyoqui Aréchiga, F. J., Suástegui Macías, J. A., Bonilla, D. & Magaña Almaguer, H. D. Evaluation of energy efficiency actions in Mexican aqueducts with an approach on their performance over time: Case study Northwest Mexico aqueduct. *Heliyon* **10**, 2024 (2024).
8. Ahadi, R., Samani, Z. & Skaggs, R. Evaluating on-farm irrigation efficiency across the watershed: A case study of New Mexico's Lower Rio Grande Basin. *Agric. Water Manag.* **122**, 2013 (2013).
9. Chávez, C., Limón-Jiménez, I., Espinoza-Alcántara, B. & Trejo-Alonso, J. Water-use efficiency and productivity improvements in surface irrigation systems. *Agronomy* **10**, 2020 (2020).