Agrociencia

eISSN: 2521-9766

VOLUME 60, NUMBER 4 | May 16 - June 30 , 2026 | MEXICO

AGRICULTURA

SECRETARÍA DE AGRICULTURA

All correspondence (subscription information, sales, advertising, author contributions, etc.) should be addressed to:

Central Office:

Agrociencia
Guerrero No. 9, Esquina con Avenida Hidalgo,
San Luis Huexotla, Texcoco 56220,
State of Mexico. MEXICO
Tel.: +52-595 92 84427

https://agrociencia-colpos.org/index.php/agrociencia

Cover: *Ficus carica* L.
Photograph by Digital library of traditional Mexican medicine

AGRICULTURA
SECRETARÍA DE AGRICULTURAY DESARROLLO RURAL

Contents

Volume 60, Number 4 | May 16 - June 30, 2026

Water-Soil-Climate

Wildlife

Agrociencia

PROXIMAL AND GROWTH ANALYSIS OF FORAGE SOYBEAN (*Glycine max* (L.) Merr.) VARIETIES

Joel **Ventura-Ríos**[1], María de los Ángeles **Maldonado-Peralta**[2], Adelaido Rafael **Rojas-García**[2*], Aldenamar **Cruz-Hernández**[3], Herminio Aniano **Aguirre**[4]

[1]Universidad Autónoma Agraria Antonio Narro. Calzada Antonio Narro 1923, Buenavista, Saltillo, Coahuila, Mexico. C. P. 25315.
[2]Universidad Autónoma de Guerrero. Facultad de Medicina Veterinaria y Zootecnia No. 2. Cuajinicuilapa, Guerrero, Mexico. C. P. 41940.
[3]Universidad Juárez Autónoma de Tabasco. División Académica de Ciencias Agropecuarias. Carretera Villahermosa-Teapa km 25, R/A, La Huasteca Segunda Sección, Villahermosa, Tabasco, Mexico. C. P. 86290.
[4]Tecnológico Nacional de México. Campus Pinotepa Nacional. Avenida Tecnológico 1155, Sección Primera, Colonia La Soledad, Santiago Pinotepa Nacional, Oaxaca, Mexico. C. P. 71602.
* Author for correspondence: rogarcia@uagro.mx

Citation: Ventura-Ríos J, Maldonado-Peralta MA, Rojas-García AR, Cruz-Hernández A, Aguirre HA. 2026. Proximal and growth analysis of forage soybean (*Glycine max* (L.) Merr.) varieties. **Agrociencia** 60(4): 467-479. https://doi.org/ 10.47163/agrociencia.v60i4.3577

Editor in Chief:
Dr. Fernando C. Gómez Merino

Received: September 30, 2025.
Approved: May 20, 2026.
Published in Agrociencia: May 28, 2026.

ABSTRACT

Due to their broad productivity and nutritional value, legumes are used in animal production systems. The aim of this study was to evaluate the productive behavior and the chemical quality of different varieties of forage soybean (*Glycine max* (L.) Merr.) using a growth analysis in the dry tropics of Mexico. The treatments included the following evaluated varieties: Salcer, Ojo de Tigre, Valente, and Albina. The variables determined were total dry matter (TDM), yield by component, plant height, intercepted radiation, plants per square meter, weight per stem, crude protein, acid detergent fiber, and neutral detergent fiber. The variety with the highest TDM yield was Ojo de Tigre, with 5562 kg ha^{-1}, whereas Salcer had the lowest, with 4626 kg ha^{-1} ($p < 0.05$). The intercepted radiation increased with the age of the plant in all four varieties until day 61 after harvest, reaching an average of 96 %, a value that remained until day 68. The genotype with the best attributes in dry matter yields, structural characteristics, and plant height was Ojo de Tigre. The optimum moment for harvest was established at day 61, when an average of 95 % intercepted radiation and better quality were obtained, which defines the optimum cutting point for the conditions of the dry tropics of Mexico.

Keywords: legumes, yield, intercepted radiation, crude protein.

INTRODUCTION

In tropical regions of Mexico, prairies are mainly composed of forage grasses (Olivares-Pérez *et al*., 2005) and some perennial legumes, planted in protein banks (Pamo *et al*., 2007; Rojas-García *et al*., 2021) or associated with grasses (Jiménez *et al*., 2002; Benítez-Bahena *et al*., 2010). However, they face prolonged droughts and management limitations in grazing areas. To mitigate this situation, improved grasses have been

introduced (Torres-Salado *et al*., 2020; Ventura-Ríos *et al*., 2021; Rojas-García *et al*., 2024). Nevertheless, the need to identify native or introduced forage legumes that provide higher nutritional value and dry matter (DM) yields persists (Maldonado-Peralta *et al*., 2023). Traditional animal production has negative impacts on natural ecosystems (Benítez-Bahena *et al*., 2010), and therefore, alternatives such as the use of annual legumes, including forage soybean, are considered (Rojas-García *et al*., 2025; Maldonado-Peralta *et al*., 2025).

Soybean is an annual legume that produces high nutritional quality forage (Pagano and Miransari, 2016), with green matter yields between 19 and 30 Mg ha^{-1} and grain yields of up to 3500 kg ha^{-1} (Tobía and Villalobos, 2004; Tobía, 2011; Hernández *et al*., 2013). This forage is appealing to cattle, with a high crude protein content (16.4–17.3 %), total digestible nutrient levels of 58–68 %, and metabolizable energy between 2.2 and 2.5 Mcal kg^{-1} of DM (Garay-Martínez *et al*., 2021), which helps reduce the use of concentrates in the diets of the animals and improves profitability (Teixeira *et al*., 2016). Ríos-Hilario *et al*. (2023), when evaluating the Salcer variety in the dry tropics in Guerrero, Mexico, found yields of 4462 kg ha^{-1} of DM with high planting densities (250 000 plants ha^{-1}) and 2958 kg ha^{-1} with lower densities (41 250 plants ha^{-1}).

Most available studies focus on soybean cultivars for the production of grain (Vega and Andrade, 2000; Enciso-Maldonado *et al*., 2021) planted at different densities and harvest dates, where yield was affected by the planting season and the population density. In other legumes, such as chepil (*Crotalaria longirostrata* Hook. and Arn.), Maldonado-Peralta *et al*. (2023) reported 95 % of intercepted radiation 36 d after cutting, with crude protein, acid detergent fiber, and neutral detergent fiber of 22, 58, and 37 %, respectively; subsequently, the intercepted radiation dropped to 62 % on day 79.

In contrast, studies on forage soybean varieties in tropical regions of Mexico are scarce; therefore, agronomic studies and proximal chemical analyses are needed. In this work, the proposed hypothesis is that, over time, soybean yield increases while its chemical quality decreases, making the optimal cutting time between days 61 and 68, depending on the genotype. Accordingly, the aim of this study was to evaluate total dry matter yield and yield by component, plant height, intercepted radiation, population density, stem weight per stem, crude protein, and acid and neutral detergent fiber of different forage soybean varieties through growth analysis in the dry tropics of Mexico.

MATERIALS AND METHODS

Study area

The evaluation of productive behavior was carried out in the town of El Pitahayo, Cuajinicuilapa, Guerrero, Mexico (16° 32′ 5.49″ N and 98° 30′ 28.87″ W), from September 1, 2020, to February 28, 2021, at an altitude of 50 m (INEGI, 2010). Weather data were obtained from the station located in Cuajinicuilapa, Guerrero, 10 km from

the experimental site (CONAGUA, 2022). The area has a warm, subhumid climate, with the highest rainfall during the study period recorded in September, October, November, and December, 2020, accumulating 832 mm, and a mean temperature of 26 °C (Figure 1).

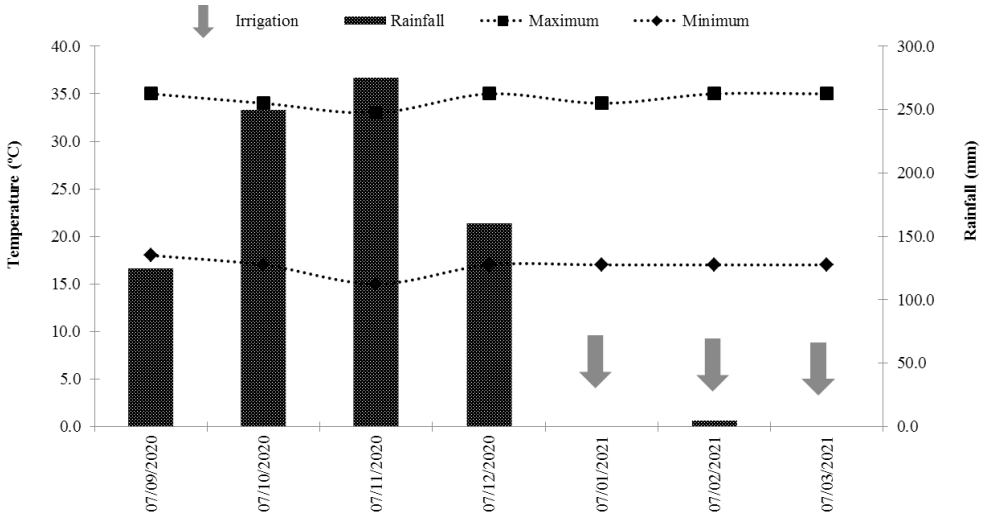

Figure 1. Rainfall (mm) and maximum and minimum temperatures (°C) during the study period in the municipal area of Cuajinicuilapa, Guerrero, Mexico.

Experimental plots

The Salcer variety of forage soybean (*Glycine max* (L.) Merr.) was evaluated, along with three genotypes: Ojo de Tigre, Valente, and Albina, developed by the Center for Professional Studies of the Higher Agricultural College of the State of Guerrero (CEP-CSAEGro) through different selection processes. Salcer is the only variety registered within the National Seed Inspection and Certification Service (SNICS); the remaining genotypes are in the evaluation process.

The experiment was established under a completely randomized block design with three repetitions. Every genotype was planted in a 10 × 10 m plot, corresponding to each treatment. The soil had a silty-sandy texture, a 15 % slope, a slightly alkaline pH (7–8), and low organic matter content (0.2 %). The land was prepared using conventional tillage: plowing once, harrowing twice, furrowing, and layout marking. Planting was carried out on September 7, 2020, during the rainy season, with a density of 230 000 seeds ha^{-1}. Each seed was placed at the top of the furrow, in a double row, with a topological arrangement of 80 cm between furrows, 10 cm between rows, and 12 cm between seeds.

Chemical fertilization was carried out on days 20 and 30 after planting (DAP), using the formula 60-40-00 (N-P-K). Ammonium sulfate served as the nitrogen source (Fertiquim, Mexico). For phosphorus, diammonium phosphate (18-46; Fertiquim, Mexico) was applied by squirting it 5 cm from the stem. For foliar fertilization, Gro Green Campbell 20-30-10 (Gor Green, Mexico) and Supra Fe (Arvensis, Mexico, as a source of iron) were sprayed using a motorized backpack sprayer (Honda, Japan) at 20, 30, and 40 DAP.

For pest control, cypermethrin was applied three times (CIMA, Química Sagal S.A., Mexico) at a dose of 1 mL L^{-1} of water, aimed at the black cutworm (*Agrotis ípsilon* Hufnagel, 1766), using a motorized backpack sprayer (Honda, Japan). Weed control was carried out twice, chemically and manually. The chemical control consisted of the targeted application of a non-selective contact herbicide (paraquat) at a dose of 1 mL L^{-1} of water, using a manual spray with a hooded nozzle. Manual control was carried out using a hoe.

Agronomic samplings were carried out at 26, 33, 40, 47, 54, 61, 68, 75, 82, and 89 DAP. The proximal chemical analyses were performed at 26, 40, 54, 68, and 82 DAP. In the months of minimum rainfall (January, February, and March 2021), underground water was used to irrigate every 8 d, with a sheet of approximately 32 mm per irrigation.

Variables evaluated

Total yield and yield by component

One day before each cut, two samples were taken from each plot, considering 1 m per furrow. The plants were separated by components (leaf, stem, and pod) and weighed on a digital precision scale (METALTEX, VE-200RT, Switzerland). Subsequently, they were dried at 55 °C until they reached a constant weight in a forced-air oven (Faithful, WGLL-230BE, China). Later, they were weighed once again to determine the content of dry matter and to estimate the yield of dry matter per hectare.

Plant height

The readings of 20 plants were taken at random in each experimental plot one day before every cut, using a measuring tape (Truper, Mexico) 3 m in length. The measurement was carried out at ground level until the last morphological component.

Intercepted radiation

Five readings were taken for every experiment. A 100-cm wooden ruler was placed under the canopy, pointing south-north, and the centimeters with shade were counted, which represented the percentage of radiation intercepted by the plant's canopy.

Population density (plants per m^2)

At the beginning of the experiment, 1 m^2 areas were traced at random in each plot using wooden stakes. In each sampling, the number of plants in the area was counted to determine the population density and evaluate the persistence of each genotype.

Weight by stem

Five stems were cut at random from each plot, at ground level and with leaves, which were dried in a forced-air oven until their weight was constant. Subsequently, the dry weight of every stem was registered.

Proximal chemical analysis

The proximal chemical analysis was carried out in the animal nutrition laboratory of the School of Veterinary Medicine and Animal Husbandry No. 2 of the Autonomous University of Guerrero, located in the municipal seat of Cuajinicuilapa. The determinations were carried out on days 26, 40, 54, 68, and 82. The samples used to estimate the dry matter yield were used to determine the crude protein (CP) content using the method described by the AOAC (2005). All analyses were carried out in triplicate. The neutral detergent fiber (NDF) and the acid detergent fiber were determined following the method proposed by van Soest *et al*. (1991).

Statistical analysis

The data gathered was analyzed under a completely randomized block design with three repetitions and a divided plot arrangement, in which the large plot corresponded to the cuts (days after planting) and the small plot, to the soybean genotype evaluation. The analysis was carried out with the PROC GLM procedure by SAS (SAS Institute, 2011), where the effects of harvest time were considered fixed. The multiple comparison of means was carried out with Tukey's test ($\alpha = 0.05$).

RESULTS AND DISCUSSION

T

otal yield and yield by component

The total dry matter yield of the four soybean varieties at different cutting ages during the rainy season (Figure 2) increased with plant age until day 75 after planting, reaching an average of 9422 kg ha^{-1}. Subsequently, it decreased until day 89, with 7033 kg ha^{-1} ($p < 0.05$). The variety with the highest yield was Ojo de Tigre, with 5562 kg ha^{-1}, whereas the one with the lowest yield was Salcer, with 4626 kg ha^{-1} ($p < 0.05$). However, Albina displayed the highest yield on day 75, with 12 542 kg ha^{-1} ($p < 0.05$). The four varieties displayed a similar component pattern. The leaf component predominated with more than 60 % of the yield until day 40 after planting; thereafter, the stem became dominant until day 89. The pod component appeared on day 61, with a greater proportion in the Albina variety and a lower one in Salcer ($p < 0.05$) (Figure 2). Aponte *et al*. (2015) reported similar values in 10 forage lines in Puerto Rico, with DM yields of 5300 and 6573 kg ha^{-1} on days 54 and 68, respectively. Tobía *et al*. (2006), in evaluations performed in Venezuela and Costa Rica, recorded yields between 5000 and 13 000 kg ha^{-1} on day 75, comparable to those obtained in this study. In turn, Ríos-Hilario *et al*. (2023) reported values similar to those found in this study on day 60 of

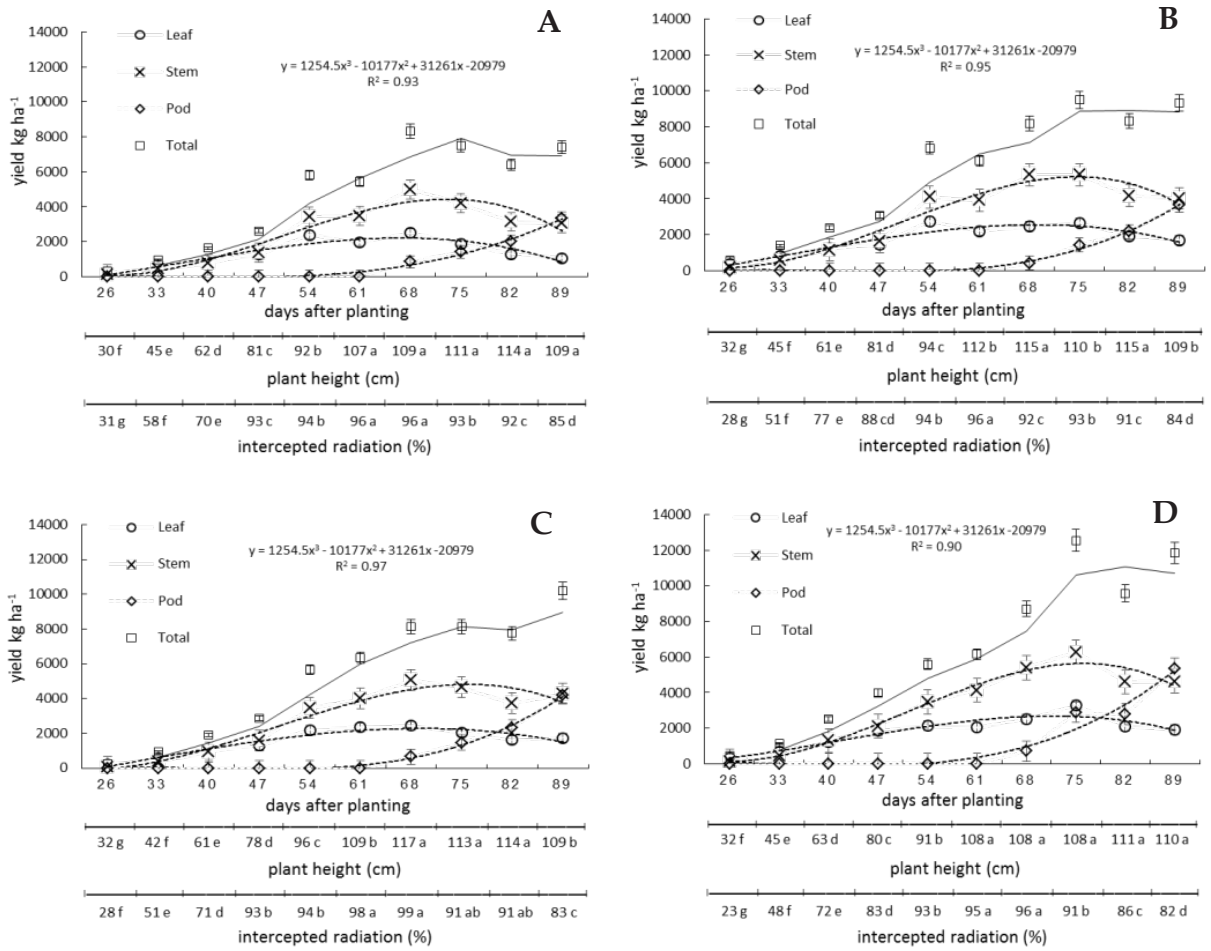

Figure 2. Total dry matter (DM) and yield by component (kg ha^{-1} of DM), plant height (cm), and intercepted radiation (%) of different forage soybean (*Glycine max* (L.) Merr.) genotypes in the dry tropics of Mexico. A: Salcer; B: Ojo de tigre; C: Valente; D: Albina.

plant growth, with 5500 kg ha^{-1} at a population density of 250 000 plants ha^{-1} in the Salcer variety.

In Brazil, Wilson-Cortez *et al*. (2011) evaluated the Monsoy 5942 variety at different population densities and fertilization depths and reported an average DM yield of 5634 kg ha^{-1} on day 68, lower than that obtained in this study on the same day after planting. In Turkey, Acikgoz *et al*. (2007) evaluated three forage soybean varieties on days 54, 68, and 82, with yields of 8467, 11 356, and 14 225 kg ha^{-1}, respectively, which were higher than those recorded in this study, possibly due to more favorable weather conditions, since plant growth varies depending on temperature and nutrient availability in the soil. Altogether, these results show that yield in forage soybean varies depending on the genotype, cutting date, geographic location, and soil fertility.

Plant height

The plant height of the four forage soybean varieties at different cutting ages during the rainy season (Figure 2) increased as plant age advanced, regardless of the genotype. The maximum height was recorded on day 82, with 114 cm, and later decreased to 109 cm on day 89. The lowest values were observed between days 26 and 54 ($p < 0.05$). On average, the tallest genotypes were Valente and Ojo de Tigre (87 cm), whereas Salcer and Albina presented the lowest values (86 cm).

Tosquy-Valle *et al*. (2009) reported lower heights in two soybean varieties evaluated in the Mexican tropics at different planting densities, which depend on the variety and regional temperature. After day 61, they recorded heights ranging from 44 to 54 cm. In Cuba, Romero-Arias *et al*. (2013) evaluated seven varieties and reported 49 cm on day 60 after planting. In Venezuela, Hernández *et al*. (2013) found average heights of 67 cm on day 75 in nine forage soybean varieties, which were also lower than those found in this study. Ríos-Hilario *et al*. (2023) reported an average height of 58 cm for the Salcer variety when evaluating different population densities and cutting ages. In this study, Valente and Ojo de Tigre displayed the greatest average height (87 cm), surpassing Salcer and Albina, which showed lower growth.

Intercepted radiation

The intercepted radiation of the four forage soybean varieties at different cutting ages during the rainy season (Figure 2) increased as phenological development progressed, reaching an average of 96 % on day 61 in the four genotypes, which remained until day 68. Subsequently, it decreased until day 89, with an average of 84 %, regardless of the variety. Albina reached 99 % on day 68 ($p < 0.05$), although, on average, it presented the lowest intercepted radiation (77 %). The recommended intercepted radiation is 95 % (Rojas-García *et al*., 2018), as it represents the best ratio between dry matter yield and chemical quality at that point. In this study, that average was reached on day 54.

Giayetto *et al*. (2011) pointed out that intercepted radiation in maize and soybean has a direct influence on plant components and yield. Leguizamón and Verdelli (2011) evaluated intercepted radiation in soybean-maize under intercropping and monoculture systems and obtained a maximum average of 45 %, different from the averages reported in this study, possibly due to differences in plant density and weather conditions. Díaz *et al*. (2003), in a maize-soybean intercrop in Argentina, reported 95 % intercepted radiation on day 54, which coincides with the 94 % registered in this study at the same age. In turn, Ríos-Hilario *et al*. (2023) reported the highest intercepted radiation, 86 %, at a population density of 250 000 plants ha^{-1} in the Salcer variety on day 60 after planting.

Population density

The population density of the four forage soybean genotypes at different physiological ages during the rainy season (Figure 3) showed, on average, a decrease as phenological growth advanced, with similar behavior among the four genotypes, declining from 24

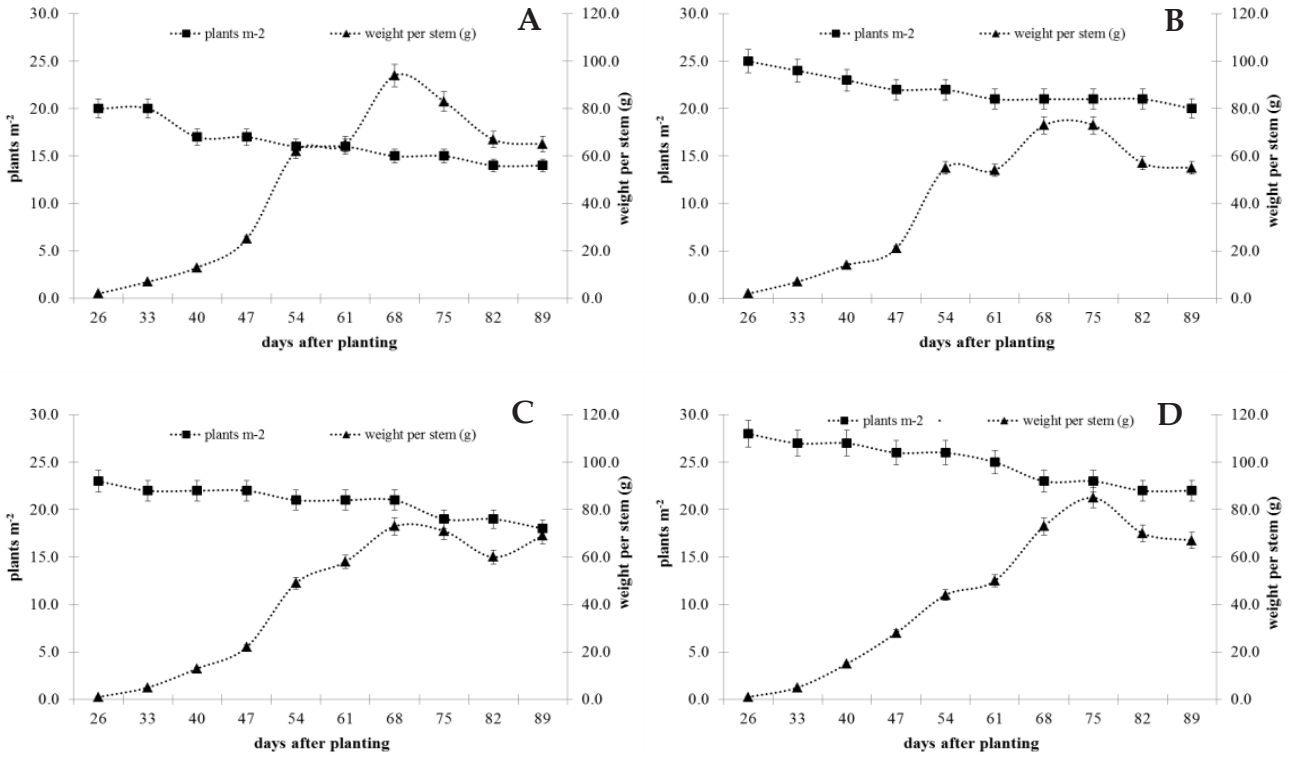

Figure 3. Population density **(p**lants m^{-2}) and weight per stem (g) of different forage soybean (*Glycine max* (L.) Merr.) genotypes in the dry tropics of Mexico. A: Salcer; B: Ojo de tigre; C: Valente; D: Albina.

plants m^{-2} on day 26 to 19 plants m^{-2} on day 89. Albina displayed the highest average density during the cycle (25 plants m^{-2}; $p < 0.05$), whereas Salcer presented the lowest value (16 plants m^{-2}). Ojo de Tigre and Valente registered intermediate and similar values, with 22 and 21 plants m^{-2}, respectively. These differences depend on the adaptive capacity of each genotype to produce under the site's weather conditions, which influence forage yield.

Roche *et al*. (1990) evaluated 74 varieties with 20 plants m^{-2} and reported DM yields of 4360 kg ha^{-1} on day 65 and 7880 kg ha^{-1} on day 78. These values are lower than those obtained in this study, in which 9422 and 8006 kg ha^{-1} of DM were reached with similar densities on days 68 and 75, respectively. Hintz *et al*. (1992) evaluated three varieties with 28 and 89 plants m^{-2} in the United States and obtained similar yields on day 75, with 7500 and 7300 kg ha^{-1} of DM. Tobía and Villalobos (2004) used 19 plants m^{-2} in Costa Rica and reported 4800 kg ha^{-1} of DM on day 61, lower than the values recorded in this study at the same age. Tosquy-Valle *et al*. (2009) evaluated densities of 20, 30, and 40 plants m^{-2} and found better agronomic characteristics with 20 plants m^{-2}. Gaso (2018) evaluated densities from 10 to 60 plants m^{-2} and found a higher final

yield with 30 plants m^{-2}, whereas the highest yield in this study was obtained with 20 plants m^{-2} on day 75.

Stem weight

The stem weight of the four forage soybean genotypes at different cutting ages (Figure 3) increased as plant age advanced, reaching an average of 78 g across the four varieties on days 68 and 75 after planting ($p < 0.05$). Subsequently, it decreased on days 82 and 89, regardless of the genotype, with an average of 64 g. The genotype with the highest average stem weight was Salcer, with 48 g, whereas the lowest was found in Ojo de Tigre, with 41 g. Likewise, Salcer recorded the highest weight on day 68, with 94 g ($p < 0.05$).

Maqueira *et al*. (2016) studied the growth and yield of four soybean varieties and reported stem weights of 9 g on day 89, lower than the values observed in this study. Díaz *et al*. (2003) recorded average weights of up to 77 g on day 90 in Cuba, similar to the 64 g obtained on day 89 in this study. Tobía and Villalobos (2004), under adverse tropical conditions, reported weights of up to 100 g on day 89, which are higher than the values found in this study. Hernández *et al*. (2013) recorded average weights of 82 g on day 89 and pointed out that stem weight later decreases due to the translocation of nutrients to the production of senescent pods and leaves, a behavior that coincides with the results obtained in this study.

Proximal chemical analysis

The crude protein (CP) content of the four forage soybean genotypes in the growth analysis during the rainy season (Table 1) showed that, regardless of the variety, the highest value was recorded on day 26, with 20 %, and the lowest on day 82, with 13 %. On average, Salcer presented the highest content (18 %), whereas Albina had the lowest (16 %) ($p < 0.05$). Hernández *et al*. (2013) reported an average CP content of 15.7 % for leaf, stem, and pod on day 75, similar to the values obtained in this study. Díaz *et al*. (2003) reported 21.7 % CP on day 54 in the whole plant, comparable to the value observed on day 26 in this study. Tobía and Villalobos (2004) reported 20 % CP on day 68. Maldonado-Peralta *et al*. (2023), when evaluating Crotalaria longirostrata Hook. and Arn. at different planting densities, found that CP content decreased as the physiological stage advanced ($p < 0.05$).

On average, the highest neutral detergent fiber (NDF) content was recorded on day 82, with 60 %, and the lowest on day 26, with 52 % ($p < 0.05$) (Table 1). The Valente genotype presented the lowest average NDF content (55 %), whereas Albina and Ojo de Tigre presented the highest values (57 %). Acid detergent fiber (ADF) was 31 % on day 26 and reached its peak on day 54, with 39 %; Albina presented the highest value on day 54 (40 %) ($p < 0.05$).

Tobía and Villalobos (2004) reported an average NDF content of 40 % when evaluating a soybean variety under adverse conditions. Aponte *et al*. (2015) reported 42.6 % NDF on day 82 in 10 forage soybean lines in Puerto Rico, with values lower than those

Table 1. Crude protein and neutral and acid detergent fiber (%) of four forage soybean (*Glycine max* (L.) Merr.) varieties in a growth analysis in the dry tropics of Mexico.

Plant age (days)	Varieties				
	Crude protein (%)				
	Salcer	Ojo de tigre	Valente	Albina	Promedio
26	21 Aa	20 Aa	19 Aa	20 Aa	20 A
40	20 Aa	17 ABa	17 Aa	16 Aa	18 B
54	15 B Cab	16 Ba	15 ABab	11 Bb	14 C
68	16 Ba	15 BCa	18 Aa	18 Aa	17 B
82	13 Ca	11 Ca	12 Ba	11 Ba	13 C
Average	18 a	16 b	16 b	16 b	
	Neutral detergent fiber (%)				
26	55 Ba	53 Ba	53 Ca	56 Ba	55 B
40	53 Ba	55 ABa	54 BCa	54 Ba	54 B
54	60 Aa	60 ABa	59 ABa	61 Aa	60 AB
68	58 ABa	63Aa	57 ABCa	63 ABa	60 AB
82	65 Aa	63 Aa	61 Aa	61 ABa	63 A
Average	58 b	59 a	57 c	59 a	
	Acid detergent fiber (%)				
26	35 Ba	30 Cc	30 Cbc	34 Bab	33 D
40	35 Ba	36 Ba	37 ABa	36 Ba	36 B
54	38 Aa	39 Aa	38 Aa	40 Aa	39 A
68	33 Bb	38 ABa	35 ABab	37 Bab	36 B
82	36 Ba	34 Ba	34 BCa	35 Ba	35 C
Average	35 b	35 b	35 b	36 a	

Means with the same lowercase letter in the same row (a, b, c) and column (A, B, C) are not statistically different (Tukey; $\alpha = 0.05$).

found in this study. Díaz *et al*. (2003) reported 58.3 % NDF in four varieties evaluated in Cuba, similar to the results obtained in this study. Hernández *et al*. (2013) found average NDF contents of 47 % in pods, 38.5 % in leaves, and 66.4 % in stems in 10 forage soybean varieties. In the Comarca Lagunera region of Mexico, average ADF values of up to 38 % have been reported on day 82 (Reta-Sánchez *et al*., 2013), which coincides with the observations made in this study.

CONCLUSIONS

The forage soybean genotype with the best attributes in dry matter yield, structural characteristics, and plant height was Ojo de Tigre. The optimum harvesting time was established on day 61, when an average intercepted radiation of 95 % was obtained and before pod production began in the four genotypes, while maintaining acceptable quality levels. The proximal chemical analysis showed variation depending on the genotype; therefore, further research on forage soybean is necessary.

REFERENCES

Acikgoz E, Sincik M, Oz M, Albayrak S, Wietgrefe G, Turan ZM, Goksoy AT, Bilgili U, Karasu A, Tongel O, Canbolat O. 2007. Forage soybean performance in Mediterranean environments. Field Crops Research 103 (3): 239–247. https://doi.org/10.1016/j.fcr.2007.06.006

AOAC (Association of Official Analytical Chemists). 2005. Official methods of analysis (18th edition). Washington, DC, USA. 1928 p

Aponte A, Valencia-Chin E, Beaver J. 2015. Producción de biomasa y valor nutritivo de líneas de soya forrajera (*Glycine max* L. Merr.) en el noroeste de Puerto Rico. The Journal of Agriculture of the University of Puerto Rico 99 (1): 19–36.

Benítez-Bahena B, Bernal-Hernández A, Cortés-Díaz E, Vera-Castillo G, Carrillo-Anzures F. 2010. Producción de forraje de guaje (*Leucaena* spp.) asociado con zacate (*Brachiaria brizantha*) para ovejas en pastoreo. Revista Mexicana Ciencias Agrícolas 1 (3): 397–411.

CONAGUA (Comisión Nacional del Agua). 2022. Estación 12208, Cuajinicuilapa, Estado Guerrero. Gobierno de México. Comisión Nacional del Agua. Ciudad de México, México. https://www.gob.mx/conagua/acciones-y-programas/guerrero-74888 (Retrieved: March 2026).

Díaz MF, González A, Padilla C, Curbelo F. 2003. Comportamiento de variedades de *Glycine max*, sembradas en junio, en producción de forrajes, forrajes integrales y granos. Revista Cubana de Ciencia Agrícola 37 (1): 59–64.

Enciso-Maldonado G, Sanabria-Velázquez A, Fernández-Riquelme F, Díaz-Nájera J, Fernández-Salinas P, Lugo-Pereira W. 2021. Soybean yield components at different densities and planting seasons in Paraguay. Agronomía Colombiana 39 (1): 12–21. https://doi.org/10.15446/agron.colomb.v39n1.88979

Garay-Martínez JR, Maldonado-Moreno N, Ascencio-Luciano G, Joaquín-Cancino S, Bautista-Martínez Y, Granados-Rivera LD. 2021. Potencial forrajero de líneas experimentales de soya (Glycinemax). Ecosistemas y Recursos Agropecuarios (2): e2932. https://doi.org/10.19136/era.a8nii.2932

Gaso D. 2018. Respuesta del rendimiento de soja a la densidad de siembra en ambientes de productividad contrastante. Agrociencia 22 (2). https://doi.org/10.31285/agro.22.2.7

Giayetto OFE, Guisasola FDM, Balboa GR, Esposito GP. 2011. Intercultivos en franjas de maíz y soja. Efecto de la radiación interceptada. *In* Quinto Congreso de la Soja del Mercosur. Rosario, Argentina. 4 p.

Hernández A, Guerra R, Tobía C, Villalobos E. 2013. Evaluación del potencial forrajero de diez cultivares de soya (*Glycine max* (L.) Merr.) en Venezuela. Agronomía Costarricense 37 (2): 45–54.

Hintz WR, Albrecht AK, Oplinger SE. 1992. Yield and quality of soybean forage as affected by cultivar and management practices. Agronomy Journal 84 (5): 795–798. https://doi.org/10.2134/agronj1992.00021962008400050007x

INEGI (Instituto Nacional de Estadística y Geografía). 2010. Compendio de información geográfica municipal de los Estados Unidos Mexicanos. Cuajinicuilapa, Guerrero. Clave geoestadística 12023. Ciudad de México, México. 10 p.

Jiménez C, Pineda L, Bernardo L, Alejandro M. 2002. Producción de maíz y soya forrajera para ensilaje y venta parcial de la cosecha de elotes o chilotes. Agronomía Mesoamericana 13 (1): 45–48.

Leguizamón ES, Verdelli DV. 2011. Rendimiento de maíz y soja en sistemas de cultivos en franjas y monocultura: efectos de la orientación de la sombra. AgriScientia 28 (2): 147–156. https://doi.org/10.31047/1668.298x.v28.n2.2792

Maldonado-Peralta MÁ, Rojas-García AR, Cristobal-Santiago O. 2023. Yield and chemical quality of chepil (*Crotalaria longirostrata* Hook. and Arn) forage at different seeding densities and cutting frequency. Agrociencia 57 (8). https://doi.org/10.47163/agrociencia.v57i8.2695

Maldonado-Peralta MÁ, Rojas-García AR, Torres-Pacheco HM, Álvarez-Vázquez P, Sabino-López JE, Palemón-Alberto F. 2025. Analysis of the *in situ* digestibility growth of forage soybean varieties. Agroproductividad 18 (12): 23–28. https://doi.org/10.32854/j73nh252

Maqueira LLA, Torres NW, Roján HO, Pérez MS, Toledo D. 2016. Response of the growth and yield of four soybean (*Glycine max*. L.) during the cold season in Los Palacios town. Cultivos tropicales 37 (4): 98–104. https://doi.org/10.13140/RG.2.2.17255.65447

Olivares-Pérez J, Jiménez-Guillén R, Rojas-Hernández S, Martínez-Hernández PA. 2005. Uso de las leguminosas arbustivas en los sistemas de producción animal en el trópico. Revista Electrónica de Veterinaria 6 (5): 1–19.

Pagano MC, Miransari M. 2016. The importance of soybean production worldwide. *In* Miransari M. (ed.), Abiotic and Biotic Stresses in Soybean Production. Academic Press: Washington, DC, USA. https://doi.org/10.1016/b978-0-12-801536-0.00001-3

Pamo ET, Boukila B, Fonteh FA, Tendonkeng F, Kana JR, Nanda AS. 2007. Nutritive value of some grasses and leguminous tree leaves of the Central region of Africa. Animal Feed Science and Technology 135 (3–4): 273–282. https://doi.org/10.1016/j.anifeedsci.2006.07.001

Reta-Sánchez DG, Castellanos-Galván PC, Olague-Ramírez J, Quiroga-Garza HM, Serrato-Corona JS, Gaytan-Mascorro A. 2013. Potencial forrajero de cuatro especies leguminosas en el ciclo de verano en la Comarca Lagunera. Revista Mexicana de Ciencias Agrícolas 4 (5): 659–671.

Ríos-Hilario JJ, Maldonado-Peralta MÁ, Rojas-García AR, Hernández-Castro E, Sabino-López JE, Segura-Pacheco HR. 2023. Comportamiento productivo del cultivo de soya variedad Salcer a diferentes densidades de población y momentos de cosecha. Revista Fitotecnia Mexicana 46 (1): 3–10. https://doi.org/10.35196/rfm.2023.1.3

Roche RSV, Hernández JE, Alonso F. 1990. Evaluación inicial de variedades de soya (*Glycine max* (L.) Merril) para la producción de forraje en la época poco lluviosa. Estación Experimental de Pastos y Forrajes "Indio Hatuey" Matanzas, Cuba. Pastos y Forrajes 13 (2): 127–136.

Rojas-García AR, Maldonado-Peralta MA, Hernández-Guzmán FJ, Cruz-Hernández A, Chay-Canul AJ, Ventura-Ríos J. 2021. Potencial forrajero de guaje (*Leucaena leucocephala*) a diferentes densidades de plantas y edades de rebrote. Ecosistemas y Recursos Agropecuarios 8 (2): e2919. https://doi.org/10.19136/era.a8nII.2919

Rojas-García AR, Maldonado-Peralta MÁ, Ortega-Acosta SÁ, Palemón-Alberto F, Pérez-Hernández H, Ventura-Ríos J. 2024. Dinámica de formación de tallos, rendimiento y análisis bromatológico del pasto Mulato II (*Urochloa* híbrido) en el trópico seco de México. Forrajes Tropicales 12 (1): 1–10. https://doi.org/10.17138/tgft(12)1-10

Rojas-García AR, Maldonado-Peralta MÁ, Torres-Pacheco HM, Aguirre HA, Espinosa-Rodríguez M, Ortega-Acosta SA. 2025. Forage production in forage soybean varieties (*Glycine max* L. Merr.). Agroproductividad 18 (12): 15–21. https://doi.org/10.32854/nyzp9t53

Rojas-García AR, Torres-Salado N, Maldonado-Peralta MÁ, Sánchez-Santillán P, García-Balbuena A, Mendoza-Pedroza SI, Hernández-Garay A. 2018. Curva de crecimiento y calidad del pasto Cobra (*Brachiaria* híbrido BR02/1794) a dos intensidades de corte. Agroproductividad 11 (5): 34–38.

Romero-Arias A, Ruz R, González M. 2013. Evaluation of seven soybean (*Glycine max*) cultivars under the edaphoclimatic conditions of the Majibacoa municipality, Las Tunas. Pastos y Forrajes 36 (4): 459–463.

SAS Institute. 2011. Statistical analysis systems. SAS/STAT. User's guide (Nineth edition). Cary, NC. USA.

Teixeira EJ, Soares MB, Tavanti TR, Palú AER, Souza JE. 2016. Avaliação da produtividade de cultivares de soja em plantio de segunda safra. Revista Científica-Cultural 1 (1): 1–4.

Tobía C, Villalobos E, Rico E. 2006. Uso del forraje de soya (*Glycine max* L.Merr.) Variedad Cigras 06 en la nutrición de los rumiantes. *In* X Seminario de Pastos y Forrajes. Zulia, Venezuela, pp: 77–86.

Tobía C, Villalobos E. 2004. Producción y valor nutricional del forraje de soya en condiciones tropicales adversas. Agronomía Costarricense 28 (1): 17–25.

Tobía C. 2011. Efecto de cinco niveles de suplencia hídrica sobre la biomasa y la composición química del forraje de soya (*Glycine max* L. Merr.) cv. CIGRAS 06 UCR-UCLA. Universidad Centroccidental "Lisandro Alvarado". Barquisimeto, Venezuela. 100 p.

Torres-Salado N, Moctezuma-Villar M, Rojas-García AR, Maldonado-Peralta MÁ, Gómez-Vázquez A, Sánchez-Santillán P. 2020. Comportamiento productivo y calidad de pastos híbridos de *Urochloa* y estrella pastoreados con bovinos. Revista Mexicana de Ciencias Agrícolas 24: 35–46. https://doi.org/10.29312/remexca.v0i24.2356

Tosquy-Valle OH, Esqueda-Esquivel VA, Zetina-Lezama R, Ascencio-Luciano G. 2009. Densidad y distancia de siembra en dos variedades de soya de temporal en Veracruz, México. Agronomía Mesoamericana 21 (1): 63–72.

van Soest, PJ, Roberton JB, Lewis BA. 1991. Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. Journal of Dairy Science 74 (10): 3583–3597. https://doi.org/10.3168/jds.s0022-0302(91)78551-2

Vega C, Andrade F. 2000. Densidad de plantas y espaciamiento entre hileras. *In* Andrade FH, Sadras VO. (eds.), Bases para el Manejo del Maíz, el Girasol y la Soja. Balcarce, Argentina, pp: 69–97.

Ventura-Ríos JM, Santiago-Ortega A, Maldonado-Peralta MÁ, Álvarez-Vázquez P, Maldonado-Peralta R, Barrera-Martínez I, Wilson-García CY. 2021. Biomasa de *Urochloa humidicola* como materia prima para producir biocombustible. Revista Fitotecnia Mexicana 44 (4A): 797–804. https://doi.org/10.35196/rfm.2021.4-A.797

Wilson-Cortez J, Angeli-Furlani CE, da Silva RP, Alandia-Román RA. 2011. Características agronómicas de la soya en función de las densidades de siembra y profundidad de deposición de abono. Revista Ceres 58 (1): 62–68.

Agrociencia

A DATA-DRIVEN APPROACH TO VERTICAL HARVESTING: ARTIFICIAL INTELLIGENCE AND INTERNET OF THINGS-BASED AUTOMATION IN AGRICULTURE

Hemlathadhevi **Annadhurai**[1], Lakshmikanth **Paleti**[2], Ramesh Kumar **Chinnakunnu**[3], Surendran **Rajendran**[4*]

[1]Panimalar Engineering College. Department of Computer Science and Engineering. Poonamallee, Chennai, Tamil Nadu 600123, India.
[2]RVRJC College of Engineering. Department of Computer Science and Business Systems. Chowdavaram, Guntur, Andhra Pradesh 522019, India. lakshmikanthpaleti@gmail.com.
[3]Galgotias University. School of Computer Science and Engineering. Greater Noida, Uttar Pradhesh 203201, India.
[4]Saveetha School of Engineering. Saveetha Institute of Medical and Technical Sciences, Department of Computer Science and Engineering. Chennai, Tamil Nadu 602105, India.
* Author for correspondence: surendranr.sse@saveetha.com

Citation: Annadhurai H, Paleti L, Chinnakunnu RK, Rajendran S. 2026. A data-driven approach to vertical harvesting: artificial intelligence and internet of things-based automation in agriculture. **Agrociencia** 60(4): 480-498. https://doi.org/ 10.47163/agrociencia.v60i4.3555

Editor in Chief:
Dr. Fernando C. Gómez Merino

Received: October 10, 2025.
Approved: May 10, 2026.
Published in Agrociencia: May 26, 2026.

ABSTRACT

Smallholder farmers cultivating limited landholdings often experience low yields and reduced returns, as restricted plot sizes limit crop rotation and expansion. Vertical farming provides a practical alternative by enabling intensive production in compact and densely populated areas, supporting year-round cultivation and increased food output. Artificial Intelligence (AI) and the Internet of Things (IoT) play a central role in modern agriculture by enabling precision farming and data-driven decision-making. The proposed system integrated soil-based, hydroponic, and aeroponic techniques within a unified vertical framework. Advanced irrigation and monitoring technologies, including moisture sensors and image-based crop analysis, optimized water usage, nutrient delivery, and crop health management. The approach automated irrigation and nutrient control and achieved a disease detection accuracy of 96 %, demonstrating improved performance compared to conventional machine learning models. Real-time monitoring through sensors and imaging devices reduced manual intervention, improved resource utilization, and supported sustainable agricultural practices. Overall, the system enhanced productivity across diverse farming conditions while promoting resource efficiency, sustainability, and climate resilience.

Keywords: automatic drip irrigation system, sustainable agriculture, crop growth, disease prediction.

INTRODUCTION

Agriculture involves the cultivation of crops and the rearing of animals to supply food and fiber, which is essential for human needs and economic development. These activities depend on effective management of soil, water, and biodiversity to maintain

productivity. This sector plays a central role in the globalized food system, generating value for rural communities worldwide. However, rising population growth continues to reduce arable land per capita. Developing horticultural areas face labor shortages due to rural-to-urban migration, alongside persistent inefficiencies and economic constraints. At the same time, rapid urbanization, climate variability, environmental contamination, and freshwater depletion intensify pressure on natural resources and global food security (Gürsu, 2024).

Farmers operating on restricted agricultural land face significant limitations in maximizing production efficiency. Constrained space restricts expansion and proper crop rotation, negatively affecting soil productivity, yields, and economic stability. Limited land availability also reduces the feasibility of adopting contemporary farming technologies that require additional space, such as precision agriculture and automated systems, thereby limiting improvements in yield and efficiency. Consequently, smallholders struggle to remain competitive compared to larger-scale producers.

Vertical farming has emerged as an innovative food production method that enhances productivity while reducing resource consumption and environmental impact. By arranging crops in vertically stacked layers, it maximizes land use efficiency and is particularly suitable for urban environments and land-limited regions. Vertical farming supports continuous year-round production, strengthens food security, and reduces transportation-related costs associated with conventional agriculture (Deepa *et al*., 2024a), while addressing land scarcity through controlled environment cultivation.

In practical applications, vertical farming systems operate within controlled environments where crops are grown in stacked layers equipped with advanced monitoring technologies. Sensors positioned across layers continuously measure moisture, nutrient concentration, temperature, humidity, and light exposure. Internet of Things (IoT) platforms transmit these data in real time to centralized processing systems for analysis, while Artificial Intelligence (AI) processes sensor-generated data to optimize irrigation scheduling, regulate nutrient delivery, and detect early signs of disease or stress through image-based monitoring.

The integration of IoT and AI enables precise vertical crop management and data-driven automation (Abraham *et al*., 2025). AI-based predictive analytics combined with continuous sensor monitoring allow early identification of pests and diseases, minimizing crop damage while preserving resources and improving yield performance. AI-supported monitoring incorporates computer vision and image recognition to assess plant growth, detect abnormalities, and implement prompt corrective actions. Automated irrigation systems integrated with AI analytics deliver precise quantities of water and nutrients according to plant developmental stages and environmental conditions, reducing manual labor and enhancing resource efficiency.

Harvesting strategies in vertical farming are designed to maximize efficiency and productivity within limited space. Manual harvesting remains appropriate for delicate or high-value crops requiring careful handling, whereas automated systems

integrating robotics and AI reduce labor costs and increase operational speed. Precision harvesting uses sensors and data analytics to determine optimal harvest timing, ensuring maximum yield and quality. Mechanical harvesting supports large-scale operations, while continuous and selective harvesting techniques maintain steady production by collecting mature crops and allowing immature plants to reach full development. These approaches strengthen production capacity and environmental sustainability in compact growing environments.

Recent studies have advanced vertical farming technologies across multiple dimensions. Erekath *et al*. (2024) explored AI-enabled IoT technology for sustainable environmental management, demonstrating how advanced AI supports precise environmental regulation, early pest prevention, automated operations, and data-driven agricultural decision-making. Xie *et al*. (2024) designed a vertical farming prototype to examine wind energy opportunities in urban environments, conducting parametric analyses of turbine placement along building envelopes and ventilation corridor widths, and investigating the feasibility of integrating micro wind turbines into façade and roof structures. Gürsu (2024) presented a system aimed at improving productivity for small growers working on restricted agricultural areas, including the "PETREE" vertical planting system that converts durable fifth LT-PET packages into plant pots and recycled plastic frames to create mobile sustainable agriculture units.

Strategic and resource-oriented evaluations have also been conducted. Akbari *et al*. (2024) applied the Quantitative Strategic Planning Matrix (QSPM) together with SWOT analysis to assess the feasibility of vertical farming in Gilan Province, Iran. Walia *et al*. (2024) proposed a comprehensive rainwater harvesting system incorporating catchments, mesh filters, gutters, conduits, first-flush devices, and storage tanks, and described *in situ* techniques such as contour farming, strip cropping, vegetative hedges, tied ridging, bunds, terracing, and runoff harvesting for arid and semi-arid regions. Talbot and Monfet (2024) developed a simulation framework for evaluating building performance that concurrently assesses energy demand and crop productivity in compact vertical farming systems, incorporating a dynamic crop model. Fasciolo *et al*. (2024) categorized and analyzed Key Performance Indicators (KPIs) in vertical farming, compiling a structured set of 78 KPIs from the literature.

Technological advancements in automation and robotics have further strengthened vertical farming systems. Fei and Vougioukas (2024) introduced a real-time optimization system to adjust platform travel speed and lift height for improved harvest efficiency, validated through simulations and field trials in a commercial apple orchard. Wang *et al*. (2023) proposed a biomimetic robotic harvesting system for strawberries in vertical cultivation environments, employing deep neural networks and conditional Generative Adversarial Networks to detect ripe fruit using synthetic data. Kabir *et al*. (2023) reviewed developments in sensor technologies, monitoring and control systems, unmanned aerial systems, and the growing influence of AI in optimizing vertical farming operations. Carotti *et al*. (2023) examined water use efficiency in lettuce (*Lactuca sativa* L.) under ebb-and-flow substrate culture and

high-pressure aeroponic systems in an experimental vertical farm at the University of Bologna. Ciriello *et al*. (2023) investigated light quality, quantity, and photoperiod effects for energy-efficient basil cultivation. Vandewetering *et al*. (2023) developed an innovative wood-based photovoltaic racking solution using an open-source hinge system to reduce mechanical pressure.

Advances in data analytics and intelligent modeling also contribute to agricultural innovation. Sivasubramanian *et al*. (2025) proposed an Al-Biruni Earth Radius Optimization-based deep transfer learning framework for remote sensing scene image classification, integrating SqueezeNet for feature extraction and a deep autoencoder neural network optimized through the Al-Biruni strategy to address high dimensionality and limited labeled data. Deepa *et al*. (2024b) introduced Agri-Ontology, an ontology-driven framework for organizing heterogeneous agricultural data using natural language processing techniques, employing BERT for semantic representation, Jaccard similarity for relationship extraction, and a BiGAN-based model for evaluation, achieving 94.64 % accuracy.

While these studies greatly improve vertical farming technologies, many only look at specific parts like energy systems, water management, structural design, robotics, or data representation, without bringing together AI-driven disease detection and real-time IoT-based automation into a complete vertical harvesting system. Some approaches present scalability challenges for small or urban farms, while others involve high operational costs or complex infrastructure. Limited real-time decision-making and delayed disease identification further reduce system effectiveness.

The present study positions vertical farming as a strategic response to contemporary agricultural challenges. The primary aim was to develop an intelligent vertical harvesting system by integrating AI-based crop monitoring with IoT-enabled automation to enhance productivity and resource efficiency. The objectives were to design a vertical farming framework combining soil-based, hydroponic, and aeroponic techniques; to monitor environmental and crop conditions using IoT sensors and imaging devices; to apply machine learning models for early crop disease detection; and to automate irrigation and nutrient delivery based on real-time data. Motivated by increasing land scarcity, climate variability, and the need for sustainable food production that reduces water usage, labor dependency, and crop losses, this work focused on controlled vertical farming environments with sensor-based monitoring, AI-driven disease detection, and automated decision-making for efficient crop management.

MATERIALS AND METHODS

An advanced vertical harvesting system was implemented to combine multiple technologies for enhancing agricultural yield (Figure 1). The system integrates three cultivation methods (aeroponics, hydroponics, and traditional soil-based farming) arranged in vertically stacked tiers. Crops grown under aeroponic conditions were

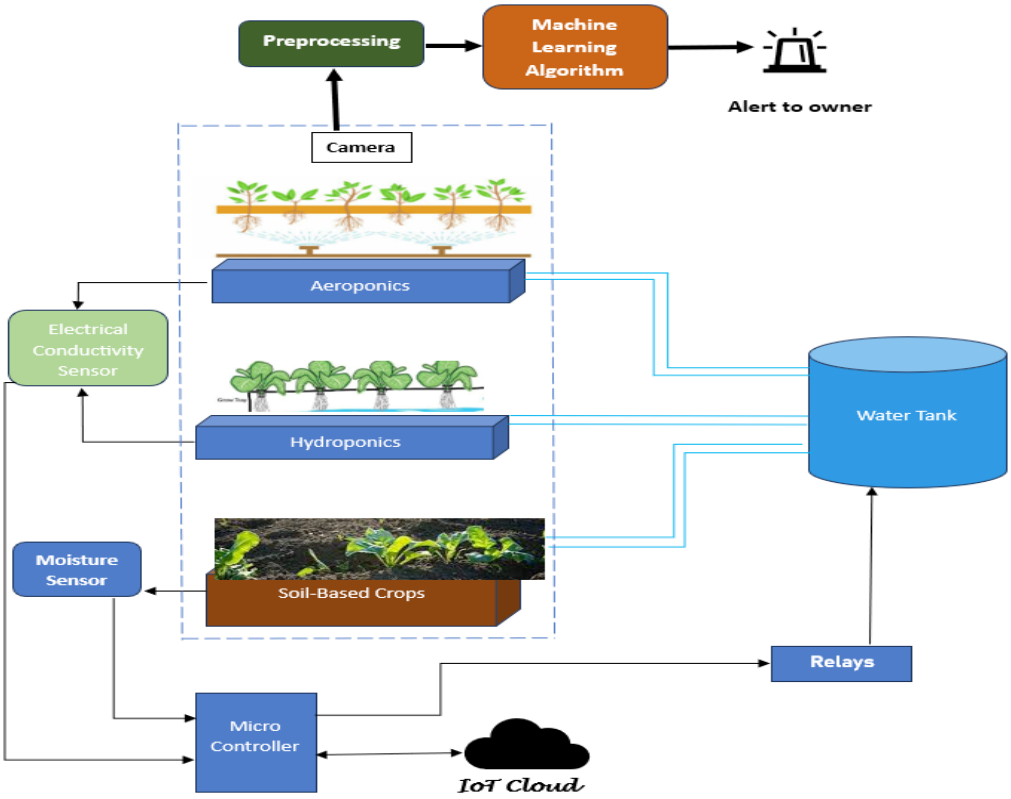

Figure 1. Integrated Artificial Intelligence (AI) and Internet of Things (IoT)-based vertical harvesting system architecture combining aeroponic, hydroponic, and soil-based cultivation.

monitored using video cameras, hydroponic crops were equipped with electrical conductivity sensors to verify nutrient solution adequacy, and soil-based crops utilized moisture sensors to regulate water levels.

The hydroponic and soil-based units were supplied by a central water tank containing nutrient-enriched water, with relays allowing selective distribution according to crop type. Sensor data were collected by a microcontroller that preprocessed the information using machine learning algorithms and generated alerts when anomalies were detected. Mahaveerakannan and Gade (2025) explained that this integration of AI and IoT technologies supports real-time monitoring, management, and optimization across diverse farming systems.

The problem addressed in this work is formulated as a quantitative optimization task aimed at maximizing crop disease detection accuracy and resource utilization efficiency under constrained farming conditions. Let the sensor data vector be defined as

$$S = \{m, t, h, ec\}$$

where *m* denotes soil moisture, *t* represents temperature, *h* indicates humidity, and *ec* corresponds to electrical conductivity. The objective was to maintain these parameters within predefined threshold ranges while minimizing water and nutrient waste.

Soil based crops

Soil-based farming relies on plant production in natural soil systems, where crop performance depends on nutrient availability and soil-borne microorganisms that support healthy growth. Major crops cultivated under this model include rice (*Oryza sativa* L.), maize (*Zea mays* L.), vegetables such as tomato (*Solanum lycopersicum* L.), root crops including potato (*Solanum tuberosum* L.) and carrot (*Daucus carota* L.), and legumes such as soybean (*Glycine max* (L.) Merr.). Rice requires flooded fields and specific soil conditions to achieve maximum yield. Maize depends on fertile, well-drained soils with consistent moisture availability. Tomatoes are sensitive to soil pH and nutrient balance, requiring well-conditioned soils for optimal fruit development. Potatoes and carrots require loose, well-aerated soils to facilitate proper tuber and root formation. Soybeans enhance soil fertility by fixing nitrogen through symbiotic relationships with soil bacteria, benefiting subsequent crops in rotation systems.

Moisture sensors are used for continuous monitoring of soil water content. These sensors collect soil condition data and transmit it to the IoT cloud for storage and analysis. When moisture levels fall below defined thresholds, a relay mechanism activates an automated drip irrigation system to supply the required amount of water. Optimal soil moisture ranges vary by crop: wheat performs best at 50–75 % of field capacity, rice at 80–100 % with periodic flooding, tomatoes at 60–80 %, carrots at 50–70 %, potatoes at 60–80 %, and lettuce at 70–90 % of field capacity. These thresholds guide irrigation management by defining the moisture levels necessary to support optimal crop growth while avoiding over- or under-irrigation.

Hydroponics

Hydroponics is a soil-less cultivation method in which plants receive nutrients from a circulating solution delivered directly to their roots. This system is suitable for urban agriculture and controlled environment systems such as greenhouses. Crops commonly grown using hydroponics include leafy vegetables such as lettuce, spinach (*Spinacia oleracea* L.), and kale (*Brassica oleracea* L.), which require nutrient-rich solutions supplied consistently through the system. Herbs such as basil (*Ocimum basilicum* L.) and cilantro (*Coriandrum sativum* L.) also perform well due to continuous nutrient availability and regulated environmental conditions. Vine crops, including tomato and cucumber (*Cucumis sativus* L.), are frequently cultivated in hydroponic systems, often supported by trellising structures during growth and fruiting. Hydroponics enables precise control over nutrient delivery and environmental conditions, improving production efficiency.

The operational parameters and threshold values for hydroponic crops (Table 1) include essential nutrients such as nitrogen, potassium, calcium, and magnesium, as well as

Table 1. Water pH, electrical conductivity (EC), and nutrient solution ranges used for the selected hydroponic crops.

Crop	Water pH Range	Electric conductivity (mS cm^{-1}) range	Nutrient solution (mg L^{-1})
Basil (*Ocimum basilicum* L.)	5.5–6.5	1.5–2.5	18–22
Tomatoes (*Solanum lycopersicum* L.)	5.5–6.5	2.0–2.5	20–24
Strawberries (*Fragaria* × *ananassa* Duchesne)	5.5–6.5	1.8–2.5	18–24
Cucumbers (*Cucumis sativus* L.)	5.8–6.2	2.0–2.5	22–26
Lettuce (*Lactuca sativa* L.)	5.5–6.5	1.2–2.0	18–22

water quality indicators that influence crop growth and productivity. Monitoring and optimizing these factors according to crop type are necessary to maintain system stability and achieve optimal yield under hydroponic cultivation.

Aeroponics

Aeroponics is a form of hydroponics that does not involve the use of soil or any substrate. Plants are placed in growing pods and intermittently sprayed with a nutrient-saturated mist. This method allows precise control over the amount of oxygen, nutrients, and water vapor supplied to the root system. Aeroponic systems use approximately 80 % less water compared to soil-based farming and even conventional hydroponic systems. One advantage of aeroponics is that nutrients are delivered to the roots as fine moist particles, promoting faster growth and reducing material requirements compared to soil-based cultivation. This approach is particularly suitable in regions with limited arable land or infertile soils. Because there is no soil contact, the risk of soil-borne diseases is minimized, and plant health monitoring is more straightforward. Despite their advanced technology and high initial investment costs, aeroponic systems are sustainable and support year-round production in controlled environments such as greenhouses.

Sensor operation and application

Electrical conductivity (EC) sensors operate by measuring the ability of a solution to conduct electric current, which is directly proportional to the concentration of dissolved ions, particularly salts and nutrients. These sensors consist of electrodes, typically made of stainless steel or graphite, placed in the solution of interest. A small electrical current is passed between the two electrodes, and the resulting conductivity is measured. Higher ion concentrations correspond to increased conductivity of the solution. EC sensors are widely applied in hydroponic systems and soil-based farming to assess nutrient availability. They allow farmers to adjust nutrient solutions or soil amendments to ensure that appropriate levels of essential nutrients are supplied to

crops. Real-time EC monitoring supports precise control of irrigation and fertilization practices, and prevents nutrient imbalances.

Sensors measure parameters and generate raw data. The measurement process involves calibration and conversion procedures to translate raw readings into meaningful units, such as converting voltage outputs from an EC sensor into conductivity values:

$$EC\left(\frac{mS}{cm}\right) = \frac{Voltage_{out}}{K * R}$$

where $Voltage_{out}$ is the sensor output voltage, K is the cell constant of the sensor, and R is the reference resistance.

Similar to moisture sensors connected to the cloud, EC sensors monitor ion concentration in nutrient solutions. When nutrient levels fall below optimal thresholds, the sensors trigger automated drip systems. This function ensures that plants receive appropriate nutrient supplementation, thereby optimizing agricultural productivity and resource efficiency.

Dataset preparation and image preprocessing

Image dataset

A camera was used for real-time plant monitoring, and the captured images were transmitted for preprocessing and subsequent analysis. To train and validate the disease detection model, a publicly available dataset sourced from Kaggle was utilized. This dataset comprises approximately 87 000 Red-Green-Blue (RGB) images of healthy and diseased crop leaves distributed across 38 distinct classes. It was used to train the model to discriminate features associated with different plant diseases and healthy conditions. The dataset was partitioned into training and validation sets in a 70:30 ratio while preserving the original directory structure. In addition, a separate directory containing 33 test images was created to evaluate the model's predictive performance.

Preprocessing

The images were preprocessed to improve the performance and accuracy of the machine learning model. Image dimensions and pixel values were standardized to ensure uniformity across the dataset. Feature representation was enhanced through normalization and augmentation techniques, and noise was reduced to enable clearer data interpretation. This preprocessing procedure improved robustness and generalization capability in the computer vision task.

Image resizing

The images were resized to a uniform spatial resolution using bilinear interpolation to preserve spatial continuity and reduce distortion. Specifically, the bilinear

interpolation equation was applied to compute the intensity value of each resized pixel as a weighted average of its four nearest neighboring pixels:

$$I'(x', y') = (1 - \Delta x), (1 - \Delta y) I(x_1, y_1) + \Delta x (1 - \Delta y) I(x_2, y_1) + (1 - \Delta x) \Delta y I(x_1, y_2) + \Delta x \Delta y I(x_2, y_2)$$

where $I'(x', y')$ is the interpolated pixel value; $I(x_1, y_1)$, $I(x_2, y_1)$, $I(x_2, y_2)$ and $I(x_1, y_2)$are the pixel values of the four nearest neighbors; and Δx Δy correspond to the fractional offsets in the horizontal and vertical directions, respectively.

Normalization

Normalization scales pixel values to a specified range, commonly [0, 1]. This was achieved by directly applying the normalization equation to each pixel in the dataset:

$$I_{norm}(x, y) = \frac{I(x, y) - I_{min}}{I_{max} - I_{min}}$$

where $I(x, y)$ represents original pixel value, and I_{max} and I_{min} represent the maximum and minimum pixel intensities within the image, respectively.

The preprocessed dataset was split into training (70 %), validation (15 %), and testing (15 %) sets. The training set was used to learn model parameters by minimizing the loss function. The validation set was used to tune hyperparameters and monitor overfitting, with early stopping applied when validation performance declined. The testing set was used after training to evaluate the model's generalization on unseen data. This split ensured reliable performance assessment and model robustness.

Machine learning model

Crop disease detection was implemented using You Only Look Once version 6 (YOLOv6). The optimized backbone network and anchor-free design enhanced feature extraction and improved detection of disease symptoms in crops. Model performance was strengthened through training techniques such as mosaic augmentation and label smoothing. In field deployment, YOLOv6 processes images captured by cameras in real time. Enhanced Non-Maximum Suppression (NMS) further improves detection by pruning overlapping bounding boxes. This integrated system supports continuous field monitoring, reduces manual inspection effort, and promotes precise treatment application to improve crop health and yield.

The process for detecting crop diseases using YOLOv6 is carried out in sequential stages. First, a convolutional neural network backbone, such as CSPDarknet53, extracts discriminative features from input images. These features enable identification of disease symptoms, including lesions and discoloration on leaves or stems. YOLOv6 utilizes an anchor-free mechanism that directly predicts bounding box coordinates

relative to reference points within the image. During training, mosaic augmentation and label smoothing improve the model's ability to distinguish between healthy crops and various disease classes. In post-processing, non-maximum suppression eliminates overlapping detections to ensure precise localization of affected regions. When deployed in real time, this integrated pipeline allows immediate intervention, helping farmers protect crops and minimize damage efficiently.

Feature extraction with convolutional neural network (CNN) backbone

YOLOv6 employs a CNN backbone (e.g., CSPDarknet53) to extract feature representations F from input images I:

$$F = CNN_{backbone}(I)$$

The convolutional layers apply filtering operations followed by Rectified Linear Unit (ReLU) activation, batch normalization, and other nonlinear activation functions to generate feature maps that capture patterns indicative of crop diseases.

Anchor-free bounding box prediction

YOLOv6 predicts bounding boxes $B = (b_x, b_y, b_h, b_w)$ directly relative to reference points within the image:

$$b_x, b_y, b_h, b_w = CNN_{head}(F)$$

Here, b_x and b_y represent the center coordinates, while b_h and b_w are the width and height of the bounding box, predicted based on features extracted by the backbone network.

Training with augmentation and label smoothing

During training, mosaic augmentation combines four images into one to improve robustness:

$$I_{mosaic} = MosaicAugmentation(I_1, I_2, I_3, I_4)$$

Label smoothing modifies ground-truth labels Y to reduce overfitting and enhance generalization:

$$Y_{smooth} = LabelSmoothing(Y)$$

Loss function calculation

YOLOv6 minimizes a combined loss function L consisting of bounding box regression, confidence, and classification losses:

Bounding box regression loss (L_{box}),

$$L_{box} = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2]$$

Confidence loss (L_{conf}),

$$L_{conf} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} [\sigma(b_i) - b_i^{true}]^2$$

Classification loss (L_{class})

$$L_{class} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} \sum_{c \in classes} [\sigma(c/b_i) - c^{true}]^2$$

Here, λ_{coord} balances the importance of coordinate predictions, 1^{ij}_{obj} indicates if an object is present in grid cell i and bounding box j, σ represents the sigmoid function, and b_i^{true} and c^{true} are ground-truth values.

Post-processing with non-maximum suppression (NMS)
After prediction, NMS removes redundant overlapping bounding boxes:

$$NMS(B) = \{b \in B | IoU(b, b') < threshold\}$$

where IoU denotes the Intersection over Union between predicted boxes b and b', and the threshold defines the allowable overlap.

Real-time deployment
In real-time operation, YOLOv6 processes input images I from field cameras using trained parameters θ and produces disease detection outputs:

$$Detection_{disease} = YOLOv6(I, \theta)$$

This pipeline enables timely detection and intervention to mitigate crop disease spread and improve agricultural productivity. Following detection, the system sends alerts directly to the farm owner through a dedicated mobile application. When symptoms such as lesions or discoloration are identified, the model processes the captured images in real time and evaluates them against predefined detection thresholds. Upon

confirming disease presence, the application generates and transmits an alert message to the owner's phone. This immediate notification enables prompt intervention, limiting disease spread and protecting crop yield. The integration of automated alerts strengthens responsiveness and supports informed, data-driven agricultural management.

Experimental setup

The framework of the proposed AI-IoT-based vertical harvesting system (Figure 2) utilized an EC sensor to measure nutrient concentrations in aeroponic and hydroponic systems. This sensor was calibrated to detect variations that are critical for optimal plant growth. A moisture sensor was deployed to monitor soil water content in soil-based crops. Both sensors were interfaced with a microcontroller (Arduino Mega 2560), which processed analog inputs and controlled digital outputs. The microcontroller

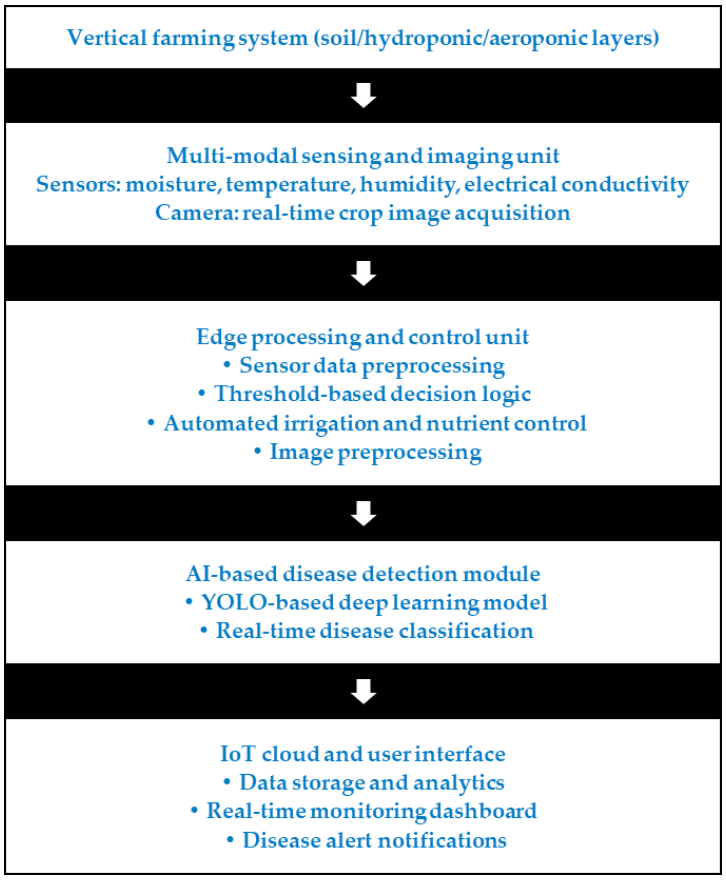

Figure 2. System architecture of the proposed Artificial Intelligence-Internet of Things (AI-IoT)-based vertical harvesting and monitoring system.

transmitted data to the cloud via Amazon Web Services Internet of Things (AWS IoT), where incoming data streams were stored and analyzed in real time. Cloud-based algorithms evaluated predefined thresholds, such as low nutrient levels, and triggered automated alerts when necessary.

A Raspberry Pi high-quality camera module captured real-time images of crop conditions. These images were preprocessed and analyzed using the YOLOv5 machine learning model to enable accurate disease detection and classification. Upon detection of disease symptoms, alerts were sent via email and Short Message Service (SMS) through Amazon Web Services Simple Notification Service (AWS SNS), ensuring timely notification for intervention.

The proposed system was implemented using a multi-layer vertical structure comprising soil-based, hydroponic, and aeroponic units. Moisture sensors were installed within soil layers, while EC sensors were positioned in hydroponic nutrient tanks to continuously monitor nutrient concentration. The camera module was mounted to capture real-time crop images for disease detection. All sensing devices were connected to a central microcontroller responsible for data acquisition, preprocessing, and transmission to the cloud platform. Relay modules interfaced with irrigation pumps and nutrient valves allowed automated control actions based on predefined threshold values and AI-based analysis.

A water tank was constructed to provide a continuous water supply to the crops. The system automatically activated the water tank and irrigation mechanism to deliver the required water when it detected inefficient nutrient levels or inadequate moisture conditions. If disease presence was predicted, the system generated immediate alerts to notify the owner, enabling prompt corrective action.

Performance metrics

These metrics quantitatively evaluate the accuracy and effectiveness of machine learning and image processing models. They measure accuracy, precision, recall, and overall performance by comparing predicted outputs with ground-truth labels and determine whether the model is reliable and suitable for tasks such as crop monitoring and disease prediction.

Accuracy (A): Measures the proportion of correctly classified instances among all instances.

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision (P): Quantifies the correctness of positive predictions.

$$P = \frac{TP}{TP + FP}$$

Recall (R): Measures the ability to correctly identify positive instances.

$$R = \frac{TP}{TP + FN}$$

F1-Score (F1): Represents the harmonic mean of precision and recall, providing a balanced metric.

$$F1 = 2 * \frac{Precision.Recall}{Precision + Recall}$$

Here, *TP* and TN denote True Positive and True Negative, while *FP* and *FN* represent False Positive and False Negative, respectively.

RESULTS AND DISCUSSION

The vertical harvesting system optimized land use through vertically stacked crop layers, increasing production per unit area compared to conventional cultivation. The integration of hydroponic and aeroponic techniques reduced water consumption and nutrient loss while improving growth rates and yield consistency. Continuous real-time monitoring using camera systems and machine learning enabled timely disease detection and intervention. The system also reduced reliance on pesticides and herbicides, lowered transportation-related carbon emissions due to proximity to urban areas, and supported year-round cultivation, contributing to improved food security and environmental sustainability.

The performance of the real-time crop condition classification model was evaluated using a confusion matrix (Figure 3). The results showed difficulty in distinguishing between the four classes: Poor, Average, Good, and Excellent. A substantial proportion of "Excellent" instances (67 %) were misclassified as "Poor," while 57 % of "Poor" instances were incorrectly predicted as "Excellent." The "Good" category achieved 41 % correct classification with notable dispersion into other classes. Similarly, only 27 % of "Average" instances were correctly identified, with frequent misclassification into "Poor" and "Good." These findings indicate that the model particularly struggled with extreme categories, meaning it requires further refinement to improve classification reliability.

Hydroponic nutrient dynamics were analyzed over time (Figure 4), evaluating nutrient concentration (mg L^{-1}) variations for potassium, phosphorus, calcium, magnesium, and nitrogen. The model compared observed nutrient levels with predefined threshold values to regulate automated nutrient supply. Phosphorus and magnesium remained relatively stable over time, while potassium and nitrogen showed slight variations. Calcium fluctuated within shorter intervals, reflecting dynamic nutrient adjustments.

The EC of the hydroponic solution (Figure 5) initially measured 2.6 mS cm^{-1} and

Figure 3. Confusion matrix for real-time crop condition classification.

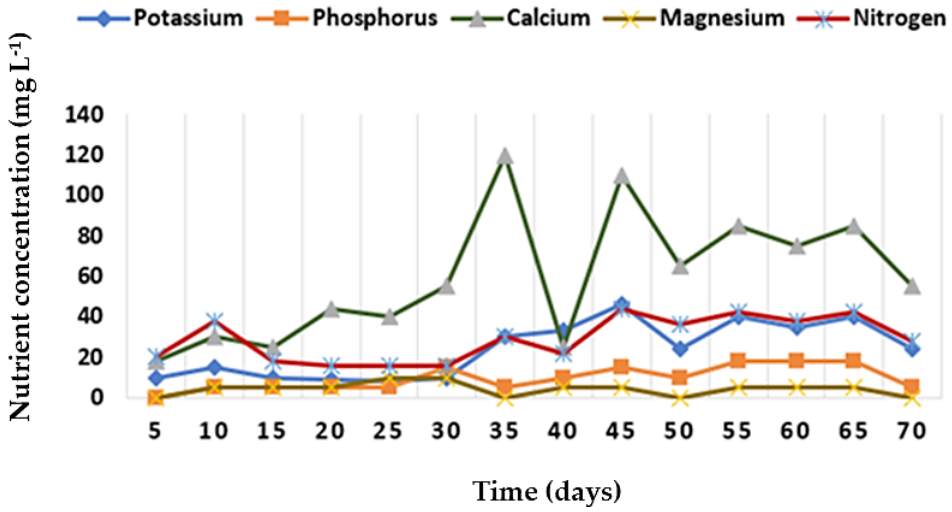

Figure 4. Temporal variation of hydroponic nutrient concentrations.

gradually increased to a peak of 4.5 mS cm^{-1}, indicating higher dissolved salt concentration, and later decreased to 2.9 mS cm^{-1}. The pH variation of the hydroponic solution (Figure 6) through time showed an increase from 5.5 to a maximum of 6.5, and then decreased to 5.6. This pattern indicated a gradual shift from acidic to near-neutral conditions before returning toward slightly acidic levels. Such variations were

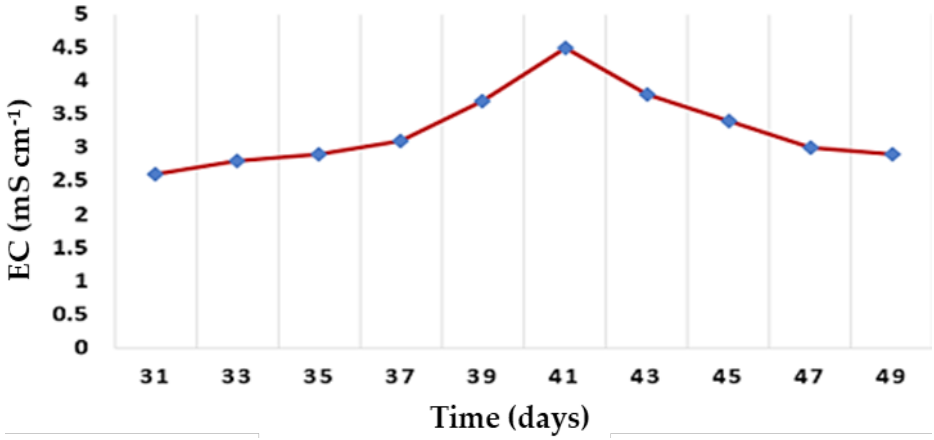

Figure 5. Temporal variation of electrical conductivity (EC) in hydroponic solution.

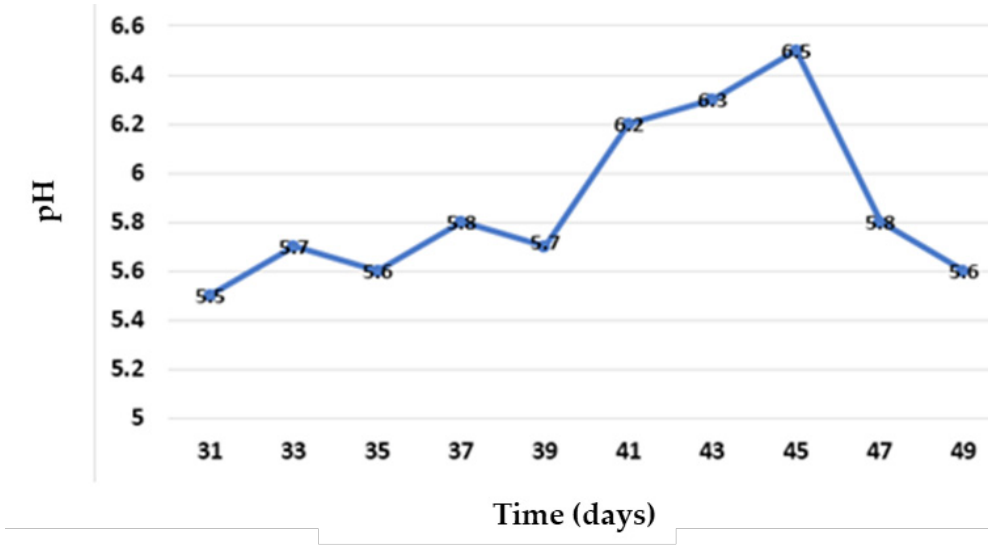

Figure 6. Temporal variation of pH Levels in hydroponic nutrient solution

likely influenced by factors including water source changes and nutrient adjustments. Quantitative evaluation demonstrated strong disease detection performance. The proposed method achieved an accuracy of 96 %, with 95.5 % precision, 96 % recall, and an F1-score of 94 %. These results confirmed balanced precision and recall, indicating robust and reliable performance in real-time crop monitoring.

Comparative analysis with baseline models further validated the effectiveness of the proposed approach. The Convolutional Neural Network (CNN) achieved 94 %

accuracy, the Support Vector Machine (SVM) achieved 95.8%, Extreme Gradient Boosting (XGBoost) achieved 88 %, Long Short-Term Memory (LSTM) achieved 90 %, and Long Short-Term Memory-Recurrent Neural Network (LSTM-RNN) achieved 95 %. The proposed YOLOv5 model outperformed all comparison models with 96 % accuracy. In addition to superior accuracy, it maintained strong precision (95.5 %) and recall (96 %), demonstrating consistent and balanced performance across evaluation metrics.

Sensor-based automation further reduced unnecessary irrigation events by maintaining moisture and EC levels within optimal thresholds, confirming the effectiveness of the quantitative optimization framework. Overall, the integrated system demonstrated reliable monitoring, accurate disease detection, efficient nutrient regulation, and improved agricultural resource management.

CONCLUSION

The proposed system is well suited for controlled vertical farming environments and has the potential to improve productivity under land and resource constraints. Despite its strong performance, certain limitations remain. The evaluation was conducted under controlled environmental conditions, and large-scale deployment may require further calibration to accommodate diverse climatic variations. The framework depends on continuous sensor connectivity and camera availability, which may increase initial setup costs for small-scale farmers. In addition, disease detection performance relies on dataset diversity; rare diseases or previously unseen crop varieties may reduce prediction accuracy.

Future work includes large-scale deployment across varied climatic and operational conditions to evaluate robustness and scalability. Integrating additional sensors, such as CO_2 and light intensity sensors, could further enhance precision farming capabilities. Extensions may also involve advanced deep learning architectures, expanded and more diverse datasets, and the incorporation of predictive analytics, cloud-edge computing, and drone-based monitoring to support fully autonomous and scalable smart farming systems.

ACKNOWLEDGEMENTS

Dataset Source: https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset

REFERENCES

Abraham S, Thomas AJ, Ansleen S, Kumar D, Rishee PLG, Senthil K. 2025. Precision agriculture enhanced by spatio-temporal attention model for real-time crop monitoring and prediction. *In* 2025 7th International Conference on Intelligent Sustainable Systems. Institute of Electrical and Electronics Engineers: Tirunelveli, India, pp: 1123–1130. https://doi.org/10.1109/iciss63372.2025.11076290

Akbari Z, Yari Hesar A, Siamian N, Fürst C, Värnik R, Azadi H. 2024. Feasibility of using vertical farming in northern Iran: A multiple necessity. Journal of Environmental Management 354: 120232. https://doi.org/10.1016/j.jenvman.2024.120232

Carotti L, Pistillo A, Zauli I, Meneghello D, Martin M, Pennisi G, Gianquinto G, Orsini F. 2023. Improving water use efficiency in vertical farming: Effects of growing systems, far-red radiation and planting density on lettuce cultivation. Agricultural Water Management 285: 108365. https://doi.org/10.1016/j.agwat.2023.108365

Ciriello M, Formisano L, Rouphael Y, de Pascale S, Kacira M. 2023. Effects of daily light integral and photoperiod with successive harvests on basil yield, morpho-physiological characteristics, and mineral composition in vertical farming. Scientia Horticulturae 322: 112396. https://doi.org/10.1016/j.scienta.2023.112396

Deepa R, Das M, Thilakavathy P, Prabhu BA, Rajendran S. 2024b. Harvesting knowledge: An innovative ontology framework for agricultural advancement using NPL methodologies. *In* 2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies. Institute of Electrical and Electronics Engineers: Trichirappalli, India. https://doi.org/10.1109/iceeict61591.2024.10718542

Deepa T, Maragathasundari S, Kameswari M, Swedheetha C. 2024a. Optimizing smart agriculture: Integrating queuing theory and IoT for enhanced resource management and efficiency. *In* 2024 International Conference on Communication, Computing, Smart Materials and Devices. Institute of Electrical and Electronics Engineers: Chennai, India. https://doi.org/10.1109/icccsmd63546.2024.11015216

Erekath S, Seidlitz H, Schreiner M, Dreyer C. 2024. Food for future: Exploring cutting-edge technology and practices in vertical farm. Sustainable Cities and Society 106: 105357. https://doi.org/10.1016/j.scs.2024.105357

Fasciolo B, Awouda AMM, Grasso N, Bruno G, Chiabert P, Lombardi F. 2024. An evaluation of research interests in vertical farming through the analysis of KPIs adopted in the literature. Sustainability 16 (4): 1371. https://doi.org/10.3390/su16041371

Fei Z, Vougioukas SG. 2024. A robotic orchard platform increases harvest throughput by controlling worker vertical positioning and platform speed. Computers and Electronics in Agriculture 218: 108735. https://doi.org/10.1016/j.compag.2024.108735

Gürsu H. 2024. Waste-based vertical planting system proposal to increase productivity in sustainable horticulture; "PETREE." Sustainability 16 (8): 3125. https://doi.org/10.3390/su16083125

Kabir MSN, Reza MN, Chowdhury M, Ali M, Samsuzzaman AMR, Lee KY, Chung SO. 2023. Technological trends and engineering issues on vertical farms: A review. Horticulturae 9 (11): 1229. https://doi.org/10.3390/horticulturae9111229

Mahaveerakannan R, Gade NVR. 2025. Smart agriculture comparative study of k-nearest neighbours and Random Forest for real-time crop monitoring and yield prediction. *In* 2025 International Conference on Digital Innovations for Sustainable Solutions. Institute of Electrical and Electronics Engineers: Faridabad, India. https://doi.org/10.1109/icdiss68238.2025.11320737

Sivasubramanian S, Kumar SGV, Thanarajan T, Rajendran S. 2025. Al-Biruni earth radius optimization for enhanced environmental data analysis in remote sensing imagery. Agrociencia 59 (5). https://doi.org/10.47163/agrociencia.v59i5.3380

Talbot MH, Monfet D. 2024. Analysing the influence of growing conditions on both energy load and crop yield of a controlled environment agriculture space. Applied Energy 368: 123406. https://doi.org/10.1016/j.apenergy.2024.123406

Vandewetering N, Hayibo K, Pearce J. 2023. Open-source vertical swinging wood-based solar photovoltaic racking systems. Designs 7 (2): 34. https://doi.org/10.3390/designs7020034

Walia SS, Kaur K, Kaur T. 2024. Water harvesting: Importance, its techniques. *In* Rainfed Agriculture and Watershed Management. Springer: Singapore, pp: 125–137. https://doi.org/10.1007/978-981-99-8425-1_16

Wang F, Urquizo RC, Roberts P, Mohan V, Newenham C, Ivanov A, Dowling R. 2023. Biologically inspired robotic perception-action for soft fruit harvesting in vertical growing environments. Precision Agriculture 24 (3): 1072–1096. https://doi.org/10.1007/s11119-023-10000-4

Xie S, Martínez-Vazquez P, Baniotopoulos C. 2024. Wind flow characteristics on a vertical farm with potential use of energy harvesting. Buildings 14 (5): 1278. https://doi.org/10.3390/buildings14051278

Agrociencia

VEGETATIVE GROWTH AND YIELD OF FIG (*Ficus carica* L.) WITH DIFFERENT PLANT ARCHITECTURE

Antonio **Vázquez-Moisen**[1], Alfonso **Muratalla-Lúa**[2], Alfredo **López-Jiménez**[1*], Sergio Humberto **Chávez-Franco**[1], Javier **Suárez-Espinosa**[3]

[1]Colegio de Postgraduados Campus Montecillo. Posgrado en Recursos Genéticos y Productividad-Fruticultura. Carretera México-Texcoco km 36.5, Montecillo, Texcoco, State of Mexico, Mexico. C. P. 56264.
[2]Colegio de Postgraduados Campus Montecillo. Posgrado en Recursos Genéticos y Productividad-Genética. Carretera México-Texcoco km 36.5, Montecillo, Texcoco, State of Mexico, Mexico. C. P. 56264.
[3]Colegio de Postgraduados, Campus Montecillo. Posgrado en Socioeconomía, Estadística e Informática-Estadística. Carretera México-Texcoco km 36.5, Montecillo, Texcoco, State of Mexico, Mexico. C. P. 56264.
* Author for correspondence: lopezja@colpos.mx

Citation: Vázquez-Moisen A, Muratalla-Lúa A, López-Jiménez A, Chávez-Franco SH, Suárez-Espinosa J. 2026. Vegetative growth and yield of fig (*Ficus carica* L.) with different plant architecture. **Agrociencia** 60(4): 499-510. https://doi.org/ 10.47163/agrociencia.v60i4.3411

Editor in Chief: Dr. Fernando C. Gómez-Merino

Received: November 20, 2025.
Approved: June 20, 2026.
Published in Agrociencia: June 29, 2026.

ABSTRACT

The fig (*Ficus carica* L.) is a crop of high commercial and nutritional value that has motivated research into agronomic practices to improve its production and quality. Pruning is an essential practice to reduce tree size and promote intensive production in protected systems. In this study, the effect of the number of productive branches and pruning at different numbers of nodes on the growth and yield of fig cv. Nezahualcóyotl was evaluated. The experiment was conducted in a tunnel greenhouse at the Postgraduate College (19° 27′ 34′′ N and 98° 54′ 12′′ W, at an altitude of 2247 m). Plants were established from the first year in 20 L pots with a substrate of tepetate, pine needles, and forest soil (1:1:1 v/v). A generalized randomized complete block design was used with three blocks (4, 6, and 8 branches per plant) and four pruning treatments (4, 6, and 8 nodes and no pruning), with five replicates per treatment. Vegetative growth and yield variables were evaluated, and means were compared using Tukey's test ($p \leq 0.05$) in RStudio. Pruning every eight nodes with four branches per plant resulted in the greatest branch length (75.9 cm) and internode length (2.6 cm). Pruning every four nodes with the same number of branches promoted a greater number of lateral shoots (23 shoots), while pruning every four nodes and the treatment without pruning increased the total number of nodes per branch (34 nodes). The combination of six branches pruned every eight nodes resulted in the highest number of fruits per plant (83 fruits). No significant differences were observed in average fruit weight or total yield per plant.

Keywords: Moraceae, Ficeae, pruning, intensive production, plant architecture.

INTRODUCTION

Fig (*Ficus carica* L.) is one of the first plants domesticated by humans, originating in the Mediterranean region of Turkey and propagated worldwide since the 16th century. Its fruit, consumed fresh, dried, or processed, is highly valued for its nutritional content and health benefits, including its association with longevity and food security (Kiralan *et al*., 2023). Global fig production has increased due to its high commercial potential, ease of handling, and growing demand for the benefits it offers (Abbas *et al*., 2019). In Mexico, its cultivation has contributed to rural economic development, although it faces challenges such as pests, diseases, and adverse climatic factors, which have driven the adoption of protected production systems such as greenhouses (Mendoza-Castillo *et al*., 2017).

Intensive greenhouse production allows precise control of key factors such as planting density, substrate management, and pruning, resulting in higher yields compared to open-field systems. However, inadequate practices in fertilization, water management, or vegetative growth control can significantly reduce productivity (Nosir, 2023). Agronomic management, including pruning, is crucial for adapting plant architecture to modern production systems, which aim to optimize fruit production for fresh consumption. Practices such as branch pruning, the use of growth regulators, and the generation of lateral shoots are essential to maximize productivity and facilitate cultural work (Ferrara and Mazzeo, 2023).

Leaves develop at the nodes of fig tree branches, and vegetative and floral buds originate in their axils. These buds give rise to the fruit, which may correspond to the current season (figs) or the previous season (early figs). For this reason, the number of nodes on a branch is directly related to the number of fruits and yield per tree. This is especially important in intensive pruning systems, which renew the canopy of fig trees annually. These systems require adequate branch growth to ensure a sufficient number of nodes and reproductive buds (Gonçalves *et al*., 2006). The common fig has a natural tendency toward vertical growth, making it necessary to implement specific pruning and management practices to limit upward growth and encourage the development of lateral shoots to increase fruit production (Ferrara and Mazzeo, 2023). Based on their tolerance to pruning, fig cultivars can be divided into two categories: those that do not tolerate heavy annual pruning and those that can be intensively pruned while maintaining high yields. Cultivars in the first group tend to branch freely, producing a large number of lateral shoots. However, excessive pruning generates vigorous but unproductive shoots, significantly reducing plant yield. In contrast, as early fig crops are insignificant and produce vigorous fruit, the previous season's wood is unnecessary in cultivars of the second group (Nosir, 2023). Therefore, the objective of this study was to evaluate the effect of the number of branches per plant and pruning according to the number of nodes under protected conditions in order to optimize vegetative growth and fig yield.

MATERIALS AND METHODS

Location

This study was conducted from September 2022 to December 2023 at the "San José" experimental orchard, in the Postgraduate College Montecillo Campus, in the municipality of Texcoco, State of Mexico, Mexico (19° 27′ 34′′ N and 98° 54′ 12′′ W, at an altitude of 2247 m). The experiment was carried out in a 260 m^2 tunnel greenhouse, covered with aphid-proof netting and 20 % shading.

Plant material

Five-year-old fig (*Ficus carica* L.) cv. Nezahualcóyotl plants were used, obtained from a collection established in 1982 by the Fruit Growing Center of the Postgraduate College. This variety is characterized by its environmental plasticity, high yield, and ease of management under intensive greenhouse conditions. The fruits are pear-shaped, with a large ostiole, red pulp, a neck, and a long peduncle. Under controlled conditions, they can reach yields of up to 120 Mg ha^{-1} (García-Ruiz *et al*., 2013). The plants were grown in 20 L pots using a substrate composed of tepetate, pine needles, and forest soil in a 1:1:1 (v/v) ratio.

The physical and chemical characteristics of the substrate were: bulk density of 1.2 g cm^{-3}, sandy loam texture (55.5 % sand, 34.9 % silt, and 9.6 % clay), pH of 5.5, electrical conductivity of 3.23 dS m^{-1}, organic matter content of 6.69 %, total nitrogen (N) of 0.31 %, and available phosphorus (P) of 554.3 mg kg^{-1}. Exchangeable bases, expressed in cmol(+) kg^{-1}, include calcium (Ca) at 5.64, magnesium (Mg) at 1.99, potassium (K) at 0.35, and sodium (Na) at 0.23. No carbonates were present.

Treatments and experimental design

Four pruning treatments were evaluated on the main branches: T1, pruning at four nodes; T2, pruning at six nodes; T3, pruning at eight nodes; and T4, no pruning. The experimental design was a generalized randomized complete block design, with three blocks defined according to the number of branches per plant: B1 (four branches), B2 (six branches), and B3 (eight branches). Each treatment had five replicates, resulting in a total of 60 experimental units, with one plant per experimental unit.

After pruning, the apical shoots were allowed to grow until they progressively reached the number of nodes corresponding to each treatment, while the lateral shoots were limited to a maximum growth of five nodes (Figure 1).

Crop management

The experiment began on September 26, 2022, with pruning at 20 cm above ground level, leaving only the main stem, from which the required branches emerged. Fertilization was carried out manually in three stages. During vegetative growth (December 2022 to March 2023), 10 g of Ultrasol-Growth (SQM, Mexico) was applied per plant every 30 d. During the growth and fruiting stage (April to June 2023), 5 g of

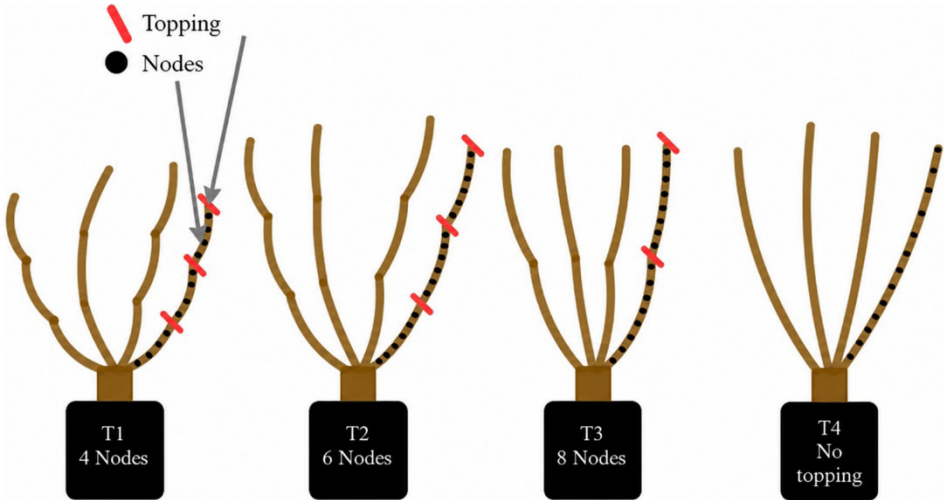

Figure 1. Pruning of branches of fig (*Ficus carica* L.) cv. Nezahualcóyotl. The distance between prunings was defined by the number of nodes (T1, T2, T3) and the control (T4). Pruning was done on plants with 4, 6, and 8 production branches.

Ultrasol-Growth and 10 g of Ultrasol-Production were applied per plant, while during the final fruiting stage (July to December 2023), 10 g of each fertilizer was applied per plant every 30 d.

Irrigation was applied using a drip system with 8 L h^{-1} pressure-compensating emitters connected to a Hunter X-Core XC-200i four-stake timer distributor (Hunter Industries Inc., USA). The water volume applied was 2.7 L d^{-1} per plant from December 2022 to March 2023 and 3.2 L d^{-1} per plant from April to December 2023. Harvesting was carried out twice per week from April to December 2023. Fruits were harvested when more than two-thirds of the surface had darkened and the fruit had softened.

Variables evaluated

Several variables related to vegetative growth and plant yield were evaluated. Branch length was measured using a measuring tape (Truper, Mexico), and the average was calculated according to the number of branches per plant; data were expressed in centimeters. The number of nodes was counted on each branch and averaged across all branches per plant, while internode length was obtained by dividing branch length by the number of nodes. Lateral shoots were recorded at the end of the experiment, and their number was averaged across the branches of each plant.

The number of fruits was determined by adding the fruits recorded at each harvest to obtain the total per plant. Fruit weight was measured individually in freshly harvested fruits using a digital scale, and the data were expressed in grams. Finally, yield per plant was calculated by multiplying the total number of fruits by their average weight.

The cumulative values at the end of the experiment were expressed in kilograms per plant.

Statistical analysis

Statistical analysis was performed using RStudio software version 4.2.1. The assumptions of homogeneity of variances (Bartlett's test at 5 %) and normality (Shapiro-Wilk test at 5 %) were verified. Differences among treatments were determined using Tukey's test ($p \leq 0.05$).

RESULTS AND DISCUSSION

Vegetative growth

Branch length

The yield of figs managed in intensive orchards is closely linked to branch growth, since the fruit develops at the nodes, specifically in the leaf axils (Gaaliche *et al*., 2011). In this study, pruning at different numbers of nodes and branch numbers per plant showed significant differences ($p \leq 0.05$) in final branch length. Plants with four branches and pruning at eight nodes achieved the greatest branch length, with an average value of 75.9 cm, similar to that produced by the treatment without pruning (Table 1). These data suggest a compensatory growth response to intermediate pruning severity.

Table 1. Vegetative growth of fig (*Ficus carica* L.) cv. Nezahualcóyotl (means ± standard deviation) in response to the number of branches in the plant and the number of nodes at the tip of the branches.

Combination		Branch length (cm)†	Nodes at the branch	Internode length (cm)†	Lateral shoots†
Branches on the plant	Nodes at the tip				
Four branches	Four nodes	54.8 bcd ± 6.7	22 c ± 2.9	2.5 ab ± 0.3	23 a ± 6.4
	Six nodes	51.4 bcde ± 5.9	23 bc ± 2.0	2.3 abc ± 0.2	18 ab ± 3.1
	Eight nodes	75.9 a ± 16.6	29 ab ± 3.6	2.6 a ± 0.4	14 ab ± 4.7
	No pruning	66.2 ab ± 17.2	34 a ± 5.3	1.9 bcd ± 0.2	2 c ± 2.0
Six branches	Four nodes	37.1 de ± 5.1	16 d ± 1.9	2.3 abc ± 0.1	16 ab ± 6.1
	Six nodes	45.4 bcde ± 9.6	20 cd ± 1.7	2.3 ab ± 0.4	18 ab ± 5.8
	Eight nodes	60.2 abc ± 16.2	24 bc ± 3.7	2.5 a ± 0.4	12 b ± 4.2
	No pruning	43.3 cde ± 3.8	28 ab ± 2.4	1.5 d ± 0.1	2 c ± 2.4
Eight branches	Four nodes	35.0 de ± 5.0	16 d ± 0.9	2.3 abc ± 0.3	17 ab ± 4.7
	Six nodes	33.4 e ± 5.0	16 d ± 1.4	2.1 abcd ± 0.3	12 b ± 3.6
	Eight nodes	43.1 cde ± 5.9	20 cd ± 1.3	2.2 abc ± 0.3	16 ab ± 1.9
	No pruning	40.0 cde ± 5.0	23 bc ± 1.8	1.7 cd ± 0.2	2 c ± 1.5

†Average values per column with a different letter are different ($p \leq 0.05$).

Pruning temporarily reduces the apical dominance that the apex exerts over lateral buds, reconfiguring the hormonal signaling network and critically altering the source-sink relationship immediately after cutting (Beveridge, 2023; Barbier *et al*., 2019; Madhumala *et al*., 2024). Pruning at eight nodes likely preserved enough leaf area and vascular connectivity to sustain carbon supply and facilitate the rapid replacement of the apex by a distal bud, thus avoiding any net penalty on elongation. This outcome contrasts with more intensive pruning or tree architectures that feature greater competition among apices. Gaaliche *et al*. (2011) reported an annual growth of 10 cm in fig branches grown under rainfed conditions, which was lower than the growth observed in this study.

Annual canopy renewal, achieved by removing the early fig crop, increases shoot length to more than 139 cm without affecting the main crop yield (Micheloud *et al*., 2018). The open-field cultivar 'Roxo de Valinhos', subjected to a severe pruning system (one fruit per node), produced branch lengths of up to 189 cm per year (Gonçalves *et al*., 2006). These differences may be attributed to environmental conditions and management practices, such as the use of intensive systems and cultivar genotype. However, in figs, this type of pruning does not affect fruit production, because the apical meristem continues to generate leaves and inflorescences along its axis (Gaaliche *et al*., 2011), as observed in this study under greenhouse conditions. Furthermore, the sequential growth of shoots prolongs the harvest period, requiring management practices such as irrigation, fertilization, and pruning to maintain continuous growth and high productivity (Leonel and Tecchio, 2010).

Nodes in the branch

The effect of the treatment combinations was significant ($p \leq 0.05$) on the number of nodes per branch. The combination of four branches on unpruned plants showed the highest average number of nodes, with 34 nodes per branch (Table 1). Under temperate climate conditions in Argentina and with intensive pruning of 'Brown Turkey' and 'Guarinta' figs, Micheloud *et al*. (2018) observed branches with average lengths of 151.6 and 139.6 cm and 34.2 and 28.8 nodes per branch, respectively, which is consistent with this study. However, these authors reported longer branches with a lower percentage of fruit-bearing nodes (49.9 and 45 %). This suggests that not all available nodes become reproductive sites, possibly due to factors such as competition for resources or genetic limitations. In contrast, Gaaliche *et al*. (2011) reported, under Mediterranean climate conditions and with light pruning, percentages of fruit-bearing nodes ranging from 50.2 to 88.5 % in five cultivars, although with lower branch growth (14.25 cm).

Internode length

The effect of blocks (number of branches per plant) and treatments (number of nodes at pruning) showed significant differences ($p \leq 0.05$) in internode length. Plants with 4 and 6 branches, combined with pruning at eight nodes, exhibited greater average

internode lengths (2.6 and 2.5 cm, respectively) (Table 1). These values were lower than those reported by Micheloud *et al*. (2018) and Garza-Alonso *et al*. (2019), who recorded internode lengths ranging from 3.34 to 4.85 cm in different cultivars under protected conditions.

In contrast, Kumar *et al*. (2014) observed an increase in internode length from 2.89 to 3.07 cm in 'Deanna' fig when the pruning level was reduced. These results may be attributed to the fact that moderate pruning preserved sufficient early leaf area and source capacity, while a multi-axis architecture increased canopy complexity, thereby modulating both the distribution of photoassimilates and the quality of light perceived by the expanding internodes. The elongation of each internode depends on the coordinated activity of cell division and elongation, processes strongly regulated by gibberellins and brassinosteroids, with auxin playing a modulating role under canopy shading conditions (Li *et al*., 2024).

Lateral shoots

The number of lateral shoots per branch differed significantly ($p \leq 0.05$) depending on the treatment combination. The treatment with four branches and pruning at four nodes showed the highest number of lateral shoots (Table 1). These results are consistent with those reported by Ateyyeh and Sadder (2006), who observed lateral shoot growth ranging from 1.9 to 6.3 shoots per branch in a traditional fig tree pruning and training system. Furthermore, these authors noted that most lateral shoots emerged from the distal nodes of the branches, whereas the basal nodes exhibited limited lateral development.

This finding is consistent with the results obtained in the present study, in which all topped branches generated lateral shoots predominantly at distal nodes. These results may be attributed to a stronger release of apical dominance and a rapid redistribution of the source-sink balance toward axillary buds near the cut point. Pruning eliminates a highly demanding apex and reduces the inhibitory control associated with the dominant axis, thereby allowing bud activation and transition to growth (Beveridge, 2023; Barbier *et al*., 2019). This process does not depend solely on a gradual decline in auxin within the stem, since early increases in promotive signals within the bud may occur after pruning, accompanied by a rapid reconfiguration of metabolites and carbon signaling (Cao *et al*., 2023).

Yield

Fruits per plant

During the warmer months, figs are harvested daily, but as temperatures decrease, fruit ripening slows and harvesting is reduced to once per week (Gariglio *et al*., 2014). Since figs ripen sequentially with shoot growth, greater vegetative development prolongs the harvest period, and the number of fruits per plant depends directly on branch length and the number of nodes per branch (Gaaliche *et al*., 2011). To maintain

stable and high productivity in fig trees, it is essential to promote shoot development through appropriate cultural practices such as irrigation, fertilization, and pruning (Micheloud *et al*., 2023).

Annual intensive pruning, in addition to reducing plant size and allowing higher planting density, promotes increased branch length, thereby increasing the number of fruits per plant and extending the harvest period. In this study, the number of productive branches per plant and the number of nodes at pruning had a significant effect ($p \leq 0.05$) on the number of fruits per plant. The combination of six branches per plant with pruning at eight nodes resulted in the highest number of fruits per plant, with an average of 83 fruits, exceeding the value obtained with six branches and pruning at four nodes (43 fruits) (Table 2). These values were lower than those reported by Nienow *et al*. (2006) for the 'Roxo de Valinhos' fig, who recorded 46 fruits on plants with four branches, 50 fruits on six branches, 95 fruits on eight branches, and 113 fruits on 12 branches.

Under open-field conditions with supplemental irrigation in Argentina, Micheloud *et al*. (2018) reported values closer to those obtained in this study, with averages of 15.4 and 14.2 fruits per shoot in the fig varieties 'Brown Turkey' and 'Guarinta,' respectively. Similarly, Kumar *et al*. (2014) observed that, in the fig variety 'Deanna' under different spacings, plants pruned to eight nodes per branch produced an average of 85 fruits per plant, whereas those pruned to six nodes yielded only 68 fruits. These results are consistent with those of the present study, in which pruning management and branch number were key factors in maximizing fruit production.

Table 2. Yield of fig (*Ficus carica* L.) cv. Nezahualcóyotl (means ± standard deviation) in response to the number of branches on the plant and the number of nodes at the tip of the branches.

Combination		Fruits per plant[†]	Fruit weight (g)	Yield (kg per plant)
Branches on the plant	Nodes at the tip			
Four branches	Four nodes	48 ab ± 16.6	44.7 a ± 3.3	2.15 a ± 0.8
	Six nodes	55 ab ± 11.7	43.2 a ± 6.4	2.42 a ± 0.7
	Eight nodes	81 ab ± 20.3	42.2 a ± 3.5	3.40 a ± 0.9
	No pruning	71 ab ± 26.8	47.3 a ± 8.1	3.47 a ± 1.9
Six branches	Four nodes	43 b ± 12.6	43.5 a ± 5.5	1.86 a ± 0.5
	Six nodes	53 ab ± 18.0	44.2 a ± 2.8	2.36 a ± 0.9
	Eight nodes	83 a ± 25.3	48.9 a ± 5.1	4.14 a ± 1.7
	No pruning	66 ab ± 9.0	44.4 a ± 6.1	2.95 a ± 0.8
Eight branches	Four nodes	57 ab ± 19.7	44.1 a ± 3.1	2.52 a ± 0.9
	Six nodes	58 ab ± 15.2	45.3 a ± 5.0	2.67 a ± 0.9
	Eight nodes	71 ab ± 17.4	43.6 a ± 2.9	3.13 a ± 1.0
	No pruning	62 ab ± 9.7	46.7 a ± 6.6	2.93 a ± 0.8

[†]Mean values per column with a different letter are different ($p \leq 0.05$), means ± standard deviation.

Because axillary buds can rapidly differentiate into reproductive structures, the number of active nodes acts as an immediate structural limit to the productive potential of the shoot (Singh, 2023). Therefore, more nodes per branch increased the number of available fruiting sites and favored the continuity of vegetative-reproductive development, which depends on the supply of photoassimilates. In contrast, fewer nodes per branch, despite stimulating bud emergence through reduced apical dominance, tend to prioritize vegetative regrowth. This reallocation of carbon and nutrients toward new growth may reduce the proportion of reproductive structures.

Fruit weight

Fruit weight is one of the most important characteristics in the fresh fruit market, as it directly influences market prices. In addition, it plays a key role in fruit packaging and transport (Pérez-Sánchez *et al*., 2016). Smaller fruits are typically destined for processing and packaging, whereas larger fruits are preferred for fresh consumption. In this study, the interaction between the number of branches per plant and the number of nodes at pruning had no effect ($p \leq 0.05$) on average fruit weight. All treatment combinations ranged from 42.2 to 48.9 g (Table 2). These results are consistent with those reported by Mendoza-Castillo *et al*. (2017), who worked with plants of the same cultivar having between 3 and 8 branches and likewise found no significant differences in fresh fruit weight, although they obtained slightly higher average values (52.37 to 58.45 g).

On the other hand, Kumar *et al*. (2014) evaluated 'Deanna' figs with plants containing 4, 6, and 8 branches and reported average fruit weights ranging from 42.6 to 65.7 g. Final fruit weight depends on the coordination between cell division and cell expansion, processes regulated by hormonal networks and influenced by environmental conditions and management practices (Zhao *et al*., 2021). In perennial fruit trees, changes in canopy architecture are generally reflected more in fruit load than in consistent reductions in individual fruit weight (Rossouw *et al*., 2024). Likewise, in intensive systems, fruit weight responds more strongly to cultivar, season, and fruit load per shoot than to the training system itself because of local source-sink restrictions (Galán *et al*., 2025).

Yield per plant

The yield of fruit trees is critically influenced by the balance between vegetative growth and fruit production. In fig cultivation, this balance is especially important, as fig and early fig fruits develop during the continuous growth of annual shoots (Gaaliche *et al*., 2011). Annual canopy renewal through intensive pruning is an effective strategy to promote shoot development, extend the harvest period, and increase annual yield. Currently, intensive greenhouse production systems that combine high planting densities, the use of containers and substrates, management of the number of branches per plant, and severe pruning can increase yields by up to 20-fold compared with open-field plantations (Mendoza-Castillo *et al*., 2017).

In this study, the combinations of branches per plant and nodes at pruning showed no significant differences ($p \leq 0.05$) in yield. However, the combination of six branches per plant with pruning at eight nodes resulted in a higher annual yield than the other combinations, reaching 4.14 kg per plant (Table 2). This is similar to the results reported by Kumar *et al*. (2014), who observed yields of 4.04 kg per plant with pruning at eight nodes and 3.29 kg per plant with pruning at six nodes in the 'Deanna' fig. In contrast, Nienow *et al*. (2006), working in Brazil with 'Roxo de Valinhos' figs at a spacing of 1.5 × 1.9 m and plants with 8 to 12 branches, reported considerably higher average yields, ranging from 11.34 to 11.65 kg per plant. Conversely, the results of the present study were higher than those reported by Leonel and Tecchio (2010), who found yields of 1.13 kg per plant without irrigation and 2.5 kg per plant with irrigation. Shoot length is closely related to the number of nodes, reproductive buds, and yield per plant (Gonçalves *et al*., 2006; Gaaliche *et al*., 2011).

Arrangements with fewer nodes tend to redirect resources toward vegetative reconstruction after apical removal, which may reduce the proportion of reproductive structures that develop into harvestable fruit, particularly when light interception and canopy microenvironment become determining factors (Singh *et al*., 2025). Therefore, the integration of appropriate pruning practices, together with modern technologies and proper agronomic management, is fundamental to optimize fig productivity in intensive systems.

CONCLUSIONS

Managing the number of branches per plant and pruning at different numbers of nodes proved to be an effective strategy for improving vegetative growth and yield of the 'Nezahualcóyotl' fig under protected conditions. Pruning at eight nodes with four branches maximized key vegetative parameters such as branch length and internode length, while pruning at four nodes with four branches favored lateral shoot development. The combination of six branches and pruning at eight nodes was the most efficient in terms of yield, achieving the highest number of fruits and the greatest yield per plant.

ACKNOWLEDGEMENTS

We thank the Secretariat of Science, Humanities, Technology and Innovation (SECIHTI) for the scholarship awarded to the first author during his academic training, at the Postgraduate College, Montecillo Campus, State of Mexico, Mexico.

REFERENCES

Abbas M, Hussain S, Ahmad M, Junaid M, Ali Y, Alamgir M, Jalal A. 2019. Testing the performance of fig (*Ficus carica*) cultivars planted in high-density system under agro-climatic condition of Islamabad. Scholars Bulletin 5 (2): 37–41.

Ateyyeh AF, Sadder MT. 2006. Preliminary study on the vegetative and reproductive growth of six common fig (*Ficus carica* L.) cultivars in Jordan. Jordan Journal of Agricultural Sciences 2 (1).

Barbier F, Dun E, Kerr S, Chabikwa T, Beveridge C. 2019. An update on the signals controlling shoot branching. Trends in Plant Science 24 (3): 220–236. https://doi.org/10.1016/j.tplants.2018.12.001

Beveridge C, Rameau C, Wijerathna-Yapa A. 2023. Lessons from a century of apical dominance research, Journal of Experimental Botany 74 (14): 3903–3922. https://doi.org/10.1093/jxb/erad137

Cao D, Chabikwa T, Barbier F, Dun EA, Fichtner F, Dong L, Kerr SC, Beveridge CA. 2023. Auxin-independent effects of apical dominance induce changes in phytohormones correlated with bud outgrowth, Plant Physiology 192 (2): 1420–1434. https://doi.org/10.1093/plphys/kiad034

Ferrara G, Mazzeo A. 2023. Bud structure and evolution. *In* Ramadan MF. (ed.), Fig (*Ficus carica*): Production, Processing, and Properties. Springer Nature: Cham, Switzerland, pp: 109–127. https://doi.org/10.1007/978-3-031-16493-4_5

Gaaliche B, Lauri P, Trad M, Costes E, Messaoud M. 2011. Interactions between vegetative and generative growth and between crop generations in fig tree (*Ficus carica* L.). Scientia Horticulturae 131: 22–28. https://doi.org/10.1016/j.scienta.2011.09.022

Galán AJ, Domínguez MG, Pérez-López M, Galván AI, Pérez-Gragera F, López-Corrales M. 2025. Agronomic performance and fruit quality of fresh fig varieties trained in espaliers under a high planting density. Horticulturae 11 (7): 750. https://doi.org/10.3390/horticulturae11070750

García-Ruiz MT, Mendoza-Castillo VM, Valadez-Moctezuma E, Muratalla-Lúa A. 2013. Initial assessment of natural diversity in Mexican fig landraces. Genetics and Molecular Research 12 (2): 3931–3943. https://doi.org/10.4238/2013.september.23.12

Gariglio NF, Favaro JC, Forte R. 2014. Higuera. *In* Gariglio NF, Bouzo CA, Travadelo MR. (eds.), Cultivos Frutales y Ornamentales para Zonas Templado-Cálidas. Experiencias en la Zona Central de Santa Fe. Universidad Nacional de San Luis: Buenos Aires, Argentina, pp: 109–127.

Garza-Alonso CA, Olivares-Sáenz E, Gutiérrez-Díez A, Vázquez-Alvarado RE, López-Jiménez A. 2019. Visual symptoms, vegetative growth, and mineral concentration in fig tree (*Ficus carica* L.) under macronutrient deficiencies. Agronomy 9 (12): 787. https://doi.org/10.3390/agronomy9120787

Gonçalves CAA, Lima LCO, Lopes PSN, de Souza MT. 2006. Poda e sistemas de condução na produção de figos verdes. Pesquisa Agropecuária Brasileira 41 (6): 955–961. https://doi.org/10.1590/S0100-204X2006000600009

Kiralan M, Ketenoglu O, Sezer KS, Mehmet YF. 2023. Composition and functional properties of fig (*Ficus carica*) phenolics. *In* Ramadan MF. (ed.), Fig (*Ficus carica*): Production, Processing, and Properties. Springer Nature: Cham, Switzerland, pp: 369–394. https://doi.org/10.1007/978-3-031-16493-4_17

Kumar R, Ganesh S, Chithiraichelvan R, Upreti KK, Sulladmath VV. 2014. Effect of spacing and pruning on growth, yield and quality of cv. Deanna fig (*Ficus carica* L.). Journal of Horticultural Sciences 9 (1): 31–37.

Leonel S, Tecchio MA. 2010. Pruning time and irrigation on the fig trees 'Roxo de Valinhos', in the Botucatu region, state of São Paulo, Brazil. Bragantia 69 (3): 571–580. https://doi.org/10.1590/s0006-87052010000300008

Li L, Wonder J, Helming T, van Asselt G, Pantazopoulou CK, van de Kaa Y, Kohlen W, Pierik R, Kajala K. 2024. Evaluation of the roles of brassinosteroid, gibberellin and auxin for tomato internode elongation in response to low red:far-red light. Physiologia Plantarum 176 (5). https://doi.org/10.1111/ppl.14558

Madhumala K, Kumar V, Kumar K. 2024. Physiology and mechanism of pruning in fruit crops. International Journal of Advanced Biochemistry Research 8 (12): 443–445. https://doi.org/10.33545/26174693.2024.v8.i12f.3185

Mendoza-Castillo VM, Vargas-Canales JM, Calderón-Zavala G, Mendoza-Castillo MDC, Santacruz-Varela A. 2017. Intensive production systems of fig (*Ficus carica* L.) under greenhouse conditions. Experimental Agriculture 53 (3): 339–350. https://doi.org/10.1017/s0014479716000405

Micheloud N, Favaro JC, Castro D, Buyatti M, Favaro MA, García MS, Gariglio N. 2018. Fig production under an intensive pruning system in the moist central area of Argentina. Scientia Horticulturae 234: 261–266. https://doi.org/10.1016/j.scienta.2018.02.035

Micheloud N, Gabriel P, Favaro JC, Gariglio N. 2023. Agronomic strategies for fig cultivation in a temperate-humid climate zone. *In* Ramadan MF. (ed.), Fig (*Ficus carica*): Production, Processing, and Properties. Springer Nature: Cham, Switzerland, pp: 193–214. https://doi.org/10.1007/978-3-031-16493-4_8

Nienow AA, Chaves A, Lajús CR, Calvete EO. 2006. Produção da figueira em ambiente protegido submetida a diferentes épocas de poda e número de ramos. Revista Brasileira de Fruticultura 28 (3): 421–424. https://doi.org/10.1590/S0100-29452006000300018

Nosir W. 2023. Cultivars and agricultural practice of fig (*Ficus carica*). *In* Ramadan MF. (ed.), Fig (*Ficus carica*): Production, Processing, and Properties. Springer Nature: Cham, Switzerland, pp: 215–246. https://doi.org/10.1007/978-3-031-16493-4_9

Pérez-Sánchez R, Morales-Corts RM, Gómez-Sánchez ÁM. 2016. Agro-morphological diversity of traditional fig cultivars grown in central-western Spain. Genetika 48 (2): 533–546. https://doi.org/10.2298/gensr1602533p

Rossouw GC, Orr R, Bennett D, Bally ISE. 2024. The roles of non-structural carbohydrates in fruiting: a review focusing on mango (*Mangifera indica*). Functional Plant Biology 51 (4). https://doi.org/10.1071/FP23195

Singh A, Kishore K, Kumar P, Khapte PS, Mishra DS, Singh D, Kothyari HS. 2023. Phenological growth and development stages of common fig (*Ficus carica* L.) under arid climate of India. Folia Horticulturae 35 (2): 395–402. https://doi.org/10.2478/fhort-2023-0028

Singh A, Kumar P, Meghwal PR, Singh A, Singh D, Kotiyari HS. 2025. Improving light interception, yield and fruit quality of fig (*Ficus carica* L.) by optimizing planting system, training system and pruning season in arid conditions. Applied Fruit Science, 67 (4): 184. https://doi.org/10.1007/s10341-025-01414-7

Zhao X, Muhammad N, Zhao Z, Yin K, Liu Z, Wang L, Luo Z, Wang L, Liu M. 2021. Molecular regulation of fruit size in horticultural plants: A review. Scientia Horticulturae 288: 110353. https://doi.org/10.1016/j.scienta.2021.110353

Agrociencia

THE CONTRIBUTION OF AGRICULTURE IN THE MEXICAN ECONOMY: A REVIEW

Miguel Ángel **Martínez-Damián**[1*], Adrián **González-Estrada**[2]

[1]Colegio de Postgraduados Campus Montecillo. Socioeconomía, Estadística e Informática. Programa de Economía. Carretera México-Texcoco km 36.5, Montecillo, Texcoco, State of Mexico, Mexico. C. P. 56230.
[2]Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias. Campo Experimental Valle de México. Carretera Texcoco-Los Reyes km 13.5, Coatlinchán, Texcoco, State of Mexico, Mexico. C. P. 56250.
* Author for correspondence: angel01@colpos.mx

Citation: Martínez-Damián MÁ, González-Estrada A. 2026. The contribution of agriculture in the Mexican economy: A review. **Agrociencia** 60(4): 511-517. https://doi.org/ 10.47163/agrociencia.v60i4.3505

Editor in Chief: Dr. Fernando C. Gómez Merino

Received: September 21, 2025.
Approved: May 22, 2026.
Published in Agrociencia: June 01, 2026.

ABSTRACT

In the mid-20th century, the theory of economic growth predicted a gradual decline in the agricultural sector's share of the overall economy. This prediction is based on specific assumptions about the sector that are not necessarily true at all times. In the case of Mexico, to test this theoretical prediction, an empirical analysis was conducted using quarterly data covering the period from 1993 to 2024. The hypothesis to be tested was that the contribution of agriculture may follow its own trajectory, independent of the prediction of a secular decline in the sector. During the period examined, using a linear adjustment of the agricultural sector's contribution with a structural break, both a declining trend and a more recent increasing trend in its contribution were found, thus refuting the prediction of classical growth theory. Although the theoretical prediction points in one direction, the empirical result is different. The share of the agricultural sector may follow its own pattern in economic evolution. In the search for an explanation of the sector's recent trend, quantity and price indices were constructed to explore whether the aforementioned adjustment is due to price or quantity. The results indicate that the price component is the basis for explaining the observed phenomenon.

Keywords: economic growth, agricultural growth, price index, quantity index.

INTRODUCTION

In economic development, there has been a decline in the share of agricultural activity in the overall economy (Anderson, 1987; Gómez-Oliver, 1995; Byerlee *et al*., 2009; Baer-Nawrocka, 2016). This trend has been predicted in economic development models such as those by Lewis (1954), Johnston and Mellor (1961), and Ranis and Fei (1961). Among the reasons cited for this trend are the dual nature of the economy, the low marginal productivity of factors such as labor in agriculture, and low income elasticity for agricultural goods, among others.

Empirical tests show that these assumptions are sometimes false, allowing for agricultural development that can maintain or increase agriculture's relative share of

the economy. In other words, although a decline in the share is expected as economic development progresses, the latter may be questionable. This phenomenon has been observed empirically by Alston and Pardey (2014), especially if the process is accompanied by improvements in factor productivity, changes in tastes and preferences, and trade flows with the rest of the world. From an empirical perspective, the objective is to examine the trend in agriculture's share of the economy using quarterly data from Mexico. It is emphasized that economic development can be consistent with different patterns of agricultural participation and not necessarily with a secular decline, which constitutes the working hypothesis.

MATERIALS AND METHODS

Quarterly data on national gross domestic product (GDP) and the agricultural sector from 1993 to 2023 were used, obtained from the Bank of Mexico's website (Banxico, 2023). The two series were used at both current prices (without adjusting for inflation) and at 2018 prices, which implicitly eliminates the effect of inflation by using that year as the valuation base.

Based on current data prices, the share of Mexican agriculture in the overall economy was calculated. In this calculation, the following ratio was used:

$$R_t = GDPA_t / GDP_t$$

where *R* is the contribution of agriculture to the economy, *GDPA* is the agricultural gross domestic product, and *GDP* is the gross domestic product of the economy. Note that both the numerator and the denominator are value-based quantities (nominal MXN); however, *R* is a dimensionless number ranging from 0 to 1, which measures the proportion or contribution of agriculture to the total economy (Gómez-Oliver, 1995; Byerlee *et al.*, 2009; Alston and Pardey, 2014).

A regression line was fitted to the ratio *R* to capture its trend over time (*t*), incorporating a breakpoint in the last quarter of 2006 to estimate the slope before and after that point. The break, however, is empirical in nature and is not associated with a specific event. Since both the numerator and denominator of the indicator consist of prices and quantities, the ratio of GDP to GDP_{2018} at 2018 prices was used to explore the explanation for this behavior, that is:

$$IGDP_t = \frac{GDP_t}{GDP_{2018}} \; IGDPA_t = \frac{GDPA_t}{GDPA_{2018}}$$

where *IGDP* and *IGDPA* are the implicit price indices for *GDP* and *GDPA*, respectively, and GDP_{2018} and $GDPA_{2018}$ represent GDPs at constant prices at the 2018 base year. These indicators reflect price growth in the economy as a whole and in the agricultural

sector, with regard to the base year to which they are normalized (Garavito-Acosta *et al*., 2011). They also allow for a comparative analysis of the relative evolution of prices in each sector.

To analyze trends in quantities, an implicit quantity index (IQ) was constructed using GDP ($GDPA_{2018}$ and GDP_{2018}) and normalized to the GDP value for the first quarter of 2018 for each variable (Guerrero-de Lizardi, 2021). This produced a quantity index normalized to 1 for the base period:

$$IQGDP_t = \frac{GDP_{2018}}{GDP_{base}} \; IQGDPA_t = \frac{GDPA_{2018}}{GDPA_{base}}$$

where GDP_{base} and $GDPA_{base}$ are the respective gross domestic products for the first quarter of 2018, which was used as the base period. Tracking these indices helps understand the behavior of the agriculture-to-total-GDP ratio, with a focus on the quantity component.

RESULTS AND DISCUSSION

The trend in agriculture's share of the total economy (Figure 1) showed regression lines with a structural break. The regression of the first quarters through the last quarter of 2006 relative to the trend over time had a negative slope of -0.026, which is statistically significant (p = 0.0001). In contrast, for the subsequent quarters, the regression of this ratio over time showed a positive slope of 0.011, which is also statistically significant (p = 0.0001). Agriculture's contribution to the economy showed a downward trend;

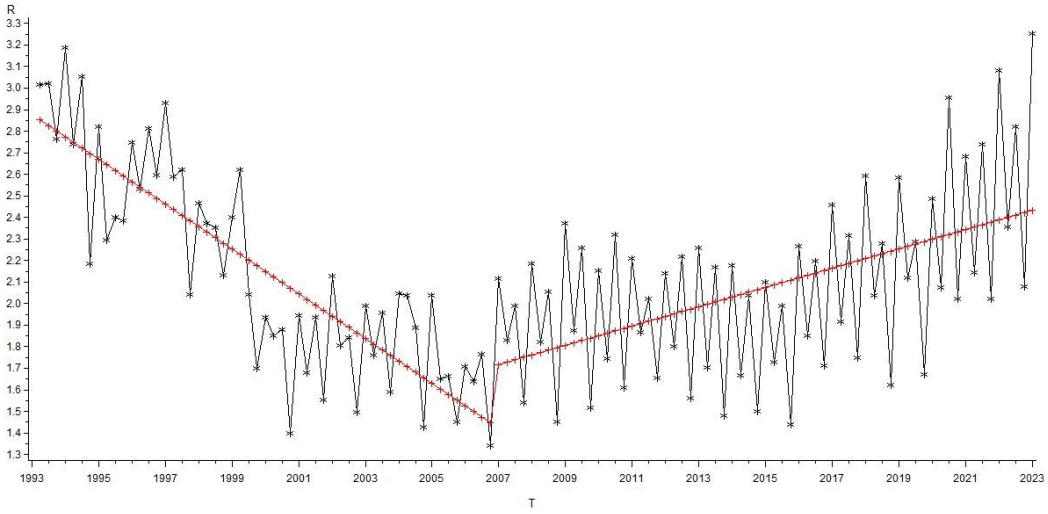

Figure 1. Agriculture's contribution to the total economy (R) in Mexico over time (T). First-period regression: R = 2.8778 – 0.0260T; second-period regression: R = 1.1379 + 0.0107T.

however, this trend breaks and gives way to an upward trend, which contradicts the predictions of agricultural development models.

Since the value of agricultural production is the product of prices and quantities, its increase can be explained by rises in prices, quantities, or a combination of both. First, the behavior of aggregate prices was analyzed by comparing implicit price indices (Church, 2016). Both indices showed a similar trend; however, toward the end of the period, an increase in the *IGDPA* was observed, indicating a relative rise in agricultural prices (Figure 2). To reinforce this analysis, the difference between the two indices was examined, which turns positive toward the end of the period studied (after 2022).

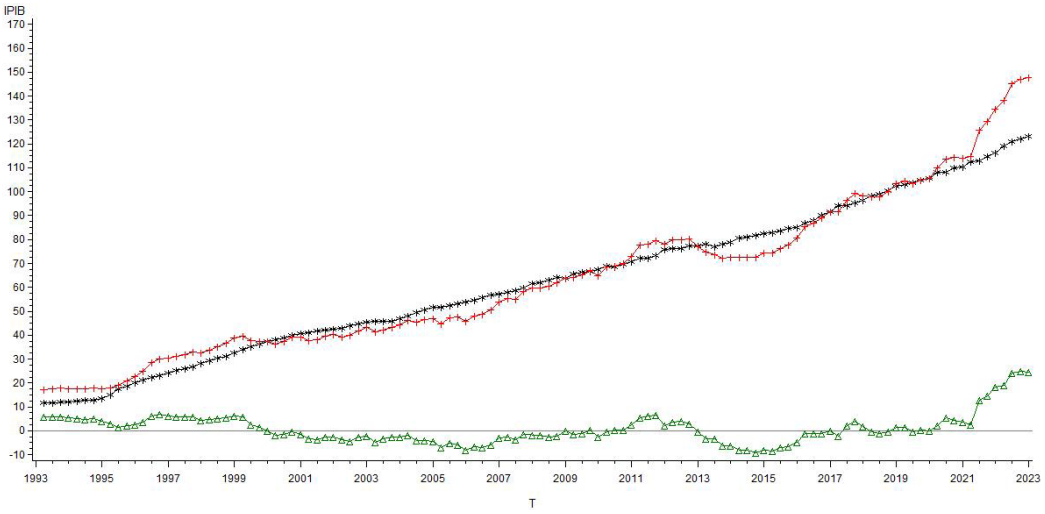

Figure 2. Implicit price indices for GDP and agriculture GDP in Mexico over time (T) (+: Agriculture, *: Total economy, Δ: Difference).

This result helps explain the recent increase in agriculture's contribution to the economy. If agricultural products become relatively more expensive, the value of production increases. This trend may be linked to factors such as changes in tastes and preferences and the effects of foreign trade (e.g., the increase in avocado exports and its impact on domestic prices) (del Moral-Barrera and Murillo-Villanueva, 2016).

Regarding quantities, the behavior of the quantity indices (Figure 3) and the differences between them were analyzed. A rising trend was observed in both indices, except for greater variability in the agricultural index. This implies that agriculture's share of the economy has a slight quantity component; however, upon examining the differences, no trend favoring one index over the other is found, but rather a slight predominance of negative differences around zero. This implies that the behavior of agriculture's share is dominated primarily by the trajectory of prices.

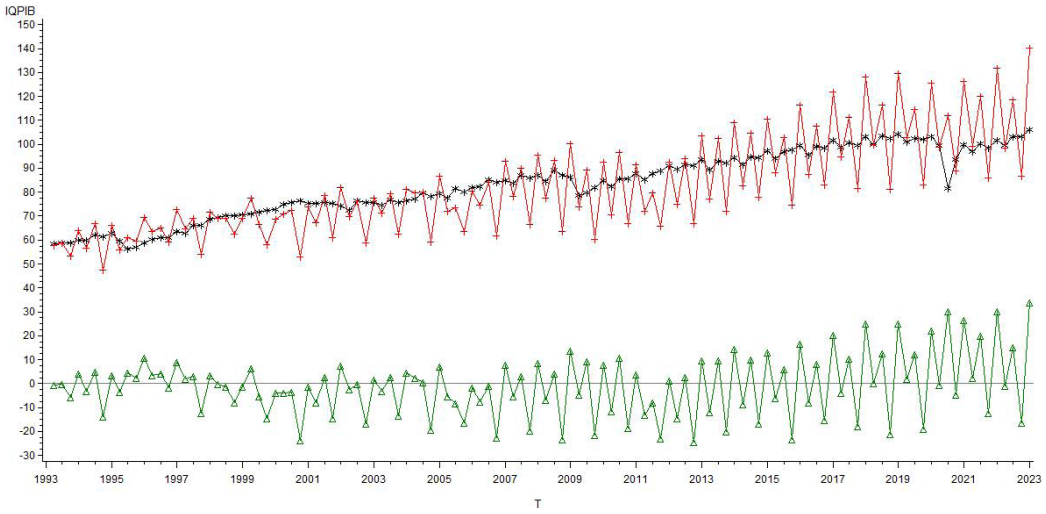

Figure 3. Quantity indices of GDP and agriculture GDP in Mexico over time (T) (+: Agriculture, *: Total economy, Δ: Difference).

A third possibility that has not been explored is a shift in crop patterns resulting from changes in relative prices, technological advances, and/or changes in agricultural policy. This goes beyond the scope of this study; the current approach assumes that such effects are already factored into changes in prices and quantities, without compromising the rigor of the analysis.

It is important to note which factors can influence prices. Assuming a constant supply, a higher price may be due to changes in tastes and preferences; that is, a greater willingness to pay for certain products such as chicken breast (Téllez-Delgado *et al*., 2016), beef, or certain fruits and vegetables like avocados and lemons (Barrera-Rojas *et al*., 2022). On the other hand, access to foreign markets can lead to shortages in the domestic market and drive prices up, as is the case with avocados. The foreign market also presents an opportunity to produce goods that would otherwise not be produced or would have limited supply, such as asparagus (Valenzuela-Romero *et al*., 2022) and cranberries (Martínez-Barajas and Torres-Zambrano, 2022).

The central idea is that, with economic development, the agricultural sector may exhibit its own distinct behavior, differing from the relative decline posited by classical theories. As early as 1961, Johnston and Mellor (1961) noted that one of agriculture's contributions to economic development is the generation of surplus for industrial development. For his part, Lenin (1992) indicates that this transfer is possible due to the existence of a large non-capitalist peasant sector. However, González-Estrada (2016) shows that, since 1970, Mexican agriculture has been characterized by a predominance of capitalist forms of enterprise production, meaning that the possibilities for continuing the transfer of economic surplus from agriculture to other sectors have come to an end.

This means that the role of agriculture in countries' economic development changes over time.

In Mexico, agriculture's contribution of economic surpluses to the development of industry and cities has already served its purpose. The capitalist sector of agriculture requires, in order to exist, an average rate of profit similar to that of the rest of the economy. The decline in agriculture's share of GDP is not an inexorable law, but rather a trend; therefore, it may well happen that, in certain periods and under certain circumstances, agriculture's contribution increases. This study presents statistical evidence of this behavior.

From 2008 onward, with the elimination of tariffs and duties on agricultural products in Mexico and the United States under the North American Free Trade Agreement, Mexican exports of fruits and vegetables to the United States began to grow rapidly, as did their prices. This explains the change in the slope of the agricultural price index trend. In this regard, Obstfeld and Rogoff (1996) showed that the dynamics of economic sectors integrated into foreign trade are greater than those of sectors not participating in such trade, and that the greater the relative participation, the higher the growth rates. Agriculture is one such sector; fruit and vegetable production is largely exported to the United States, while the staple grain sector imports over 20 million Mg from that country (USDA, 2023).

CONCLUSIONS

In Mexico, starting in 2007, there has been an increase in the contribution of agriculture to the overall economy, following a downward trend, as predicted by classical models of agricultural development. This result contradicts those predictions; without presenting an agricultural development model for Mexico, we conclude that the behavior over time of agriculture's share of the economy does not necessarily follow a pattern of secular decline in the agricultural sector's contribution to the overall economy. The observed agricultural development is consistent with an increase in its share, which is primarily due to price changes rather than changes in quantities.

REFERENCES

Alston JM, Pardey PG. 2014. Agriculture in the global economy. Journal of Economic Perspectives 28 (1): 121–146. https://doi.org/10.1257/jep.28.1.121

Anderson K. 1987. On why agriculture declines with economic growth. Agricultural Economics 1 (3): 195–207. https://doi.org/10.1016/0169-5150(87)90001-6

Baer-Nawrocka A. 2016. The role of agriculture in the national economy of EU countries. Journal of Agribusiness and Rural Development 10 (4). https://doi.org/10.17306/jard.2016.77

Banxico (Banco de México). 2023. Sistema de información económica. Producto interno bruto (CR199). Ciudad de México, México. https://www.banxico.org.mx/SieInternet/consultarDirectorioInternetAction.do?sector=2&idCuadro=CR199&accion=consultarCuadro&locale=es (Retrieved: March 2026).

Barrera-Rojas MÁ, Sánchez-Zavalegui RA, Flores-Rodríguez JR, Mex-Castillo RC, Figueroa-Chay JF. 2022. Una hipótesis sobre el incremento del precio del limón en México, 2022. Ciencia e Interculturalidad 31 (2): 98–117. https://doi.org/10.5377/rci.v31i02.15183

Byerlee D, de Janvry A, Sadoulet E. 2009. Agriculture for development: Toward a new paradigm. Annual Review of Resource Economics 1 (1): 15–31. https://doi.org/10.1146/annurev.resource.050708.144239

Church JD. 2016. Comparing the consumer price index with the gross domestic product price index and gross domestic product implicit price deflator. Monthly Labor Review 139. https://doi.org/10.21916/mlr.2016.13

del Moral-Barrera LE, Murillo-Villanueva B. 2016. Producción y preció del aguacate en México, 2011-2016. II. Economía Actual 9 (4).

Garavito-Acosta AL, Huertas-Campos CA, López-Valenzuela DC, Parra-Polanía JA. 2011. Construcción del índice de términos de intercambio para Colombia. Revista del Banco de la República 84: 1002.

Gómez-Oliver L. 1995. El papel de la agricultura en el desarrollo de México. Oficina Regional de la FAO para América Latina y el Caribe. Santiago de Chile, Chile. 52 p.

González-Estrada A. 2016. Industrialización y transnacionalización de la agricultura mexicana. Revista Mexicana de Ciencias Agrícolas 7 (3): 693–707.

Guerrero-de Lizardi CG. 2021. Revisitando a Fisher y a Friedman y Schwartz: números índices y análisis dimensional. Economía Informa 427: 48–58.

Johnston BF, Mellor JW. 1961. The role of agriculture in economic development. The American Economic Review 51 (4): 566–593.

Lenin VI. 1992. El desarrollo del capitalismo en Rusia. Editorial Progreso: Moscú, Rusia. 655 p.

Lewis WA. 1954. Economic development with unlimited supplies of labour. The Manchester School 22 (2): 139–191. https://doi.org/10.1111/j.1467-9957.1954.tb00021.x

Martínez-Barajas BM, Torres-Zambrano JP. 2022. El consumo de berries en México: análisis a través de las elasticidades mediante un sistema de demanda casi ideal. Panorama Económico 17 (36): 33–47. https://doi.org/10.29201/peipn.v17i36.109

Obstfeld M, K Rogoff. 1996. Foundations of international macroeconomics. MIT Press: Cambridge, MA, USA.

Ranis G, Fei JC. 1961. A theory of economic development. The American Economic Review 51 (4): 533–565.

Téllez-Delgado R, Mora-Flores JS, Martínez-Damián MÁ. 2016. Caracterización del consumidor de carne de pollo en la zona metropolitana del Valle de México. Estudios Sociales 26 (48): 191–209.

USDA (U.S. Department of Agriculture). 2023. Fruit and vegetable imports from Mexico continue upward trend as Mexico's growers adopt U.S. food safety rules. Washington, DC, USA. https://www.ers.usda.gov/data-products/charts-of-note/chart-detail?chartId=108128 (Retrieved: March 2026).

Valenzuela-Romero CJM, Robles-Parra JM, Tafolla-Arellano JC, Camarena-Gómez B, Preciado-Rodríguez M. 2022. Condiciones de mercado para el espárrago en el Noroeste de México. Revista Agraria 19 (1). https://doi.org/10.59741/agraria.v19i1.19

Agrociencia

CAN CARBON TRADING CATALYZE ECOLOGICAL CIRCULAR AGRICULTURE? QUASI-EXPERIMENTAL EVIDENCE FROM CHINA'S REGIONAL POLICY PILOTS

Yidan **Qiu**[1], Yunfeng **Xing**[2*]

[1]Zhejiang University of Finance and Economics. School of Economics. Hangzhou, Zhejiang 310018, China.
[2]Renmin University of China. School of Agricultural Economics and Rural Development. Beijing 100872, China.
* Author for correspondence: xingyunfengruc@ruc.edu.cn

Citation: Qiu Y, Xing Y. 2026. Can carbon trading catalyze ecological circular agriculture? Quasi-experimental evidence from China's regional policy pilots. **Agrociencia** 60(4): 518-537. https://doi.org/ 10.47163/agrociencia.v60i4.3521

Editor in Chief:
Dr. Fernando C. Gómez Merino

Received: September 12, 2025.
Approved: May 22, 2026.
Published in Agrociencia: May 29, 2026.

ABSTRACT

Ecological circular agriculture is a necessary choice for sustainable agricultural development. It is crucial for alleviating resource-environmental pressures and balancing ecological and economic progress. Within this context, the role of carbon trading policy is significant, as it acts as a catalyst in promoting the vitality of ecological circular agriculture. However, there is a dearth of research on how carbon trading affects the development of ecological circular agriculture. Therefore, this study aims to fill this gap by examining the impact of carbon trading policy on ecological circular agriculture and its underlying mechanisms. Panel data from 30 provincial-level administrative regions in China spanning 2006–2021 was used to construct a multidimensional index for ecological circular agriculture and apply a difference-in-differences (DID) approach. The findings reveal that carbon trading policy can enhance ecological circular agriculture in pilot provinces (municipalities) by 3.5 %, primarily through improved ecological technology and agricultural carbon productivity to drive the green transformation of agriculture. The effects are most pronounced in the western Chinese region and areas with stronger agricultural labor productivity. This research improves the understanding of carbon trading mechanisms in agricultural systems and provides insights for designing effective carbon trading mechanisms.

Keywords: sustainable agriculture, carbon trading policy, difference-in-differences (DID) model, mechanism analysis.

INTRODUCTION

Ecological circular agriculture is a production approach focused on the cyclic utilization of agricultural resources. It minimizes resource waste and environmental pollution by establishing closed-loop material circulation systems (Podger *et al*., 2016). Research demonstrates its significant impacts on agricultural sustainability, ecological integrity, and public health (Yue *et al*., 2022; Wang *et al*., 2025). However, ecological circular agriculture in China faces more acute challenges than developed nations, with

the excessive application of chemical fertilizers and pesticides remaining prevalent in routine farming practices (Hu and Liu, 2024).

In 2023, the fertilizer use intensity of China was recorded at 390.4 kg ha^{-1} and pesticide use intensity at 6.2 kg ha^{-1} (NBS, 2024). The fertilizer use intensity was significantly higher than the globally recommended safe level of 225 kg ha^{-1}. Moreover, pesticide overuse leads to the migration of residues into water systems and food chains through soil and rainwater, endangering soil safety, ecosystems, and human health (Sun *et al*., 2012). These practices have caused soil compaction, acidification, microbial community disruption, and fertility degradation (Qi *et al*., 2020), directly impairing crop growth and agricultural product quality (Cheng *et al*., 2025). Therefore, to achieve sustainable agricultural development and the health and well-being of residents, it is necessary to actively promote the development of ecological circular agriculture.

Previous studies indicate that the development of ecological circular agriculture is influenced by multiple factors, including agricultural technology, practitioner expertise, market demand, and policy frameworks (Pretty *et al*., 2018; Bwambale *et al*., 2022; Springmann and Freund, 2022; Shehawy and Ali Khan, 2024; Sasaki, 2025). Advanced ecological techniques in agricultural technology are important, like precision irrigation and biological pest management (Bwambale *et al*., 2022; Rad, 2025). Practitioner expertise is also critical, as farmers' awareness and knowledge affect the adoption of circular practices, while training, age, and regional contexts further shape behaviors (Pretty *et al*., 2018; Velasco-Muñoz *et al*., 2021; Wen *et al*., 2024).

Market demand is a key driver, with factors such as consumer health consciousness and government procurement accelerating the transition (Shehawy and Ali Khan, 2024; Liu *et al*., 2024b). Macro-level policies include land use planning, subsidies, and taxation. Land policies allocate resources, subsidies ease producers' financial constraints (Springmann and Freund, 2022), and taxation discourages pollution. Existing studies have explored determinants from government, technology, personnel, and market perspectives; at the policy level, they cover land use, subsidies, and taxation. However, the impact of carbon trading policy, which serves as a mandatory and guiding regulatory instrument, on ecological circular agriculture has yet to be evaluated. Additionally, its influence on the internal mechanisms of ecological circular agriculture remains unclear.

To narrow the gap, this study evaluates the efficacy of carbon trading policy in promoting ecological circular agriculture and investigates its underlying mechanisms, aiming to provide theoretical insights and practical guidance for sustainable agricultural transitions. Specifically, it makes three contributions. First, it constructs a four-dimensional index system covering resource recycling, low-carbon production technologies, ecological benefits, and economic sustainability, providing a quantitative framework for assessing ecological circular agriculture development and addressing the lack of standardized evaluation tools. Second, by using difference-in-differences (DID) and propensity score matching (PSM) methodologies, it provides empirical evidence demonstrating the causal relationship between carbon trading policy and the

advancement of ecological circular agriculture. Third, it clarifies policy mechanisms by identifying ecological technology adoption and agricultural carbon productivity enhancement as pathways, examining heterogeneous effects across regions and production structures. These insights allow for region-specific policy customization and adaptable implementation of carbon trading mechanisms.

MATERIALS AND METHODS

Research hypotheses

Theoretically, the carbon trading policy is grounded in the theory of externalities (Marshall, 2013), which argues that environmental pollution stems from the divergence between private and social costs. Agricultural production often produces negative externalities such as greenhouse gas emissions and resource depletion. Carbon trading internalizes these externalities by assigning and trading emission rights (Zeng *et al.*, 2024), using price signals to encourage emission reduction and resource efficiency. This market-based mechanism aligns private interests with social environmental goals, fostering the transition toward ecological circular agriculture (Nsabiyeze *et al.*, 2024). The comprehensive theoretical framework and the influence mechanism of carbon trading policy on ecological circular agriculture are illustrated (Figure 1).

First, by setting emission caps and creating tradable permits, carbon trading provides direct economic incentives for emission reduction. Agricultural producers can gain

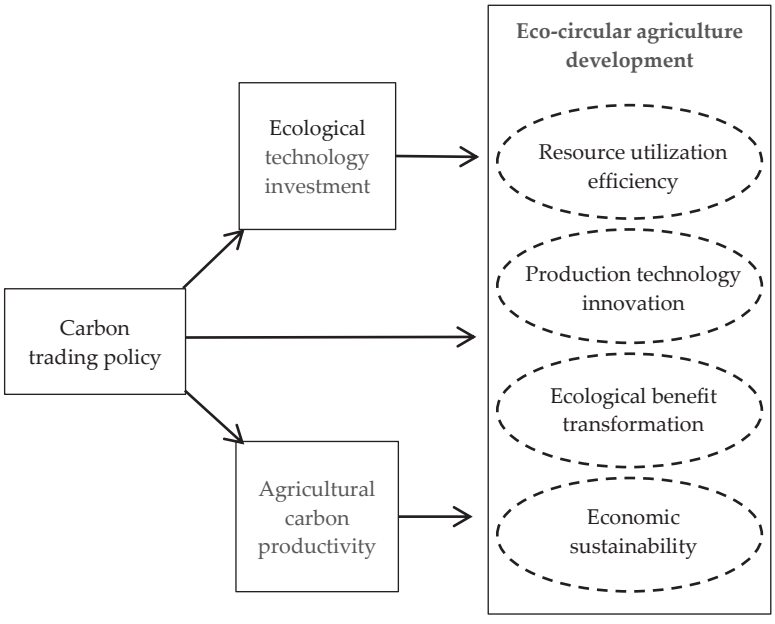

Figure 1. Influence mechanism of carbon trading policy on ecological circular agriculture (ECA).

financial returns by adopting eco-friendly practices or participating in emission reduction projects (Guo *et al.*, 2020), thereby promoting waste recycling and resource utilization in circular systems (Waldén *et al.*, 2020). Carbon trading helps reduce abatement costs and improve agricultural resource efficiency, driving ecological circular agriculture development. Consequently, it was proposed that:

H_1: Carbon trading policy exerts a statistically significant positive effect on the development of ecological circular agriculture.

Second, under carbon constraints, producers face stronger pressure to innovate. According to the Porter Hypothesis (Porter and van der Linde, 1995), properly designed environmental regulations can induce technological innovation that enhances both productivity and sustainability. By monetizing emission reductions, carbon trading stimulates investment in ecological technologies such as precision fertilization, biochar production, and manure treatment (Chen *et al.*, 2014), thus improving resource efficiency and reducing agricultural emissions. Therefore, it was proposed that:

H_2: Carbon trading policy stimulates the development of ecological circular agriculture by augmenting investments in ecological technologies.

Third, improving agricultural carbon productivity is essential for advancing low-carbon and efficient agricultural transformation, a key indicator of ecological circular agriculture (Rehman *et al.*, 2022). The adoption of ecological technologies such as precision fertilization and resource recycling reduces carbon emissions per unit of output and enhances resource efficiency. Meanwhile, carbon trading promotes industrial synergy by linking green consumption with low-carbon production, encouraging producers to adopt cleaner technologies for higher returns. Overall, higher agricultural carbon productivity embodies both economic efficiency and environmental sustainability, driving agriculture's transition toward an ecologically circular model (Huang *et al.*, 2024). Thus, it was proposed that:

H_3: Carbon trading policy stimulates the development of ecological circular agriculture by enhancing agricultural carbon productivity.

Research design

This study considered the implementation of China's carbon emissions trading policy as a quasi-natural experiment to examine the impact of carbon trading policy on the development level of ecological circular agriculture. Using a difference-in-differences (DID) approach, it constructs the following econometric model:

$$ECA_{it} = \alpha_0 + \alpha_1 did_{it} + \alpha_2 X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

$$did_{it} = Treat_{it} * Post_{it}$$

where i and t denote region and year, respectively. The explained variable ECA_{it} represents the development level of ecological circular agriculture, quantified through a composite index constructed in this study. The variable $Treat_{it}$ is a dummy

variable indicating carbon trading pilot regions, and $Post_{it}$ is a time dummy variable. The interaction term did_{it} serves as the core explanatory variable; specifically, $did_{it} = 1$ indicates that province (municipalities) i adopted the carbon trading policy in year t. The coefficient α_1 of did_{it} captures the net effect of carbon trading policy on ecological circular agriculture development. Control variables X_{it} include factors influencing ecological circular agriculture. The model incorporates province fixed effects μ_i and year fixed effects λ_t. The error term ε_{it} accounts for unobserved stochastic factors.

Robustness evaluation

To verify the reliability of the baseline regression results, this study performed a series of robustness tests.

Placebo test. To rule out the influence of unobservable random factors, a placebo test was conducted using 500 randomized simulations. Specifically, eight provinces were randomly selected as a fictional treatment group for the policy year 2013, with the remaining 22 provinces serving as the control group. This procedure was repeated 500 times to generate a distribution of estimated coefficients.

Propensity score matching (PSM-DID). A logit model was used to estimate propensity scores based on relevant covariates to address potential selection bias. This ensured that the treatment and control groups had similar pre-policy characteristics, allowing for a more accurate causal inference.

Alternative group specifications. Two additional checks were conducted: excluding Sichuan and Fujian provinces (which launched markets in 2016) to mitigate temporal heterogeneity and removing regions with high initial development levels (Beijing, Tianjin, Shanghai, and Guangdong) to eliminate potential sample bias.

Heterogeneity analysis

Given the disparities in economic conditions and resource endowments, this study examined the differential impacts of carbon trading policy through three lenses.

Regional heterogeneity. The 30 provinces were categorized into Eastern, Central, and Western regions. An inter-group difference test was performed to identify locational variations in policy efficacy.

Carbon price tiers. Pilot regions were divided into high-price and low-price groups based on the median value of average transaction prices (2013–2021) sourced from the CSMAR (2022).

Agricultural labor productivity. The sample was split into high-productivity and low-productivity groups based on the median ratio of agricultural value added to

employment. This assessed whether existing production efficiency influenced the ability of farming entities to adapt to carbon trading incentives.

Variable measurement and description

Explained variable

Ecological circular agriculture (*ECA*) was measured through a composite index structured across four dimensions based on agroecological principles (Altieri *et al*., 2018) and the "3R" framework of the circular economy (Liu *et al*., 2017). This study used the entropy weight method to objectively assign weights to each indicator to quantify provincial *ECA* development levels. Detailed formulae and data sources are provided (Table 1).

Table 1. Measurement system for ecological circular agriculture (ECA) development level.

Dimension	Specific indicator	Attribute
Ecological benefits	Carbon emission intensity per unit area	N
	Fertilizer application intensity	N
	Pesticide application intensity	N
Resource utilization efficiency	Multiple cropping index	P
	Agricultural water use efficiency	P
	Fertilizer use efficiency	P
Technological application level	Proportion of water-saving irrigated area	P
	Agricultural patent intensity	P
Economic sustainability	Green food certification count	P
	Green total factor productivity	P

P: positive indicator, where higher values indicate a higher level of ecological circular agriculture development; N: negative indicator, where lower values represent a higher level of development.

Core explanatory variable

The explanatory variable *did* represents the interaction term between *Post* and *Treat*. According to official guidelines, the initial carbon trading pilot regions established in 2013 included Beijing, Shanghai, Guangdong, Tianjin, Shenzhen, Hubei, and Chongqing. In 2016, Sichuan and Fujian were added to this list. This study categorized the provinces into treatment and control groups based on whether they participated in the pilot program.

To avoid administrative overlap, Shenzhen was excluded from the treatment group, as it is a sub-provincial city within Guangdong Province. As a result, the treatment group consists of eight provinces (municipalities): Beijing, Shanghai, Tianjin, Guangdong, Hubei, Chongqing, Sichuan, and Fujian. The control group includes 22 non-pilot

provinces and autonomous regions, with Tibet, Hong Kong, Macao, and Taiwan excluded due to data constraints. Since China's carbon trading market was gradually implemented after 2013, the policy shock year is designated as 2013. Consequently, *did* equals 1 for treatment provinces in 2013 and subsequent years and 0 otherwise, while control provinces retain a *did* value of 0 throughout the study period.

Mediating variables

The mediating variables in this study were ecological technology investment (*ETI*) and agricultural carbon productivity (*ACP*). This study adopts the density of green agricultural innovation patents granted to measure ecological technology investment and the ratio of agricultural added value to agricultural carbon emissions to measure agricultural carbon productivity.

Control variables

To control for the influence of other factors on the research results, the following variables were introduced: Per capita disposable income (*LI*), measured by the ratio of agricultural gross output value to employment in the primary sector, with logarithmic transformation applied; (2) cultivated land area (*TI*), in 10 thousand hectares; (3) rural education level (*RE*), proxied by the average years of schooling among rural residents; (4) industrial structure upgrading (*IU*), measured by the ratio of tertiary sector added value to secondary sector added value, reflecting industrial structure sophistication (Wang *et al*., 2019); (5) forest resource abundance (*FR*), represented by forest coverage rate; and (6) technological progress (*TP*), defined by the number of agricultural green technology patents granted (in thousands).

Data sources and descriptive statistics

This study utilized panel data from 30 provinces in China (excluding Tibet, Hong Kong, Macao, and Taiwan), covering the years 2006 to 2021. The data is primarily obtained from several sources, including the China Statistical Yearbook, China Rural Statistical Yearbook, China Agricultural Statistical Yearbook, China Energy Statistical Yearbook, China Population and Employment Statistical Yearbook, and statistical yearbooks from provincial or municipal offices. Missing values for specific years were filled in using linear interpolation to ensure data completeness. Additionally, all continuous variables were subjected to 1 % winsorization to minimize the impact of outliers. Descriptive statistics for the variables are summarized (Table 2).

Table 2. Descriptive statistics of the study variables.

Variable	Observations	Mean	Standard deviation	Minimum values	Maximum values
ECA	480	0.164	0.081	0.07	0.71
did	480	0.138	0.345	0	1
ETI	480	0.673	2.478	0	28.54
ACP	480	5.06	2.557	1.709	23.502
LI	480	11.423	0.588	10.097	12.713
TI	480	433.737	318.999	16.06	1586.41
RE	480	7.622	0.646	5.48	9.73
IU	480	1.232	0.674	0.561	4.768
FR	480	0.33	0.185	0.032	0.66
TP	480	0.152	0.245	0	1.627

ECA: ecological circular agriculture index; *did*: difference-in-differences interaction term for *Post* (time) × *Treat* (carbon trading); *ETI*: ecological technology investment (proxied by the density of granted green agricultural innovation patents); *ACP*: agricultural carbon productivity (agricultural added value to agricultural carbon emissions ratio); *LI*: per capita disposable income (log-transformed); *TI:* cultivated land area (10 thousand hectares); *RE*: rural education level (average years of schooling); *IU*: industrial structure upgrading index; *FR*: forest resource abundance (forest coverage rate); *TP*: technological progress (measured by the number (thousands) of granted agricultural green technology patents).

RESULTS AND DISCUSSION

Empirical results

Benchmark regression results

This study used the DID model to evaluate the impact of carbon trading policy on ecological circular agriculture development. The coefficients for the carbon trading policy remained positive and statistically significant across all regression specifications after incrementally incorporating control variables (Table 3). These results are consistent with findings from Wang *et al*. (2023), who argue that carbon trading mechanisms facilitate low-carbon transitions in agricultural practices. This positive effect is attributed to the policy's dual role in balancing cost constraints with incentives, guiding production models toward sustainability, and fostering technological innovation. Specifically, the results indicate that carbon trading policy significantly enhances ecological circular agriculture development levels in pilot provinces (municipalities) by 3.5 %, providing strong validation of H_1.

Table 3. Difference-in-differences (DID) estimates of the impact of carbon trading policy on ecological circular agriculture (ECA).

	ECA						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
did	0.046**	0.042**	0.039**	0.035**	0.036**	0.035**	0.035**
	(0.020)	(0.018)	(0.018)	(0.016)	(0.015)	(0.015)	(0.015)
LI		-0.178	-0.179	-0.112	-0.117	-0.135	-0.115
		(0.116)	(0.112)	(0.106)	(0.105)	(0.102)	(0.093)
TI			-0.001*	-0.001**	-0.001**	-0.001*	-0.001
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
IU				0.040	0.035	0.036	0.040*
				(0.025)	(0.024)	(0.024)	(0.021)
RE					0.029*	0.028*	0.031**
					(0.015)	(0.015)	(0.014)
FR						0.118	0.145*
						(0.079)	(0.075)
TP							0.047**
							(0.019)
Constant	0.158***	2.192	2.243*	1.429	1.265	1.433	1.146
	(0.003)	(1.323)	(1.276)	(1.218)	(1.161)	(1.127)	(1.009)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.884	0.890	0.893	0.898	0.901	0.903	0.909
N	480	480	480	480	480	480	480

(1)–(7) represent regression specifications where control variables are added incrementally to test the robustness of the core coefficient. *ECA*: ecological circular agriculture index; *did*: difference-in-differences interaction term for *Post* (time) × *Treat* (carbon trading); *LI*: per capita disposable income (log-transformed); *TI:* cultivated land area (10 thousand hectares); *RE*: rural education level (average years of schooling); *IU*: industrial structure upgrading index; *FR*: forest resource abundance (forest coverage rate); *TP*: technological progress (measured by the number (thousands) of granted agricultural green technology patents). Year FE and Province FE denote year and province fixed effects, respectively. Values in parentheses are standard errors. *** ** *: 1, 5, 10 % significance.

Pre-treatment parallel trends test

To validate the assumption of parallel trends, this study used the event study approach (Callaway and Sant'Anna, 2020) to plot dynamic treatment effects (Figure 2). Before policy implementation in 2013, the trends in ecological circular agriculture development between pilot and non-pilot regions exhibited no significant divergence. However, after 2013, pilot provinces began to experience a notable upward shift, with the effect becoming more pronounced over time. These results confirm that the assumption of parallel trends is valid, which enhances the credibility of the DID estimates.

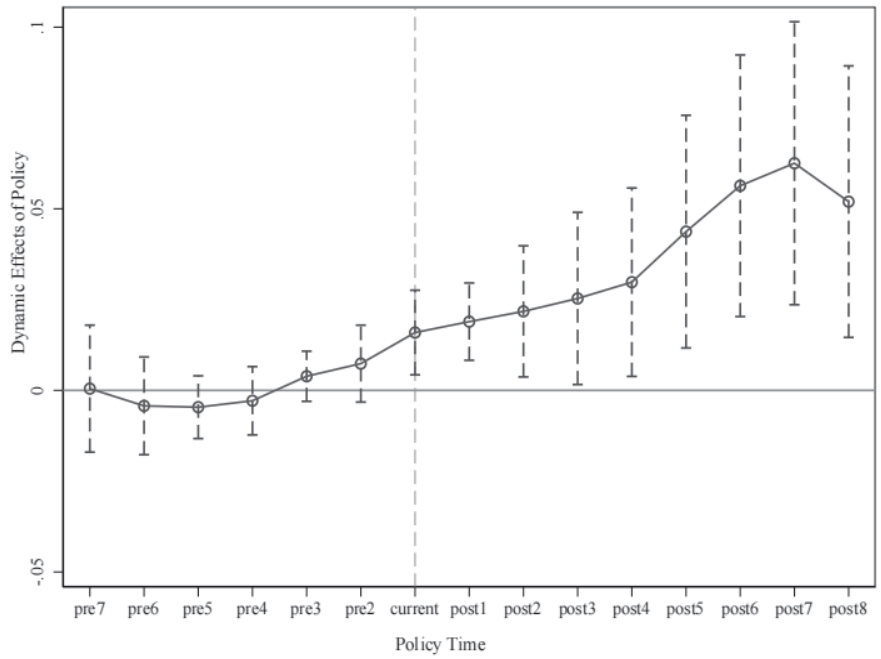

Figure 2. Event study estimates for parallel trends and dynamic treatment effects of carbon trading policy on ecological circular agriculture (ECA). The x-axis represents the time relative to the policy implementation year, which is designated as 2013; "pre" indicates the number of years prior to the policy shock (e.g., pre7 corresponds to 2006); "current" refers to the year of implementation (2013); "post" indicates the number of years following the implementation (e.g., post8 corresponds to 2021).

Placebo test

This study conducted a placebo test to determine whether changes in the development level of ecological circular agriculture are primarily due to the effects of the carbon trading policy rather than other unobservable factors (Figure 3). Specifically, 2013 was designated as the policy implementation year, and 8 out of 30 provinces were randomly selected to form a 'fictional' treatment group, with the remaining 22 provinces serving as the control group; this randomized procedure was then repeated for 500 iterations. The regression coefficients obtained through random sampling followed a normal distribution, with most coefficient values clustered around zero. Moreover, there was a significant deviation from the actual benchmark regression coefficient of 0.035. These findings suggest that based on fictional data, the hypothetical carbon trading policy does not contribute to the development of ecological circular agriculture, effectively ruling out the influence of random factors.

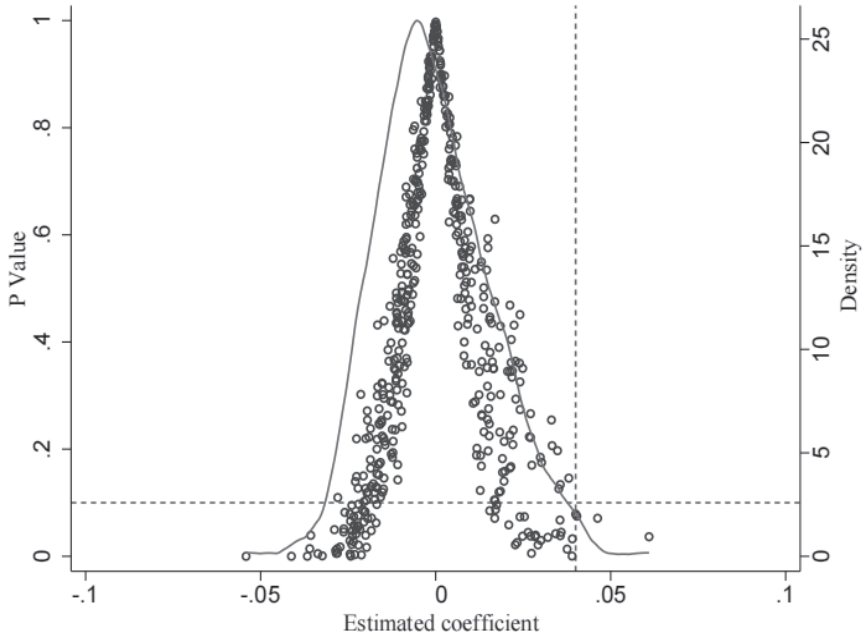

Figure 3. Placebo test based on randomized policy assignment for ecological circular agriculture (ECA).

Robustness evaluation

Propensity score matching-difference-in-differences (PSM-DID) test

To address potential selection bias, the logit model (Ventz *et al*., 2022) was used to estimate propensity scores for participation in carbon trading policies at the provincial level using covariates. As shown in Column (1) (Table 4), the core coefficient remained positive and significant, confirming the robustness of baseline findings.

Alternative treatment and control groups

Two robustness checks were conducted. First, excluding Sichuan and Fujian provinces (launched carbon markets in 2016) to mitigate heterogeneity (Column (2)); second, removing Beijing, Tianjin, Shanghai, and Guangdong (higher-than-average ecological circular agriculture levels) to eliminate potential bias (Column (3)) (Table 4). In both cases, the carbon trading policy coefficient retained positive significance, reinforcing the robustness of its positive effect on ecological circular agriculture.

Mechanism analysis

Based on the analysis of policy effects, this study further investigates the mechanisms through which carbon trading policy influences the development of ecological circular agriculture. The study focused on two mediating channels: ecological technology

Table 4. Propensity score matching-difference-in-differences (PSM-DID) and alternative specifications for the impact of carbon trading policy on ecological circular agriculture (ECA).

	ECA		
	PSM-DID	Other robustness checks	
	(1)	(2)	(3)
did	0.021***	0.045**	0.017*
	(0.005)	(0.020)	(0.009)
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
R^2	0.927	0.908	0.934
N	268	448	416

Controls refers to the same set of control variables used in Table 3, including *LI*, *TI*, *IU*, *RE*, *FR*, and *TP*. Column (1) reports the PSM-DID estimation results; Column (2) reports the results after excluding Sichuan and Fujian provinces; and Column (3) reports the results after excluding Beijing, Tianjin, Shanghai, and Guangdong. Year FE and Province FE denote year and province fixed effects, respectively. Values in parentheses are standard errors. *** ** *: 1, 5, 10 % significance.

investment (*ETI*) and agricultural carbon productivity (*ACP*). To empirically test these mechanisms, the following econometric models were constructed:

$$Media_{it} = \beta_0 + \beta_1\, did_{it} + \beta_2\, X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

$$ECA_{it} = \gamma_0 + \gamma_1\, did_{it} + \gamma_2\, Media_{it} + \gamma_3\, X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

where $Media_{it}$ represents the mediating variables and γ_2 quantifies the indirect effect of the carbon trading policy through these mechanisms. Other variables are consistent with Equation (1).

The results of the mechanism analysis (Table 5) show that columns (1) and (3) use ecological technology investment and agricultural carbon productivity as dependent variables, respectively, to examine the impact of the carbon trading policy on these mediators. The estimates reveal statistically significant positive coefficients for *did*, indicating that it substantially enhances ecological technology investment and agricultural carbon productivity.

These findings align with prior studies by Feng *et al*. (2024) and Yang *et al*. (2024), which suggest that carbon trading policy encourages agricultural producers to invest in ecological technology research and development and adoption through economic incentives. This study further focuses on the specific field of ecological

Table 5. Mechanism analysis. Effects of carbon trading policy on ecological circular agriculture (ECA) through ecological technology investment (ETI) and agricultural carbon productivity (ACP).

	(1) *ETI*	(2) *ECA*	(3) *ACP*	(4) *ECA*
did	1.123***	0.022***	0.390**	0.030***
	(0.319)	(0.005)	(0.192)	(0.005)
ETI		0.012***		
		(0.002)		
ACP				0.003***
				(0.001)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
R^2	0.737	0.942	0.880	0.930
N	480	480	480	480
Sobel test	0.0114*** (z = 3.592)		0.0064** (z = 2.550)	
Goodman test1	0.0114*** (z = 3.579)		0.0064** (z = 2.535)	
Goodman test2	0.0114*** (z = 3.604)		0.0064** (z = 2.566)	
Indirect effect	0.0114*** (z = 3.592)		0.0064** (z = 2.105)	
Direct effect	0.0222*** (z = 3.823)		0.0225*** (z = 3.715)	
Total effect	0.0336*** (z = 5.200)		0.0290*** (z = 4.469)	
Mediation proportion	0.3391		0.2217	

ECA: ecological circular agriculture index; *did*: difference-in-differences interaction term for *Post* (time) × *Treat* (carbon trading); *ETI*: ecological technology investment (proxied by the density of granted green agricultural innovation patents); *ACP*: agricultural carbon productivity (agricultural added value to agricultural carbon emissions ratio). Controls includes the full set of variables (*LI*, *TI*, *IU*, *RE*, *FR*, and *TP*) consistent with Table 3. Values in parentheses are standard errors. The Sobel and Goodman tests assessed the significance of the mediating effects. *** ** *: 1, 5, 10 % significance.

circular agriculture, clarifies how carbon trading promotes the development of ecological circular agriculture by enhancing investments in ecological technology and increasing agricultural carbon productivity, expands the research scope regarding the relationship between carbon trading policy and agricultural development, and provides more targeted theoretical support for policy-making and practice.

To formally evaluate the mediating effects, the Sobel and Goodman tests were utilized. These tests are particularly effective for calculating the Z-statistics of indirect paths and quantifying the mediation proportion, providing a clear measure of the extent to which the policy's impact is transmitted through specific channels. The results indicate significant mediation: ecological technology investment (*ETI*) accounts for 33.91 % of the total effect, while agricultural carbon productivity (*ACP*) accounts for

22.17 %. These findings validate H_2 and H_3, confirming that the carbon trading policy fosters ecological circular agriculture development through these dual mechanisms. To further ensure the robustness of these mediation pathways, the Bootstrap method was used to estimate confidence intervals. As a non-parametric approach, Bootstrap does not require the indirect effect to follow a normal distribution, offering a more reliable and rigorous statistical confirmation. The results reveal that the 95 % bias-corrected confidence intervals for both *ETI* and *ACP* exclude zero, reinforcing the significance of the identified mediation effects.

Heterogeneity analysis

In examining the carbon trading policy impact on ecological circular agriculture, a single macro-level perspective was found insufficient for policy refinement and implementation. Multifaceted factors, including economic conditions, resource endowments, and policy landscapes, influence the development of ecological circular agriculture. Meanwhile, variables like geographic positioning, carbon pricing mechanisms, and agricultural labor productivity show significant heterogeneity. Studying these variations helped understand how carbon trading policy works in different contexts and overcome the limitations of uniform policies. This analysis provides a scientific basis for governments to develop differentiated strategies, optimize resource allocation, and promote more efficient, equitable, and resilient development of ecological circular agriculture. Therefore, this study conducted three heterogeneity analyses on regional disparities, carbon price, and agricultural productivity.

Regional heterogeneity

Variations in economic conditions, resource availability, and regional policy priorities lead to divergent policy outcomes. Factors such as funding, technology, forest vegetation resources, and supporting infrastructure significantly influence the efficacy of carbon trading policy (Li *et al*., 2024; Liu *et al*., 2023, 2024a). To address this gap, regional heterogeneity was analyzed by categorizing 30 provinces into eastern, central, and western regions (Table 6), and the inter-group difference test was conducted successfully.

The impact of the carbon trading policy on ecological circular agriculture varies across regions due to locational differences. In eastern China, the policy has positively affected ecological circular agriculture thanks to strong economic and technological resources (Dai *et al*., 2025). However, as practices mature, market saturation and diminishing returns on technology investments have slowed growth (Gan *et al*., 2024). In the western region, the carbon trading policy shows strong potential. From 2021 to 2023, it reduced enterprise taxes by over CNY 400 billion through favorable tax policies. By the end of 2023, the balance of green loans in this area increased by 30.1 % (PBC, 2024) compared to the previous year. These measures encourage businesses and farmers to adopt better ecological farming practices and improve their operations. Additionally, the region's ample land, abundant renewable energy, and other resources offer great opportunities for large-scale ecological farming and breeding.

Table 6. Heterogeneity analysis of carbon trading policy effects on ecological circular agriculture (ECA) across regions, carbon price levels, and agricultural productivity groups.

	ECA						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Eastern	Central	Western	High-price group	Low-price group	High-productivity group	Low-productivity group
did	0.028***	0.006	0.033***			0.031***	0.019***
	(0.011)	(0.009)	(0.005)			(0.010)	(0.004)
DID_{high}				0.033***			
				(0.011)			
DID_{low}					0.021***		
					(0.007)		
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.908	0.964	0.962	0.904	0.902	0.909	0.958
N	176	128	176	480	480	240	240
p statistic	0.038**	0.074*	0.002***	0.010**		0.044**	

(1)–(7) represent sub-group regressions based on different regional and structural characteristics; (1)–(3) represent the Eastern, Central, and Western regions, respectively; (4)–(5) denote the high-price and low-price tiers of carbon trading pilot regions; (6)–(7) represent the high-productivity and low-productivity groups based on agricultural labor productivity. *did*: difference-in-differences interaction term for *Post* (time) × *Treat* (carbon trading); DID_{high}: the interaction term (*Post* × *Treat*) for the subgroup of pilot regions with average carbon trading prices above the sample median (including Beijing, Shanghai, Guangdong, Chongqing, and Fujian); DID_{low}: the interaction term for pilot regions with average carbon trading prices below the median (including Tianjin, Hubei, and Sichuan). Controls includes the full set of variables (*LI*, *TI*, *IU*, *RE*, *FR*, and *TP*) consistent with Table 3. Values in parentheses are standard errors. The Sobel and Goodman tests assess the significance of the mediating effects. *** ** *: 1, 5, 10 % significance.

Overall, the combined advantages make the policy impact slightly stronger in the West than in the East. However, structural and institutional barriers constrain the effectiveness of carbon trading policies in the central region. As a traditional agricultural heartland, its long-standing focus on yield stability has led to a path dependence on conventional practices, hindering the adoption of eco-friendly technologies and reducing farmers' participation (Wang *et al*., 2022). Meanwhile, the expansion of energy-intensive agri-industries such as food processing and feed production, coupled with limited profits, weakens incentives for technological upgrading (Du *et al*., 2023). Moreover, local governments' emphasis on industrial and grain output, along with weak coordination between low-carbon and agricultural policies, further limits policy effectiveness.

Heterogeneity in carbon allowance trading prices

Since pilot regions hold substantial autonomy in operating their carbon trading markets, the enforcement intensity of policies varies significantly across different pilot areas. Eight pilot regions were classified into distinct tiers based on their 2006–2021 average transaction prices, sourced from the China Stock Market and Accounting Research database (CSMAR, 2022). This study conducts separate regression analyses for high-price and low-price groups to examine the differential impacts of carbon trading prices on ecological circular agriculture. The high-price group includes Beijing, Shanghai, Guangdong, Chongqing, and Fujian, while the low-price group includes Tianjin, Hubei, and Sichuan.

The carbon trading policy exhibits significant positive effects on ecological circular agriculture across all carbon price levels (Table 6), though the magnitude of impact varies substantially with price tiers. Regions with higher carbon prices demonstrate a more substantial policy effect than low-price regions. This disparity arises because entities in high-price regions face elevated decarbonization costs, incentivizing them to actively adopt low-carbon technologies and ecological circular agriculture practices to reduce compliance expenses and capitalize on carbon credit revenues. Such price-driven incentives accelerate technological upgrading and systemic shifts toward circular agroecosystems, enhancing agricultural sustainability. In contrast, low-price regions experience weaker motivation for transformation: agricultural producers may perceive the costs of maintaining traditional practices as lower than transitioning to ecological circular agriculture, resulting in limited investments in sustainable technologies.

Heterogeneity in agricultural labor productivity

Agricultural labor productivity is a crucial metric for assessing production efficiency and shows significant disparities across different regions and farming entities. In this study, agricultural labor productivity is measured as the ratio of agricultural value added to agricultural employment. The sample was classified into high-productivity and low-productivity groups based on the median value.

The results of the heterogeneity test for different productivity levels (Table 6) reveal that the high-productivity group experienced a significantly greater improvement in ecological circular agriculture after the policy was implemented compared to the low-productivity group. This difference can be attributed to the fact that high-productivity regions typically have access to advanced production technologies, better management practices, and a more skilled labor force. These advantages allow them to quickly take advantage of new opportunities under the carbon trading policy. On the other hand, low-productivity regions face challenges due to outdated technologies, inefficient resource management, and limited technical training, which hinder their ability to fully benefit from policy incentives. The weaker response to the policy in these areas highlights the need for targeted interventions, such as precision agronomy extension services and AI-driven decision-support tools, to help bridge productivity gaps and enhance sustainability outcomes.

CONCLUSIONS

This study demonstrates that carbon trading policy significantly promotes ecological circular agriculture in China, increasing the development level of pilot regions by 3.5 %. This effect is mainly transmitted through ecological technology investment (33.91 % of the total effect) and agricultural carbon productivity (22.17 %). Policy impacts are stronger in the Eastern region, where institutional support and market maturity are higher, as well as in areas with higher carbon prices and agricultural labor productivity. These findings indicate that market-based carbon trading can align ecological sustainability with agricultural modernization through innovation and efficiency improvements.

Several policy implications emerge from these findings. First, carbon trading coverage and standardized agricultural carbon accounting should be expanded to encourage broader farmer participation through simplified registration and government-led certification programs. Second, given that technology investment and productivity are the main transmission channels, dedicated funding and fiscal incentives should support low-carbon technologies such as biochar and precision irrigation, particularly for smallholders.

Regional disparities should also be addressed through tailored interventions, including agricultural digitalization in central regions and the expansion of demonstration zones in developed areas. In addition, establishing a unified national carbon trading platform would improve market liquidity, reduce regional price disparities, and strengthen long-term incentives for ecological circular agriculture. These findings provide empirical support for both China's carbon trading policy and broader policy adaptation.

ACKNOWLEDGEMENTS

Yidan Qiu is the first author and Yunfeng Xing is the corresponding author. And the authors declare no conflict of interest. This research was supported by The Outstanding Innovative Talents Cultivation Funded Programs 2024 of Renmin University of China.

REFERENCES

Altieri MA, Farrell JG, Hecht SB, Liebman M, Magdoff F, Murphy B, Norgaard RB, Sikor TO. 2018. Agroecology. CRC Press: Boca Raton, FL, USA. https://doi.org/10.1201/9780429495465

Bwambale E, Abagale FK, Anornu GK. 2022. Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: A review. Agricultural Water Management 260: 107324. https://doi.org/10.1016/j.agwat.2021.107324

Callaway B, Sant'Anna PHC. 2021. Difference-in-differences with multiple time periods. Journal of Econometrics 225 (2): 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Chen C, Pan J, Lam KS. 2014. A review of precision fertilization research. Environmental Earth Sciences 71 (9): 4073–4080. https://doi.org/10.1007/s12665-013-2792-2

Cheng Z, Zhu M, Cai J. 2025. Reducing fertilizer and pesticide application through mandatory agri-environmental regulation: Insights from "Two Zero" policy in China. Environmental Impact Assessment Review 110: 107716. https://doi.org/10.1016/j.eiar.2024.107716

CSMAR (China Stock Market and Accounting Research Database). 2022. GTA Data. Shenzhen, China. https://data.csmar.com/ (retrieved: April 2026).

Dai L, Yang A, Song G. 2025. Regional synergy between green technology innovation and low carbon development in China: A quantitative study of provinces and economic belts. Energy Reports 13: 1085–1094. https://doi.org/10.1016/j.egyr.2024.12.074

Du Y, Liu H, Huang H, Li X. 2023. The carbon emission reduction effect of agricultural policy—Evidence from China. Journal of Cleaner Production 406: 137005. https://doi.org/10.1016/j.jclepro.2023.137005

Feng T, Wang X, Shi Y, Tu Q. 2024. The role of carbon price signal in green innovation: Evidence from China. Journal of Environmental Management 370: 122787. https://doi.org/10.1016/j.jenvman.2024.122787

Gan T, Zhou Z, Li S, Tu Z. 2024. Carbon emission trading, technological progress, synergetic control of environmental pollution and carbon emissions in China. Journal of Cleaner Production 442: 141059. https://doi.org/10.1016/j.jclepro.2024.141059

Guo J, Gu F, Liu Y, Liang X, Mo J, Fan Y. 2020. Assessing the impact of ETS trading profit on emission abatements based on firm-level transactions. Nature Communications 11 (1). https://doi.org/10.1038/s41467-020-15996-1

Hu Y, Liu Y. 2024. Impact of fertilizer and pesticide reductions on land use in China based on crop-land integrated model. Land Use Policy 141: 107155. https://doi.org/10.1016/j.landusepol.2024.107155

Huang X, Zhang T, Wang X, Zheng J, Xu G, Wu X. 2024. Regional differences of agricultural total factor carbon efficiency in China. Humanities and Social Sciences Communications 11 (1). https://doi.org/10.1057/s41599-024-03296-8

Li G, Huang Y, Peng L, You J, Meng A. 2024. Agricultural carbon reduction in China: The synergy effect of trade and technology on sustainable development. Environmental Research 252: 119025. https://doi.org/10.1016/j.envres.2024.119025

Liu B, Ding CJ, Hu J, Su Y, Qin C. 2023. Carbon trading and regional carbon productivity. Journal of Cleaner Production 420: 138395. https://doi.org/10.1016/j.jclepro.2023.138395

Liu H, Wang H, Nong H, He Y, Chen Y, Wang H, Yu M. 2024a. Opportunities and implementation pathway for China's forestry development under the "Dual Carbon" strategy. Carbon Research 3 (1). https://doi.org/10.1007/s44246-024-00144-x

Liu L, Liang Y, Song Q, Li J. 2017. A review of waste prevention through 3R under the concept of circular economy in China. Journal of Material Cycles and Waste Management 19 (4): 1314–1323. https://doi.org/10.1007/s10163-017-0606-4

Liu S, Jin Y, Zheng F. 2024b. Did product certification and e-commerce benefit agricultural producers in China? Food Policy 127: 102688. https://doi.org/10.1016/j.foodpol.2024.102688

Marshall A. 2013. Principles of economics. Palgrave Macmillan: London, UK. 731 p. https://doi.org/10.1057/9781137375261

NBS (National Bureau of Statistics of China). China statistical yearbook 2024. China Statistics Press. Beijing, China. https://www.stats.gov.cn/sj/ndsj/2024/indexeh.htm (Retrieved: April 2026).

Nsabiyeze A, Ma R, Li J, Luo H, Zhao Q, Tomka J, Zhang M. 2024. Tackling climate change in agriculture: A global evaluation of the effectiveness of carbon emission reduction policies. Journal of Cleaner Production 468: 142973. https://doi.org/10.1016/j.jclepro.2024.142973

PBC (People's Bank of China). 2024. 2023 Environmental, social and governance report. Beijing, China. 169 p. https://pic.bankofchina.com/bocappd/report/202403/P020240328722666521727.pdf (Retrieved: April 2026).

Podger D, Hoover E, Burford G, Hak T, Harder MK. 2016. Revealing values in a complex environmental program: A scaling up of values-based indicators. Journal of Cleaner Production 134: 225–238. https://doi.org/10.1016/j.jclepro.2015.08.034

Porter ME, van der Linde C. 1995. Toward a new conception of the environment-competitiveness relationship. The Journal of Economic Perspectives 9 (4): 97–118. https://doi.org/10.1257/jep.9.4.97

Pretty J, Benton TG, Bharucha ZP, Dicks LV, Flora CB, Godfray HCJ, Goulson D, Hartley S, Lampkin N, Morris C, *et al*. 2018. Global assessment of agricultural system redesign for sustainable intensification. Nature Sustainability 1 (8): 441–446. https://doi.org/10.1038/s41893-018-0114-0

Qi D, Si Z, Scott S. 2020. Can we be more collaborative? Top-down policies and urban-rural divides in the ecological agriculture sector in Nanjing, China. Society and Natural Resources 34 (2): 208–226. https://doi.org/10.1080/08941920.2020.1774951

Rad J. 2025. The rise of synthetic ecosystems in agriculture: Artificial intelligence as the future of urban food systems. Discover Sustainability 6 (1). https://doi.org/10.1007/s43621-025-00914-6

Rehman A, Alam MM, Alvarado R, Işık C, Ahmad F, Cismas LM, Mungiu Pupazan MC. 2022. Carbonization and agricultural productivity in Bhutan: Investigating the impact of crops production, fertilizer usage, and employment on CO_2 emissions. Journal of Cleaner Production 375: 134178. https://doi.org/10.1016/j.jclepro.2022.134178

Sasaki H. 2025. Sustainable agriculture and agricultural environmental policy. *In* Kuriyama K. (ed.), Economics of Sustainable Agriculture. Springer Nature: Singapore, pp: 45–66. https://doi.org/10.1007/978-981-96-3502-3_3

Shehawy MY, Ali Khan SMF. 2024. Consumer readiness for green consumption: The role of green awareness as a moderator of the relationship between green attitudes and purchase intentions. Journal of Retailing and Consumer Services 78: 103739. https://doi.org/10.1016/j.jretconser.2024.103739

Springmann M, Freund F. 2022. Options for reforming agricultural subsidies from health, climate, and economic perspectives. Nature Communications 13 (1): 82. https://doi.org/10.1038/s41467-021-27645-2

Sun B, Zhang L, Yang L, Zhang F, Norse D, Zhu Z. 2012. Agricultural non-point source pollution in China: Causes and mitigation measures. Ambio 41 (4): 370–379. https://doi.org/10.1007/s13280-012-0249-6

Velasco-Muñoz JF, Mendoza JMF, Aznar-Sánchez JA, Gallego-Schmid A. 2021. Circular economy implementation in the agricultural sector: Definition, strategies and indicators. Resources, Conservation and Recycling 170: 105618. https://doi.org/10.1016/j.resconrec.2021.105618

Ventz S, Khozin S, Louv B, Sands J, Wen PY, Rahman R, Comment L, Alexander BM, Trippa L. 2022. The design and evaluation of hybrid controlled trials that leverage external data and randomization. Nature Communications 13 (1). https://doi.org/10.1038/s41467-022-33192-1

Waldén P, Ollikainen M, Kahiluoto H. 2020. Carbon revenue in the profitability of agroforestry relative to monocultures. Agroforestry Systems 94 (1): 15–28. https://doi.org/10.1007/s10457-019-00355-x

Wang C, Zhang X, Vilela ALM, Liu C, Stanley HE. 2019. Industrial structure upgrading and the impact of the capital market from 1998 to 2015: A spatial econometric analysis in Chinese regions. Physica A: Statistical Mechanics and Its Applications 513: 189–201. https://doi.org/10.1016/j.physa.2018.08.168

Wang S, Yuan L, Gong B. 2023. China's agricultural green transition and high-quality development toward carbon neutrality. Chinese Political Science Review 8 (2): 240–272. https://doi.org/10.1007/s41111-023-00238-8

Wang X, Huang J, Liu H. 2022. Can China's carbon trading policy help achieve carbon neutrality? — A study of policy effects from the Five-sphere Integrated Plan perspective. Journal of Environmental Management 305: 114357. https://doi.org/10.1016/j.jenvman.2021.114357

Wang Z, Wang H, Zhao M, Zhao N, Lyu Y, Meng X, Wang L, Lyu P. 2025. Ecological circular disposal of agricultural waste: Integrated production of gas, electricity, heat, and fertilizer for achieving synergistic effects of pollution reduction and carbon emission reduction. Circular Economy 4 (1): 100130. https://doi.org/10.1016/j.cec.2025.100130

Wen J, Cheng N, Ma Y, Deng B. 2024. An intelligent coupled optimization model for circular agriculture considering ecological and economic benefits. Environment, Development and Sustainability. https://doi.org/10.1007/s10668-024-05553-x

Yang X, Liu Y, Bezama A, Thrän D. 2024. Agricultural carbon emission efficiency and agricultural practices: Implications for balancing carbon emissions reduction and agricultural productivity increment. Environmental Development 50: 101004. https://doi.org/10.1016/j.envdev.2024.101004

Yu D, Liu L, Gao S, Yuan S, Shen Q, Chen H. 2022. Impact of carbon trading on agricultural green total factor productivity in China. Journal of Cleaner Production 367: 132789. https://doi.org/10.1016/j.jclepro.2022.132789

Yue Q, Guo P, Wu H, Wang Y, Zhang C. 2022. Towards sustainable circular agriculture: An integrated optimization framework for crop-livestock-biogas-crop recycling system management under uncertainty. Agricultural Systems 196: 103347. https://doi.org/10.1016/j.agsy.2021.103347

Zeng S, Fu Q, Haleem F, Shen Y, Peng W, Ji M, Gong Y, Xu Y. 2024. China's carbon trading pilot policy, economic stability, and high-quality economic development. Humanities and Social Sciences Communications 11 (1). https://doi.org/10.1057/s41599-024-03646-6

Agrociencia

A METHODOLOGICAL APPROACH TO QUANTIFYING FLOOD RISK IN AGRICULTURAL AREAS: MAIZE CROP

Michell Deyanira **Cruz-Santiago**[1], Rodrigo **Roblero-Hidalgo**[1*], Héctor Alonso **Ballinas-González**[1], Juan Francisco **Gómez-Martínez**[1], José Antonio **Quevedo-Tiznado**[1], Roel **Simuta-Champo**[1]

[1]Instituto Mexicano de Tecnología del Agua. Paseo Cuauhnáhuac 8532, Colonia Progreso, Jiutepec, Morelos, Mexico. C. P. 62550.
* Author for correspondence: rodrigo_roblero@tlaloc.imta.mx

Citation: Cruz-Santiago MD, Roblero-Hidalgo R, Ballinas-González HA, Gómez-Martínez JF, Quevedo-Tiznado JA, Simuta-Champo R. 2026. A methodological approach to quantifying flood risk in agricultural areas: Maize crop. **Agrociencia** 60(4): 538-552. https://doi.org/ 10.47163/agrociencia.v60i4.3283

Editor in Chief:
Dr. Fernando C. Gómez-Merino

Received: December 16, 2025.
Approved: June 05, 2026.
Published in Agrociencia: June 22, 2026.

ABSTRACT

The agricultural sector is vulnerable to flooding, as it causes damage to soil, crops, and hydro-agricultural infrastructure, thereby limiting production. Quantifying and delineating flood-prone areas is important, as food security is at stake. This study was conducted in Irrigation District (ID) 008 Metztitlán, Mexico, which experiences recurrent flooding that affects the production system and hydro-agricultural infrastructure. The objective of this study was to develop and apply a methodology to assess flood risk in agricultural areas focused on maize cultivation. A basic hydrological and hydraulic model was developed for the proposed methodological framework for agricultural risk analysis, considering three flood factors that affect crops: A) duration, B) depth, C) velocity, and the phenological stage of growth. Based on these factors, parameters were proposed for assessing hazard, vulnerability, and exposure value to calculate risk in monetary terms using map algebra. The scenario analyzed was for a 20-year return period, determining that 94.7 % of the total area of the ID presents some degree of risk. The proposed methodology allowed for the generation of risk maps. Delineating risk zones can aid decision-making to mitigate flood damage in the agricultural sector.

Keywords: hydrological modeling, hydraulic modeling, return periods, hazard, vulnerability.

INTRODUCTION

Floods are natural disasters of hydrometeorological origin that represent 50 % of all natural disasters and significantly affect society (Mendoza-Cariño *et al*., 2018). In Mexico, they cause substantial economic losses (Hernández-Uribe *et al*., 2017). The agricultural sector is vulnerable to flooding, causing damage to the regional and national economy through the disruption of the production cycle, reduced income, unemployment, and food shortages (Vega-Serratos *et al*., 2018). In 2023, national agricultural production amounted to MXN 921 877 021.79 thousand, corresponding to a planted area of 20 023 594.61 ha and a harvested area of 18 383 456.31 ha, implying a loss due to natural disasters (floods, frosts, and winds) of 1 640 138.3 ha (SIAP, 2024). Flood risk assessment has become an increasingly common practice because it enables

disaster mitigation. Quantifying flood risk is a complex task due to the variables involved (depth, duration, and velocity) and its dynamic nature (every flood behaves differently) (Hernández-Uribe *et al*., 2017). Studies conducted in Europe indicate that the most advanced methodologies for assessing flood risks are linked to the economic evaluation of tangible direct damages (Vega-Serratos *et al*., 2018).

Flooding is an environmental stressor that negatively affects the growth of crops (Aslam and Aslam, 2023). The links between flooding and food security are extremely significant (Pacetti *et al*., 2017). The economic damage caused by flooding in the agricultural sector is less severe compared to urban areas; therefore, damage assessment is often neglected or addressed only through simple approaches and rough estimates (Merz *et al*., 2010). The impact of floods on final crop yields is not well understood due to the scarcity of datasets and the lack of quantitative models (Shirzaei *et al*., 2021). Among studies focusing on direct damage to agriculture, hazard parameters are considered the most influential in calculating direct damage. However, the parameters used are restricted to those that can be obtained from hydraulic models (Brémond *et al*., 2013).

Based on flood damage assessments in agricultural areas conducted in Europe, Förster *et al*. (2008) identified the main influencing variables as the time of year when the flood occurs, water depth, flood duration, flow velocity, and the deposition of pollutants. Brémond *et al*. (2013) indicate that the flood parameters that can be used to construct damage functions for agriculture are the seasonality of the flood, water depth, duration, flow velocity, sediment deposition, environmental pollution, and water salinity. Citeau (2003) assessed the effects of flooding on soils and crops in relation to water depth. Tariq *et al*. (2021) calculated risk in agricultural areas as the interaction of hazard with vulnerability, exposure, susceptibility, intensity, and probability, using damage curves for maize and rice. Yildirim and Demir (2022) conducted a comprehensive assessment of agricultural flood risk using flood maps, analyzing the seasonal variation in flood risk focused on maize, soybean, and alfalfa.

In Mexico, Baró-Suárez *et al*. (2007) developed a methodological framework for assessing flood damage to maize crops using a function based on the parameters of flood duration, depth, and seasonality. Vega-Serratos *et al*. (2018) developed a methodological framework for constructing damage curves using the parameters of flood duration, depth, and seasonality.

There is a need to identify and assess agricultural areas at risk of flooding using new methodological approaches; therefore, the objective of this study was to develop and apply a methodology for assessing flood risk in agricultural areas, with a focus on maize cultivation. The methodology involved proposing values for depth, duration, and velocity that could cause irreversible damage to the maize plant depending on its current phenological stage during growth. The values proposed facilitated the development of criteria for assessing hazard, vulnerability, and exposure. This approach aimed to calculate risk in monetary terms through the application of map algebra. The case study focused on Irrigation District (ID) 008 Metztitlán, which faces

recurrent flooding due to the overflow of the lagoon and the Metztitlán River, resulting in significant impacts on the production system.

MATERIALS AND METHODS

Information used

This study utilized relevant data, such as the topographic survey of the Metztitlán River from the Venados Bridge (0+000) to km 23+630, conducted using a total station with cross-sections every 20 m. In addition, the Digital Elevation Model (MDE) obtained from the Mexican Elevation Continuum (CEM) 3.0 was used, with a resolution of 15 × 15 m. For edaphology and land use data, the vector dataset from Series VI, scale 1:250 000, 2016 edition, was used (INEGI, 2016).

The precipitation data corresponds to daily precipitation from 19 conventional weather stations with an average of 30 years of records: 13015-San Agustín Metzquititlán, 13029-San Lorenzo Sayula, 13031-Santiago Tulantepec, 13033-Singuilucan, 13041-Tulancingo (Obs), 13042-Zacualtipán, 13061-Alcholoya, 13077-Metztitlán, 13082-Presa La Esperanza, 13087-San Cristóbal, 13093-Venados, 13095-Agua Blanca, 13096-Atotonilco, 13098-Huasca, 13099-Metepec, 13116-El Zembo, 13121-Santa María Amajac, 13130-Santa María Asunción, and 30359-Palo Bendito (SMN, 2024). Additionally, information on instantaneous discharge corresponding to hydrometric station 26042-Venados was obtained from the National Surface Water Data Bank of the National Water Commission (BANDAS, 2024).

The overall methodology consisted of four phases: I: data collection (MDE, vector data on soil science, land use and vegetation, daily precipitation, measured discharge, and derivation of design storms); II: hydrological modeling using HEC-HMS software version 4.2 (basin model, meteorological model, time series, control specifications, calibration, and derivation of hydrographs by sub-basin); III: hydraulic modeling using Iber software version 3.3 (geometry, roughness, boundary conditions, problem data, maximum cumulative enclosure, calibration, and determination of flood depth, velocity, and duration, which are the hazard criteria); and IV: agricultural analysis (phenology and crop characteristics, flood seasonality, flood duration, and characterization of flood-prone areas, which are part of the vulnerability criteria).

Study area

The ID 008 Metztitlán is located within the Metztitlán River basin, in the state of Hidalgo, Mexico. The basin covers an area of 2937.03 km^2. The length of the main channel is 109.54 km. The average annual precipitation is 450–700 mm, the concentration time is 1671.58 min, and the lag time is 1002.95 min (16.7 h). The ID is located at the extreme coordinates 20° 28′ 02.56″ N, 98° 40′ 15.24″ W and 20° 40′ 41.69″ N, 98° 51′ 37.92″ W, covering an area of 3553.7 ha.

Base information processing

A key part of the research involves the collection, review, and analysis of historical data on daily precipitation recorded at weather stations and daily instantaneous discharge data from the hydrometric station. Based on the MDE and using a geographic information system, 38 sub-basins and the river network were delineated. The characteristics of each sub-basin were determined: area, channel length, concentration time, delay time, and weighted curve number. For the precipitation analysis, the annual maximum precipitation (mm d^{-1}) was determined. A frequency analysis of these maximum precipitation events was conducted to obtain values associated with different return periods (RP) using probability distribution functions such as normal, lognormal, log-Pearson III, gamma, Gumbel, and double Gumbel.

Once this process was completed, the function that best fit the maximum annual precipitation was selected using the mean square error criterion. The frequency analysis was conducted at 19 conventional weather stations. For the instantaneous discharge rates at the Venados hydrometric station, which measures runoff from the upper part of the basin comprising 22 sub-basins, the daily maximum values were determined, and the corresponding frequency analysis was performed.

Hydrological modeling

The hydrological model was developed using HEC-HMS software version 4.2. For the basin component, data on area, length of the main channel, weighted number of curves, and lag time were used for each sub-basin. For the meteorological component, the precipitation weighting method was employed, using precipitation data based on the catchment area of each station within the basin. For the time series component, the rainfall distribution was modeled using centered unitary hyetographs.

The simulation was performed for return periods (RP) of 2, 5, 10, 20, 50, 100, 500, and 1000 years. Subsequently, the model was calibrated by comparing the simulated peak discharges for different return periods with those obtained from the frequency analysis of the maximum daily instantaneous discharges at the hydrometric station. The model was considered calibrated when a difference of ±15 % in discharge and an R^2 of 0.98 were obtained. The adjusted parameter was the curve number within a 10 % range.

Hydraulic modeling

The hydraulic model was developed using Iber software version 3.3. To construct the model, hydrometric station 26042-Venados was designated as the starting point of the river channel and the inlet to ID 008. The channel is 38.75 km long and has 11 tributaries along its course before emptying into Lake Metztitlán. The basin is exorheic, as it discharges into the Almolón River through two tunnels with maximum discharge rates of 53 and 11 $m^3 s^{-1}$, which are considered in the model as outlet conditions. The inflow hydrographs (19 tributaries) were obtained from the hydrological model.

The model setup included: 1) geometric definition using an unstructured mesh with

cell sizes of 10 m for the main channel and tributaries, 15 m for the lagoon, and 30 m for the remaining surfaces; 2) assignment of roughness based on land use and vegetation for each non-uniform rational B-spline (NURBS) surface; 3) mesh generation and elevation assignment based on the topographic survey of the channel, integrated with the MDE; 4) definition of initial and boundary conditions by inputting hydrographs; and 5) flow simulation over 172 800 s, with 1000-s intervals.

The model was calibrated by comparing the simulated floodplain with a satellite image corresponding to the October 5, 1999 event, which occurred from the community of Venados to Lake Metztitlán and affected 3363 ha (94.6 % of the area of ID 008). The comparison with the flood extent generated by the hydrographs for that date yielded a percentage error of 17.5 %.

Agricultural analysis

Maize cultivation is of national importance because it is the primary grain used in the Mexican diet, as reflected in the average annual per capita consumption, which stood at 196.4 kg in 2017 (SAGARPA, 2017). Maize is planted in two growing seasons: spring-summer (March 1 to September 30) and fall-winter (October 1 to February 28). Depending on the season in which flooding occurs, the effects may have a greater or lesser impact.

The plant is robust in stature, with a single, erect stem that can reach up to 4 m in height, without branches or internodes, and has a spongy pith; it bears separate male and female inflorescences on the same plant (SAGARPA, 2017). Phenological stages mark its growth according to the number of days after sowing, reaching heights of 0.5, 1, 1.5, 2, and 2.5 m at 20, 30, 40, 60, and 80 d, respectively.

Water stress caused by flooding is a limiting factor in maize production (Jiménez *et al*., 2012). Citeau (2003) indicates that the maximum submersion time the plant can withstand is one day, with a water velocity of 0.5 m s^{-1}; Baró-Suárez *et al*. (2007) note that if the plant exceeds 30 cm in height, it can withstand more than 6 d under wet conditions, whereas below that height it is more susceptible, and if flooding persists for more than 3 d, negative effects are inevitable.

Agricultural risk

To calculate risk, an equation was developed based on the National Center for Disaster Prevention (CENAPRED) definition of risk, in which risk is the combination of three factors: hazard, vulnerability, and value of exposed assets. For agricultural areas, the main factors affecting crops were considered: depth, duration, and velocity in relation to the phenological stages of maize, such that each factor contributes a component to risk, hazard (Equation 1), vulnerability (Equation 2), exposure value (Equation 3), and risk (Equation 4):

$$Hazard\ (H) = depth\ (m) \times velocity\ (m\ s^{-1}) \quad (1)$$

$$Vulnerability\ (V) = f\ (depth\ (m),\ crop\ phenology) \quad (2)$$

$$Exposure\ value\ (E) = f\ (duration\ (h),\ crop\ yield\ (Mg\ ha^{-1})) \quad (3)$$

$$Risk\ (R) = P \times V \times C \quad (4)$$

Hazard index

Hazard is defined based on the depth and velocity of the flood (Equation 1). The hazard criteria assume an average maize plant height of 2 m and a maximum tolerable velocity of 0.5 m s^{-1}, such that the product of water depth and velocity determines the hazard classification. The hazard index was constructed on a dimensionless scale ranging from 0.05 to 1, where 0.05 corresponds to low hazard, 0.25 to medium hazard, and 1 to high hazard (Table 1).

Table 1. Hazard index classification.

Depth (m)	Velocity (m s^{-1})	Depth × Velocity	Hazard	Typification (dimensionless)
0.5	0.10	0.05	High	0.05
1.0	0.25	0.25	Medium	0.25
2.0	0.50	1.00	Low	1.00

Vulnerability index

Vulnerability depends on the depth of flooding and the plant's phenology (Equation 2). For depth, the values from the hazard index were used: 0.5 m (P1), 1 m (P2), and 2 m (P3). The phenological stages considered, due to their greater sensitivity to flooding, were V7 (internode elongation, 30–55 d after sowing), VT (flowering, pollination, and fertilization, 55–65 d), and R1 (grain filling, 65–90 d). Plant height as a function of days of development was considered to define vulnerability functions.

The vulnerability index was categorized on a dimensionless scale ranging from 0.25 to 1, where 0.25 corresponds to low vulnerability (L), 0.5 to medium vulnerability (M), and 1 to high vulnerability (H). Thus, vulnerability depends on the depth of the flooding and the plant's phenological stage (Table 2).

Exposure value

The exposure value depends on the duration of the flooding and crop yield (Equation 3). Yield losses are associated with the duration of the flooding. The values used were adapted from Citeau (2003) and Baró-Suárez *et al*. (2007) (Table 3).

The exposure value will be equal to the economic losses (Equation 5), where the area is given in hectares (ha), the yield in megagrams per hectare (Mg ha^{-1}), and the price in MXN Mg^{-1}.

$$Financial\ losses = Affected\ area \times Yield \times Price \quad (5)$$

Table 2. Vulnerability index classification.

	V7	VT	VR	Vulnerability	Typification	Vulnerability function for phenological stage V7		Vulnerability function for the VT and R1 phenological stages	
						Depth of flooding (m)	Vulnerability index	Depth of flooding (m)	Vulnerability index
P1	M	B	B	Low	0.25	0.5	0.5	0.5	0.25
P2	A	M	M	Medium	0.5	1.0	1.0	1.0	0.5
P4	A	A	A	High	1.0	2.0	1.0	2.0	1.0

P1, P2, and P3: flood depth levels (0.5, 1, and 2 m, respectively); V7: phenological stage of internode elongation (30–55 d after sowing); VT: stage of flowering, pollination, and fertilization (55–65 d); R1: grain filling stage (65–90 d); B: low vulnerability; M: medium vulnerability; A: high vulnerability.

Table 3. Impact on yield in relation to the duration of the flooding.

Duration (days)	Impact on yield
Flash flood (1)	10 %
2 to 3	40 %
3 to 5	60 %
Above 6	100 %

RESULTS AND DISCUSSION

The methodology was applied to a 20-year RP using map algebra. The first step was to calculate the hazard. The rasters of the maximum flood extent corresponding to velocity and depth were multiplied, and the resulting raster was classified according to the hazard index (range from 0.05 to 1). This yielded the hazard map for a 20-year RP for ID 008 (Figure 1), where 2692.25 ha were classified as high hazard, representing 75.7 % of the district's total area.

The second step was to calculate vulnerability. To do so, the phenological stage affected by the flood was determined by analyzing monthly average and instantaneous precipitation, which showed peak values in September and October. This indicates that the affected cycle corresponds to spring-summer. According to the planting dates, the crop is in stages R1–R6. Based on the above, the vulnerability function for stage R1 was used, and the depth raster was classified, yielding the vulnerability map for a 20-year RP for ID 008 (Figure 2), in which 3027.39 ha exhibit high vulnerability, equivalent to 85.2 % of the district's total area.

The third step was to calculate the exposure value. This was done using, as reference values, a maize yield of 9.8 Mg ha^{-1} at a price of MXN 6000 Mg^{-1}, taken from the agricultural cost system for the state of Hidalgo published by the Trusts Established

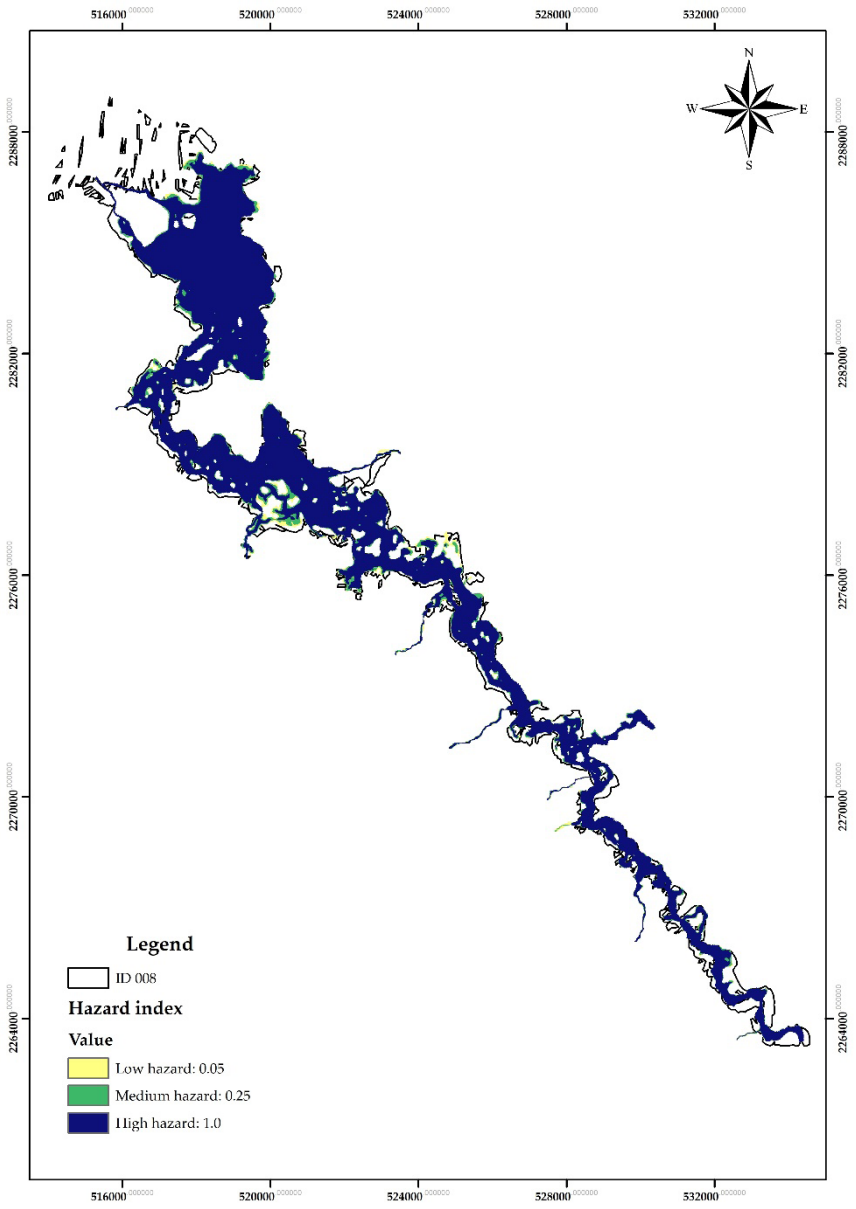

Figure 1. Twenty-year return period hazard map for Irrigation District (ID) 008 Metztitlán, Mexico.

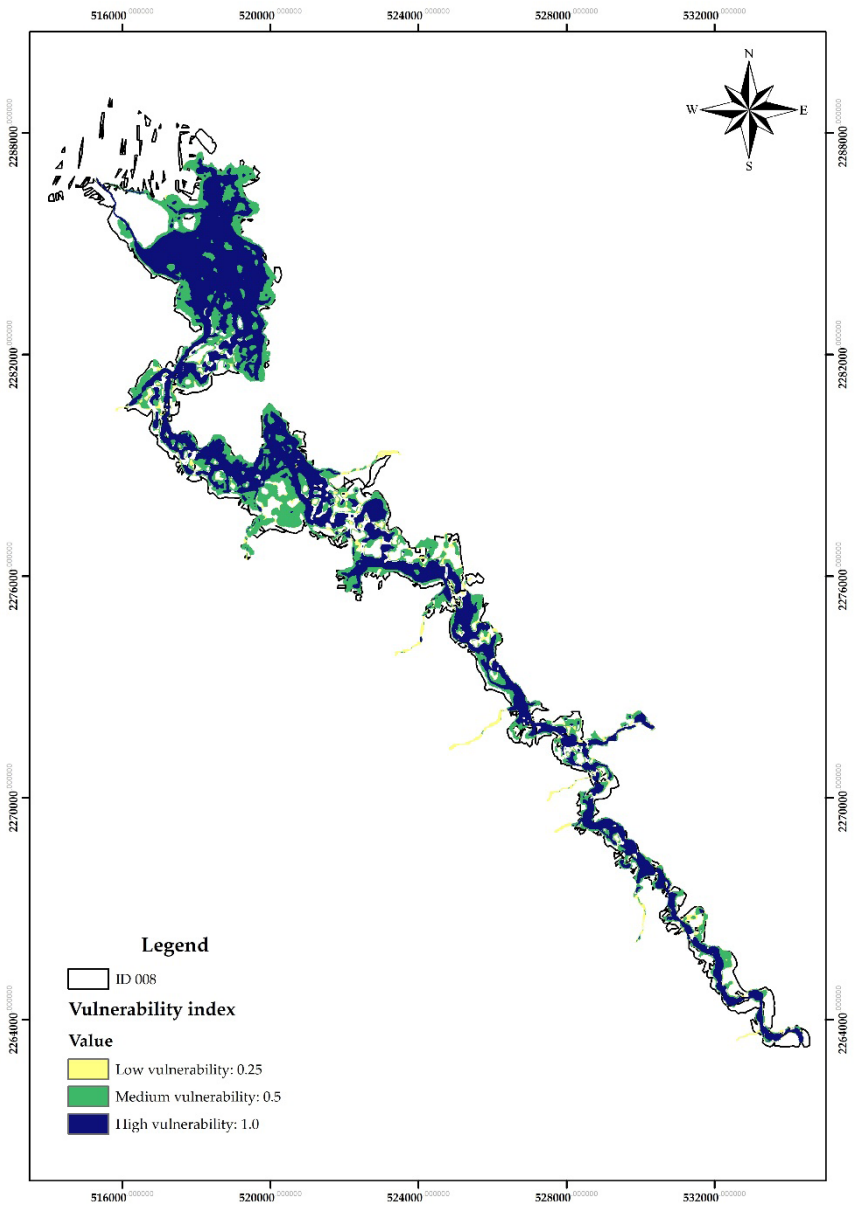

Figure 2. Twenty-year vulnerability map for Irrigation District (ID) 008 Metztitlán, Mexico.

in Relation to Agriculture (FIDA, 2024). A flood duration of 4 d was assumed, corresponding to the transit time of the complete hydrograph, which implies a 60 % reduction in yield based on the proposed values. Likewise, the hazard classification raster was used to quantify the area affected by the flood extent.

Based on these data, economic losses were estimated for each hazard classification (Table 4). These values were represented in a raster file, with each area assigned the corresponding loss value, resulting in a map of the exposure value for a 20-year return period of ID 008 (Figure 3). This indicates that, for a high exposure level, losses amount to MXN 75 819 307.2.

Table 4. Economic losses based on the classification of risk in Irrigation District 008, Metztitlán, Mexico.

Hazard	Affected area (ha)	Yield (Mg ha^{-1})	Price × Yield (MXN ha^{-1})	Economic losses (MXN)
Low	87.69	3.92	23 520.00	2 062 468.80
Medium	232.62	3.92	23 520.00	5 471 222.40
High	3223.61	3.92	23 520.00	75 819 307.20

The risk was calculated using raster algebra by multiplying the hazard, vulnerability, and exposure value rasters, which generated a risk raster in economic terms. The risk map for a 20-year RP of ID 008 (Figure 4) shows its spatial distribution based on flood exposure. The low-risk area, corresponding to losses of less than MXN 25 780.86, covers 45.2 ha, while the high-risk area, with losses between MXN 3 790 965.2 and 75 819 304, covers 1999.29 ha, equivalent to 56.25 % of the total area of ID 008. During the return period analyzed, 3348.96 ha are at some degree of flood risk, corresponding to 94.23 % of the district's total area.

The methodological approach considers the factors that influence crops, including depth, duration, rate, and seasonality, the latter of which relates to the phenological stage. According to Merz *et al*. (2010), there are few studies that incorporate velocity and seasonality in the agricultural sector; Förster *et al*. (2008) make the same point in their damage assessment proposal, which is conducted on a monthly basis considering the probability of flooding for a 100-year RP.

The agricultural vulnerability assessment methodology by Blanc *et al*. (2010) focuses on the classification and characterization of damage to the element, with an approach aimed at its reduction. In contrast, the proposed methodology is based on the characterization of the crop, the potential damage from flooding, and its tolerance, which allowed for the definition of vulnerability functions. The assessment of economic losses was carried out in a general manner to obtain damages on a large scale and over extensive areas, unlike the cost estimation by Morris and Brewin (2014), which was developed at the farm level.

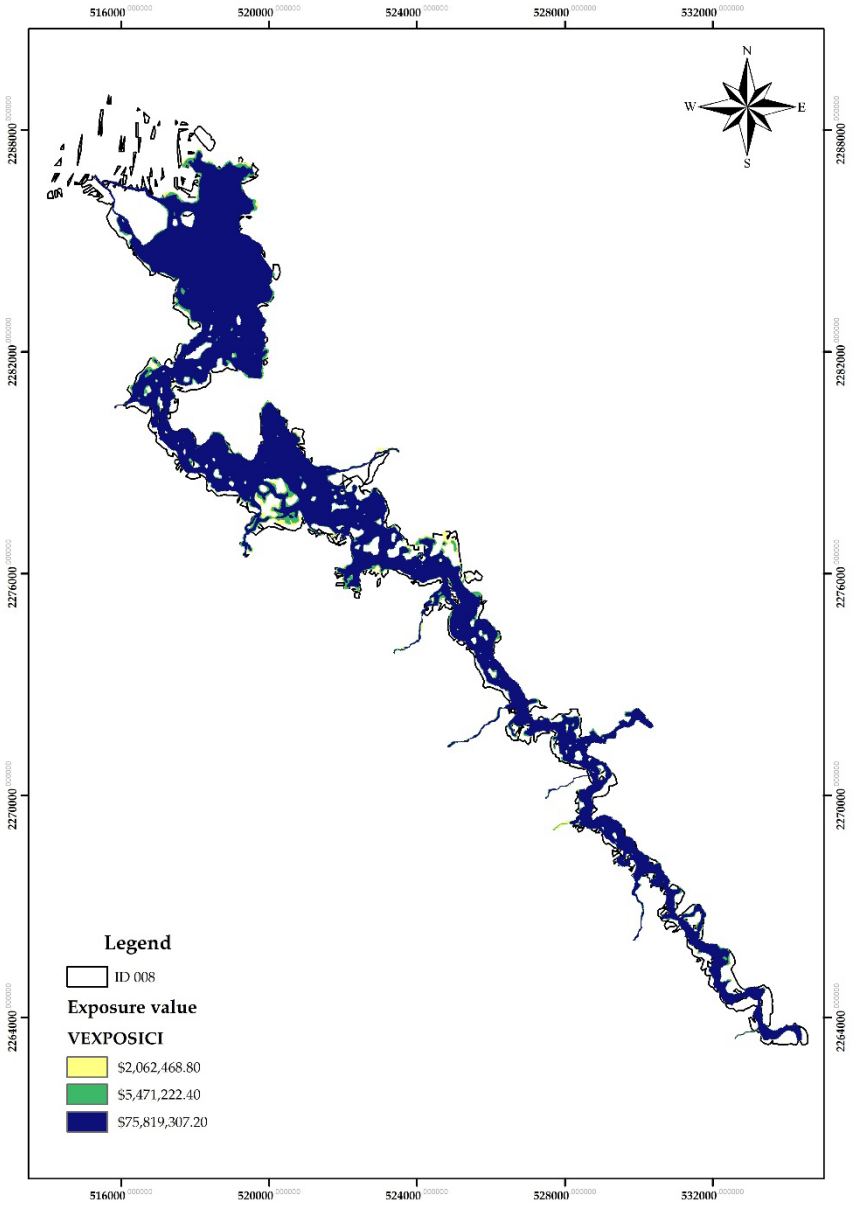

Figure 3. Exposure map for a 20-year return period (RP) for Irrigation District 008 in Metztitlán, Mexico.

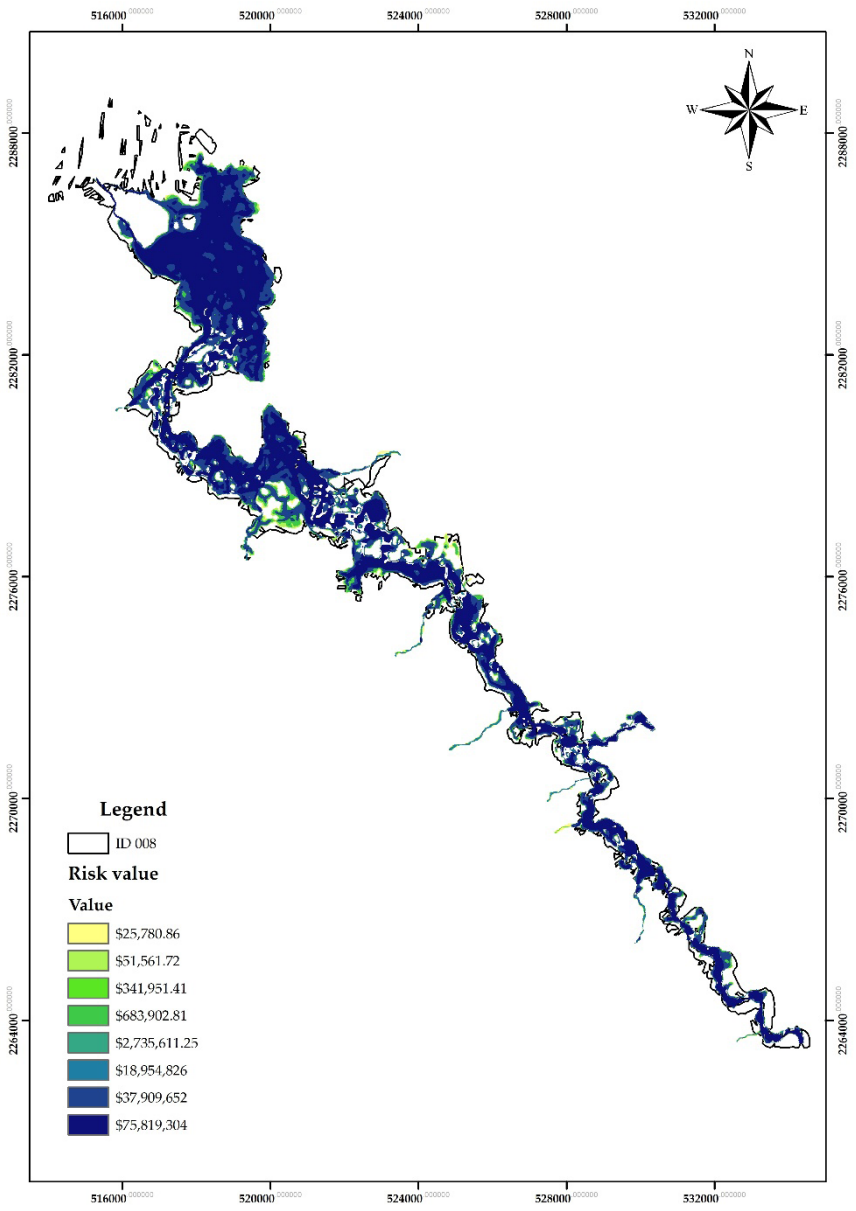

Figure 4. Risk map for a 20-year return period (RP) for Irrigation District 008 in Metztitlán, Mexico.

When comparing this methodology to approaches used in Mexico, such as the maize damage curves developed by Baró-Suárez *et al*. (2007), which are based on flood duration and apply only to plants taller than 30 cm while estimating damages in terms of minimum wages, and the methodology proposed by Vega-Serratos *et al*. (2018), which calculates average damage based on duration and probability of exceedance with monthly disaggregation, it becomes evident that the proposed methodology offers distinct advantages. It not only assesses economic damage but also allows for the calculation of hazard, vulnerability, and risk in economic terms, without being restricted to plant height. Additionally, it incorporates map algebra, simplifying the process and facilitating the generation of maps that identify zones with varying levels of hazard, vulnerability, and risk. The results illustrate a method for obtaining zoned risk maps, contrasting with the use of curves that rely on information about probability of exceedance or duration and only provide general estimates of damage.

CONCLUSIONS

The proposed methodology allows for the estimation of areas with varying degrees of risk, vulnerability, and hazard posed by flooding in the agricultural sector, specifically on plots planted with maize. The resulting risk map illustrates economic losses by area based on their exposure to flooding. Identifying the areas of greatest risk allows for the implementation of mitigation measures to reduce the effects of these events. For a more comprehensive analysis of risk quantification in an irrigation district, it would be valuable to have access to information such as crop patterns, production costs, crop value, yields, and detailed data on production cycles. Additionally, an assessment of the infrastructure is crucial; however, this information was not available in the case study.

ACKNOWLEDGMENTS

To the Mexican Institute of Water Technology and the National Council for Science and Technology for the scholarship awarded to the first author.

REFERENCES

Aslam S, Aslam S. 2023. Impact of flooding on agricultural crops—An overview. *In* Shukla, P, Singh P, Singh RM. (eds.), Water Science and Technology Library. Springer: Cham, Switzerland, pp: 255–263. https://doi.org/10.1007/978-3-031-20208-7_15

BANDAS (Banco Nacional de Datos de Aguas Superficiales). 2024. Base de datos hidrométricas. Gobierno de México. Comisión Nacional del Agua. Ciudad de México, México. https://app.conagua.gob.mx/bandas/bases.html/ (Retrieved: March 2026).

Baró-Suárez JE, Díaz-Delgado C, Esteller-Alberich MV, Calderón G. 2007. Curvas de daños económicos provocados por inundaciones en zonas habitacionales y agrícolas de México. Parte I. Propuesta metodológica. Ingeniería Hidráulica en México 22 (1): 91–102.

Blanc C, Brémond P, Grelot F. 2010. Projet EVA. Évaluation de la vulnérabilité agricole. Rapport pour la mission Rhône, DREAL Rhône-Alpes. Centre National du Machinisme Agricole, du Génie Rural, des Eaux et des Forêts. Antony, France, pp: 267.

Brémond P, Grelot F, Agenals AL. 2013. Economic evaluation of flood damage to agriculture – Review and analysis of existing methods. Natural Hazards and Earth System Sciences 13 (10): 2493–2512. https://doi.org/10.5194/nhess-13-2493-2013

Citeau JM. 2003. A new flood control concept in the Oise catchment area: Definition and assessment of flood compatible agricultural activities. TS14 New Professional Tasks - Environmental Issues and Statutory Valuation. Paris, France. 19 p.

FIDA (Fideicomisos Instituidos en Relación con la Agricultura). 2024. Agrocostos. Gobierno de México. Fideicomisos Instituidos en Relación con la Agricultura. Ciudad de México, México. https://www.fira.gob.mx/Nd/Agrocostos.jsp (Retrieved: March 2026).

Förster S, Kuhlmann B, Lindenschmidt KE, Bronstert A. 2008. Assessing flood risk for a rural detention area. Natural Hazards and Earth System Science 8 (2): 311–322. https://doi.org/10.5194/nhess-8-311-2008

Hernández-Uribe RE, Barrios-Piña H, Ramírez AI. 2017. Análisis de riesgo por inundación: metodología y aplicación a la cuenca Atemajac. Tecnología y Ciencias del Agua 8 (3): 5–25. https://doi.org/10.24850/j-tyca-2017-03-01

INEGI (Instituto Nacional de Estadística y Geografía). 2016. Conjunto de datos vectoriales de uso del suelo y vegetación. Escala 1:250 000. Serie VI. Capa unión. Ciudad de México, México. https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463173359 (Retrieved: March 2026).

Jiménez JC, Moreno LP, Magnistskiy S. 2012. Plant responses to stress due to flooding. A review. Revista Colombiana de Ciencias Hortícolas 6 (1): 96–109.

Mendoza-Cariño M, Bautista-Olivas AL, Quevedo-Nolasco A, Mendoza-Cariño D. 2018. Análisis hidrológico de largo plazo en la cuenca del río Metztitlán Hidalgo, México, y su relación con el cambio climático. Hidrobiológica 28 (1): 17–30. https://doi.org/10.24275/uam/izt/dcbs/hidro/2018v28n1/mendoza

Merz B, Kreibich H, Schwarze R, Thieken A. 2010. Assessment of economic flood damage. Natural Hazards and Earth System Science 10 (8): 1697–1724. https://doi.org/10.5194/nhess-10-1697-2010

Morris J, Brewin P. 2014. The impact of seasonal flooding on agriculture: The spring 2012 floods in Somerset, England. Journal of Flood Risk Management 7 (2): 128–140. https://doi.org/10.1111/jfr3.12041

Pacetti T, Caporali E, Rulli MC. 2017. Floods and food security: A method to estimate the effect of inundation on crops availability. Advances in Water Resources 110: 494–504. https://doi.org/10.1016/j.advwatres.2017.06.019

SAGARPA (Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación). 2017. Maíz grano blanco y amarillo mexicano. Planeación agrícola nacional 2017-2030. Ciudad de México, México. 27 p.

Shirzaei M, Khoshmanesh M, Ojha C, Werth S, Kerner H, Carlson G, Sherpa SF, Zhai G, Lee JC. 2021. Persistent impact of spring floods on crop loss in U.S. Midwest. Weather and Climate Extremes 34: 100392. https://doi.org/10.1016/j.wace.2021.100392

SIAP (Servicio de Información Agroalimentaria y Pesquera). 2024. Anuario estadístico de la producción agrícola. Gobierno de México. Servicio de Información Agroalimentaria y

Pesquera. Ciudad de México, México. https://nube.agricultura.gob.mx/cierre_agricola/ (Retrieved: March 2026).

SMN (Servicio Meteorológico Nacional). 2024. Información estadística climatológica. Gobierno de México. Servicio Meteorológico Nacional. Comisión Nacional del Agua. Ciudad de México, México. https://smn.conagua.gob.mx/es/climatologia/informacion-climatologica/informacion-estadistica-climatologica/ (Retrieved: March 2026).

Tariq MAUR, Rajabi Z, Muttil N. 2021. An evaluation of risk-based agricultural land-use adjustments under a flood management strategy in a floodplain. Hydrology 8 (1): 53. https://doi.org/10.3390/hydrology8010053

Vega-Serratos BE, Domínguez-Mora R, Posada-Vanegas G. 2018. Seasonal flood risk assessment in agricultural areas. Tecnología y Ciencias del Agua 9 (3): 92–127. https://doi.org/10.24850/j-tyca-2018-03-04

Yildirim E, Demir I. 2022. Agricultural flood vulnerability assessment and risk quantification in Iowa. Science of the Total Environment 826: 154165. https://doi.org/10.1016/j.scitotenv.2022.154165

Agrociencia

INFLUENCE OF ORGANIC MATTER ON THE PHYSICAL QUALITY OF SOILS UNDER INTENSIVE AGRICULTURAL MANAGEMENT

Víctor Manuel **Montoya-Jasso**[1], Víctor Manuel **Ordaz-Chaparro**[1*], Gerardo Sergio **Benedicto-Valdés**[1], Enrique **Ojeda-Trejo**[1], Edgar Vladimir **Gutiérrez-Castorena**[2]

[1]Colegio de Postgraduados Campus Montecillo. Posgrado en Edafología. Carretera México-Texcoco km 36.5, Montecillo, Texcoco, State of Mexico, Mexico. C. P. 56230.
[2]Universidad Autónoma de Nuevo León. Facultad de Agronomía. Pedro de Alba S/N, Ciudad Universitaria, San Nicolás de los Garza, Nuevo León, Mexico. C. P. 66451.
* Author for correspondence: ordaz@colpos.mx

Citation: Montoya-Jasso VM, Ordaz-Chaparro VM, Benedicto-Valdés GS, Ojeda-Trejo E, Gutiérrez-Castorena EV. 2026. Influence of organic matter on the physical quality of soils under intensive agricultural management. **Agrociencia** 60(4): 553-564. https://doi.org/ 10.47163/agrociencia.v60i4.3414

Editor in Chief:
Dr. Fernando C. Gómez Merino

Received: November 24, 2025.
Approved: May 15, 2026.
Published in Agrociencia: May 28, 2026.

ABSTRACT

Intensive agricultural management causes soil alteration, which leads to a decrease in some of its physical characteristics and a change in its quality and fertility. Bulk density and hydraulic conductivity are parameters used to evaluate these effects on food production. Other factors, such as soil organic matter, are also relevant for assessing soil quality, including physical quality, due to their close relationship with other soil properties. The objective of this study was to evaluate the physical quality of soils subjected to frequent changes resulting from intensive agricultural management. Five soils (S1, S2, S3, S4, and S5) classified as Vertisols from Acámbaro, in the state of Guanajuato, Mexico, were evaluated. Sampling was carried out at a depth of 0 to 20 cm to obtain a homogeneous composite sample. Nineteen physical variables and the soil organic matter content were determined. A principal component analysis was applied to rank the most important variables and calculate physical quality indices. Soil S4 had a low degradation index, associated with particle stability and high hydraulic conductivity (7.0 cm h^{-1}). Soils S1, S2, S3, and S5 were classified as having a moderate degradation index, attributable to intense mechanical effects such as subsoiling and harrowing. When the organic matter variable was integrated into the generation of the indices, the physical quality of all soils was considered high, and the degradation index decreased as a result of the high organic matter content (5.85–8.58 %) and reserves, which favor adequate physical conditions. The addition of labile organic materials, such as manure and compost, in intensive agriculture improves the physical quality of the soil over prolonged periods, with positive effects on its conservation.

Keywords: Intensive agriculture, physical fertility, soil conservation, Vertisols.

INTRODUCTION

Intensive agricultural management causes soil alteration, resulting in limited adaptability and recovery capacity. In the medium term, characteristics such as structure, water storage capacity, and resistance to environmental factors such as

wind and heavy rainfall lead to a decline in soil quality and fertility (García *et al*., 2018; de Lima *et al*., 2017). This process causes the deterioration of its physical properties, such as compaction, as well as chemical imbalances such as the formation of salt crusts and the release of environmentally significant gases (Lepore *et al*., 2024).

Authors such as Schjønning *et al*. (2015) recommend establishing studies on land use intensity and soil characterization to address degradation processes and generate preventive or corrective strategies against adverse effects on soil susceptibility. Such studies should consider soil characteristics such as bulk density, particle size distribution, and hydraulic conductivity due to their influence on the assessment of effects on crop production (Ward *et al*., 2021) and their sensitivity to changes in soil management (Novillo-Espinoza *et al*., 2018). In this context, soil organic matter is important for creating quality indices, including physical quality, due to its strong connection with other soil properties (Montoya-Jasso *et al*., 2019; Chen *et al*., 2023; Xu *et al*., 2023).

Recently, mathematical models have been created that seek to integrate various variables to define restoration and improvement strategies in the management of agricultural production systems, including soil quality criteria (Lepore *et al*., 2024). These models make it possible to identify attributes with the greatest impact that contribute to the generation of indices of the current state of the soil (Montoya-Jasso *et al*., 2022). These include models for calculating soil water deficit (Schulte *et al*., 2005), drainage rates in different texture classes (Allen *et al*., 1998), and approaches based on lysimeter observations (Diamond and Sills, 2001). However, models of physical soil quality indicators, such as the S-Index (Dexter, 2004) and the Soil Management Assessment Framework (Andrews *et al*., 2004), include organic matter as an empirical variable with arbitrary weighting, without explicitly considering quantitative evidence of its relative weight on pure physical variables. Although these models integrate physical and chemical soil variables, there is a need to combine variables with high degrees of correlation.

Based on the above, the objective of this research was to create soil indices for the proposed models, both with and without organic matter. This was conducted to assess the physical quality of soils that undergo constant changes in their characteristics due to intensive agricultural management.

MATERIALS AND METHODS

Soils

Five soils classified as Vertisols were used, originating from Jaral del Refugio, in the municipality of Acámbaro, Guanajuato, Mexico. Sampling was carried out in soils under intensive agricultural management (subsoil, harrowing, and leveling at the beginning of each agricultural cycle for more than 30 years) (Table 1), at a depth of 0–20 cm. A sub-sampling system with an auger was applied at five points, and the samples were mixed to obtain a homogeneous composite sample per plot. The soils were dried

Table 1. Description of intensive agricultural management practices at the site evaluated.

Handling practice	Description	Impacts on the soil
Conventional farming	Subsoiling, plowing, and harrowing with tractors to turn the soil to a depth of 0–30 cm, followed by leveling and sowing with precision seeders. This method is used on 70–80 % of plots used for annual crops (corn and sorghum in the spring-summer cycle, wheat, oats, and chickpeas in the fall-winter cycle).	Improves initial aeration and weed control but accelerates water erosion and loss of organic matter (González-Contreras *et al*., 2025). High dependence on machinery (100 % tractors).
Fertilization and incorporation of amendments	Intensive chemical application (320 kg ha^{-1} of N; 41–100 kg ha^{-1} of P) in furrows, followed by sowing. Transition to mixed application with organic matter (50 %) and 100 % weed management with herbicides.	It corrects deficiencies, but generates losses of up to 165 kg N ha^{-1}; inefficiency 20–40 %. In transition plots, chemical doses are reduced by 20–50 % (Bhople *et al*., 2025).
Stubble management	Incorporation of post-harvest stubble (in corn and sorghum, 5–10 Mg of residues ha^{-1}). Currently, burning is avoided in only 60 % of the soils evaluated, and manure and compost are added (farmers do not have actual data on the amount added).	It stimulates soil biology, increases organic matter, reduces compaction by machinery, and improves yields, increasing corn grain yields by up to 1020 kg ha^{-1} (González-Contreras *et al*., 2023).

in the shade at room temperature to constant weight. They were then sieved with a 2.0 mm mesh. In cases where macroaggregates were present, they were ground with a porcelain mortar to reduce their size to 2.0 mm.

Laboratory determinations

The determinations were carried out at the Soil Physics Laboratory of the Postgraduate College. To remove the organic matter from the soil and ensure representativeness in subsequent analyses, a preliminary treatment with 30 % hydrogen peroxide was carried out. In soils where organic matter needed to be quantified, its content was determined using the Walkley and Black (1947) method.

Bulk density and saturated hydraulic conductivity were determined from unaltered samples, for which PVC cores 80 mm in diameter and 10 cm in length were installed, forming soil columns. Bulk density was calculated using the following equation:

$$Bulk\ density\ (g\ cm^{-3}) = \frac{Soil\ mass\ (g)}{Core\ volume\ (cm^3)}$$

Subsequently, with the soil dried and sieved to 2.0 mm, the total porosity was calculated:

$$Total\ porosity\ (\%) = \left(1 - \left(\frac{Dap(g\ cm^{-3})}{Dr\ (g\ cm^{-3})}\right)\right) * 100$$

Textural classification was determined using the Bouyoucos (1936) method. Field capacity and permanent wilting point were obtained using the methodology proposed by Reynolds and Elrick (1986). The weighted mean diameter was estimated following the procedure described by Burés-Pastor (1997).

Likewise, the variables corresponding to the Atterberg limits (liquid limit, plasticity index, shrinkage index, and workability range) were determined. Based on these properties, the particle separability index (ISP) and the sealing-coating index (ISE) and coating index (IE) were estimated using the equations proposed by Florentino (1998) and Comerma *et al*. (1992):

$$ISP = \frac{\%\ of\ clay}{\%\ of\ silt + \%\ of\ sand}$$

$$ISE = \frac{6.7433 * \%\ of\ organic\ matter}{0.55001 * (\%\ of\ silt + \%\ of\ sand)}$$

$$IE = \frac{1.1255 * \%\ of\ silt}{(\%\ of\ clay + 10)(\%\ of\ organic\ matter)}$$

Multivariate analysis

For the analysis of determinations, SAS 9.3 statistical software (SAS Institute Inc., Cary, NC, USA) was used, and the methodology proposed by Florentino (1998) was applied, which recommends using principal component analysis (CP) to obtain a hierarchy of the most important variables and calculate the physical quality indices of the soils, taking into account their distribution and correlation along the component with the highest explained variance (Chen *et al.,* 2025).

Two CP analyses were performed. The first included the organic matter variable for all soils, while in the second this variable was excluded, with the aim of evaluating the impact of organic matter on the generation of soil physical quality indicators. Five levels of degradation (1 to 5) were identified, and the level degradation identity value (ID_N) was obtained using the following equation:

$$ID_N = \frac{V_{min}}{V_{max}}$$

where V_{min} corresponds to the level with the least degradation (1, 2, 3, or 4, as applicable) and V_{max} to the level with the greatest degradation (5).

The degradation index (*ID*) was calculated using the following expression:

$$ID = \sum_{i=1}^{n} \left(\frac{ID_N}{\#V} \right)$$

where $\#V$ corresponds to the number of variables selected by the model as the most representative ones.

Finally, the physical quality index of soils (ICFS) was calculated as:

$$ICFS = \frac{1}{ID}$$

RESULTS AND DISCUSSION

The physical characterization and organic matter content of the soil (Table 2) showed homogeneity in clay content (40 %), presumably expansive clays given the soil classification, which predisposes it to compaction (Dexter, 2004). However, the bulk density is low (<1.11 g cm^{-3}), indicating an absence of compaction despite frequent mechanization (Bondi *et al*., 2020) with subsoiling and harrowing, as well as the high organic matter content (>5.85 %), which, according to Dorji *et al*. (2019), promotes continuous improvements in physical and chemical properties, such as increased soil microporosity.

Soils S2, S3, and S5 exhibited zero hydraulic conductivity upon saturation, indicating a reduction in total pore space and suggesting structural alteration associated with intensive soil preparation for agricultural purposes (de Jong van Lier, 2017). In contrast, soils S1 and S4, with a smaller weighted mean diameter (<5.83 mm) and a crusting index <1, exhibit high stability, with no issues of surface compaction and greater mechanical resistance to losses due to irrigation water or precipitation.

In the principal component analysis of the model that included soil organic matter (MOS), CP1 explained 97.4 % of the variance, followed by CP2 with 1.41 %, which together accounted for 98.81 % of the total variability (Figure 1A). In the model excluding MOS, the component with the highest explained variance was CP1 at 97.31 %, followed by CP2 at 1.46 %, together explaining 98.76 % of the variability (Figure 1B). These results reinforce the reliability and applicability of both models for determining physical quality indices.

To calculate the degradation and physical quality indices, it is necessary to use the variables with the highest eigenvalues of CP1, as this component explains the largest percentage of variance and helps reduce measurement errors (Chen *et al*., 2025). In the

Table 2. Average values of the soil determinations made in Vertisols of Jaral del Refugio, in Guanajuato, Mexico.

Variable	Evaluated soils				
	S1	S2	S3	S4	S5
BD (g cm^{-3})	0.96	1.01	0.94	1.11	1.07
TPS (%)	61	57	55	54	57
MiP (%)	23	26	12	23	12
MaP (%)	38	31	43	31	45
Sa (%)	8	3	13	20	7
Si (%)	52	57	47	40	53
Cl (%)	40	40	40	40	40
FC (%)	48	59	38	46	39
PWP (%)	29	36	31	33	29
SHC (cm h^{-1})	1.33	0	0	7	0
WMD (mm)	5.83	6.22	6.67	4.39	6.62
OM (%)	6.63	8.58	5.85	8.19	6.24
PSI	0.667	0.667	0.667	0.667	0.667
SCI	0.495	0.563	0.383	0.188	0.548
CrI	1.355	1.753	1.195	1.673	1.275
SL	34	33	36	37	30
PL	38	37	37	41	37
LL	57	63	61	63	59
RW	19	26	24	22	22

BD: bulk density; TPS: total pore space; MiP: microporosity; MaP: microporosity; Sa: sands; Si: silts; Cl: clays; FC: field capacity; PWP: permanent wilting point; SHC: saturated hydraulic conductivity; WMD: weighted mean diameter; OM: organic matter; PSI: particle separability index; SCI: sealing-crusting index; CrI: crusting index; SL: shrinkage limit; PL: plastic limit; LL: liquid limit; RW: range of workability.

model including MOS, the representative variables were the liquid limit, sand content, permanent wilting point, crusting index, and sealing-crusting index (Figure 1A). In contrast, in the model excluding MOS, the selected variables were the liquid limit, total pore space, sand content, sealing-crusting index, and particle separability index, as they exhibited higher correlation and a high discriminatory value (Figure 1B).

Additionally, as bulk density increased, total pore space decreased (Figure 1A), indicating a decline in the soil's hydraulic properties (Chrysanthopoulos and Kallioras, 2025), including restrictions on root growth (Dexter, 2004). Likewise, MOS exhibited effects associated with a reduction in total pore space due to its high degree of recalcitrance, suggesting a gradual decrease in total porosity (Figure 1A). This behavior is consistent with that described by Minasny and McBratney (2017), who reported limiting effects on soil hydraulic capacity that favor waterlogging conditions and restrict plant growth (de Melo *et al*., 2023) when an easily mineralized organic

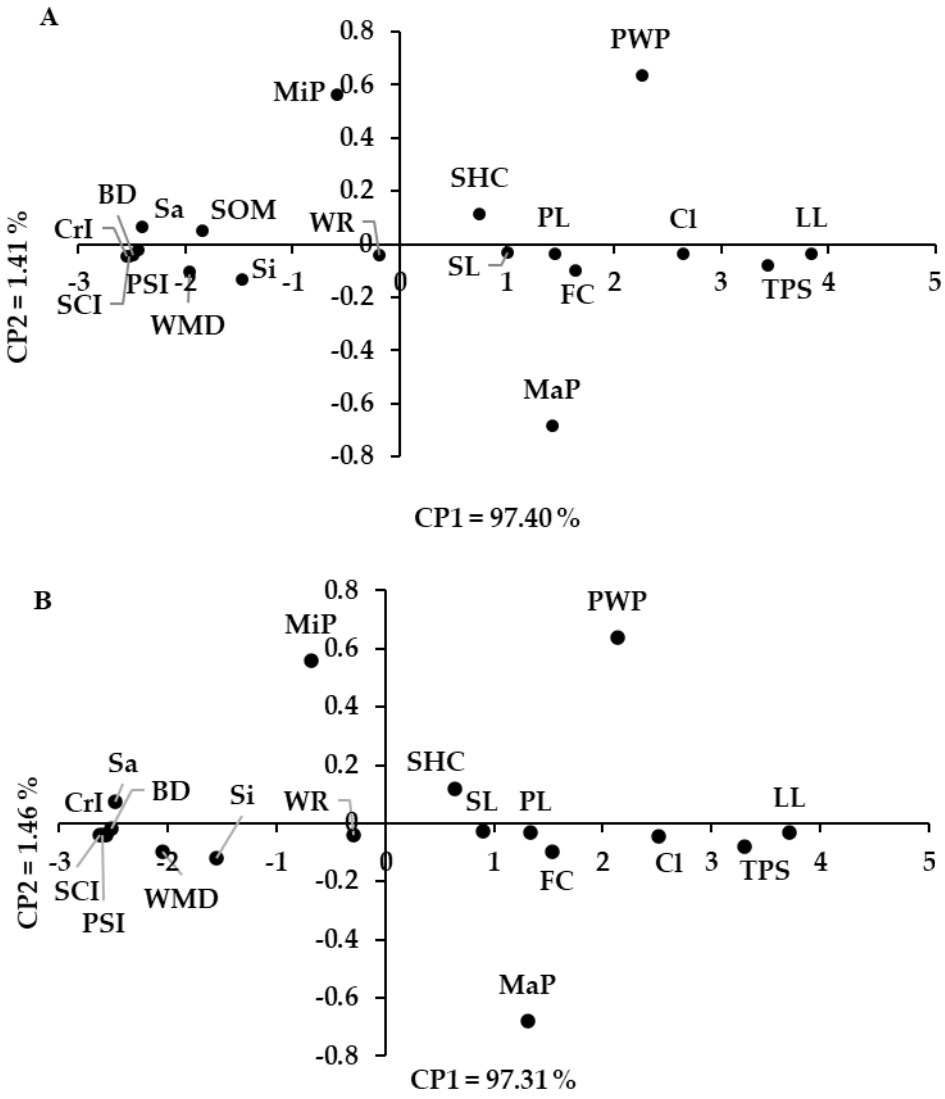

Figure 1. Biplot of the principal component (CP) analysis of soil physical properties. A) Model including soil organic matter (MOS); B) model excluding MOS. BD: bulk density; TPS: total pore space; MiP: microporosity; MaP: macroporosity; Sa: sands; Si: silts; Cl: clays; FC: field capacity; PWP: permanent wilting point; SOM: soil organic matter; SHC: saturated hydraulic conductivity; WMD: weighted mean diameter; PSI: particle separability index; SCI: sealing-crusting index; CrI: crusting index; SL: shrinkage limit; LP: plastic limit; LL: liquid limit; WR: workability range.

source is not incorporated to improve oxygen availability in the medium term (Mori *et al*., 2023).

In both models, field capacity and the permanent wilting point decreased as soil hydraulic conductivity increased (Figure 1) as a result of reduced moisture retention capacity in the surface layers (Salazar *et al*., 2022). However, this behavior is common in soils with high content of expansive clays (Pinheiro *et al*., 2019). Florentino (1998) proposed a classification of soil physical quality based on the degradation index, defined according to the intensity of the model's effect (Table 3).

Table 3. Classification of soil degradation and physical quality indices.

Class	Degradation Index	Interpretation	Physical quality index	Interpretation
1	<0.20	Very low	5.00	Very high
2	0.20–0.40	Low	5.00–2.50	High
3	0.40–0.60	Moderate	2.50–1.67	Moderate
4	0.60–0.80	High	1.67–1.25	Low
5	0.80–1.00	Very high	<1.25	Very low

Based on the classification of the degradation indices (DI) and physical soil quality indices (ICFS) for the evaluated soils (Table 4), it was observed that soil S4 had the lowest DI. This was due to aggregate stability resulting from the presence of particles with a mean diameter <4.39 mm and high hydraulic conductivity (7.0 cm h^{-1}) (Table 2). According to Reyes (2019), this reduces the risk of erosion. The remaining soils had a moderate DI, indicating a decrease in moisture retention and availability (hydraulic conductivity values <1.0 cm h^{-1}) (Minasny and McBratney, 2017) due to higher intensity of mechanized tillage (Liu *et al*., 2022).

Table 4. Assessment of the physical quality of Vertisols in Jaral del Refugio, Guanajuato, Mexico, subjected to intensive tillage.

Soil	DI	DI classification	ICFS	ICFS Classification	ICFS(t)	ICFS Classification(t)
S1	0.84	Very high	1.19	Very low	4.52	High
S2	0.84	Very high	1.19	Very low	4.78	High
S3	0.96	Very high	1.04	Very low	4.60	High
S4	0.88	Very high	1.14	Very low	4.49	High
S5	0.92	Very high	1.09	Very low	4.56	High

DI: Degradation index; ICFS: soil physical quality index; ICFS(t): transformed soil physical quality index.

Notably, ICFS values improved (Table 4) when labile organic matter was added to the soil, leading to a significant enhancement in physical quality through the formation of soil aggregates resulting from the release and formation of chelates during the mineralization of organic matter in conjunction with soil clays (Montoya-Jasso *et al*., 2021). The addition of waste creates a reserve of slowly mineralizing organic materials, which ensures the long-term preservation of soil physical quality (Głąb *et al*., 2016).

CONCLUSIONS

Intensive agricultural management reduces the ability of the topsoil to retain and distribute moisture, reflecting significant degradation effects. Despite extensive mechanization, the addition of labile organic matter in intensive agriculture improves soil physical quality over a considerable period, which has a positive impact on soil conservation. To minimize errors in interpreting the results of multivariate analysis, it is important to use reference tables specific to the soil type when evaluating soil degradation indices and physical quality.

ACKNOWLEDGMENTS

To the Secretariat of Science, Humanities, Technology, and Innovation (SECIHTI) for providing the first author with a doctoral fellowship. To the staff of the Soil Physics and Chemistry Laboratory at the Postgraduate College for their assistance in conducting the analyses. To the farmers of Jaral del Refugio, Acámbaro, Mexico, for facilitating the collection of research materials.

REFERENCES

Allen GR, Raes D, Smith M. 1998. Crop evapotranspiration. Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. Food and Agriculture Organization of the United Nations. Rome, Italy.

Andrews SS, Douglas LK, Cambardella CA. 2004. The soil management assessment framework. Soil Science Society of America Journal 68 (6): 1945–1962. https://doi.org/10.2136/sssaj2004.1945

Bhople P, Wall D, Richards K, Clough T, Brennan F, Lanigan G, Ros M, Herrmann AM, Pedersen IF, Elsgaard L, *et al*. 2025. Soil nutrient stoichiometry impacts on soil organic carbon stocks in long-term phosphorus fertilisation experiments. Geoderma 463: 117538. https://doi.org/10.1016/j.geoderma.2025.117538

Bondi G, O´Sullivan L, Fenton O, Creamer R, Marongiu I, Wall DP. 2020. Trafficking intensity index for soil compaction management in grasslands. Soil Use and Management 37 (3): 504–518. https://doi.org/10.1111/sum.12586

Bouyoucos GJ. 1936. Directions for making mechanical analysis of soils by the hydrometer method. Soil Science 42 (3): 225–230. https://doi.org/10.1097/00010694-193609000-00007

Burés-Pastor S. 1997. Sustratos. Ediciones Agrotécnicas S.L.: Madrid, España. 220 p.

Chen K, Liu K, Zhou Y, Li Y, Wu C, Gao G, Wang H, Laghrouche S, Djerdir A. 2025. State of health prognosis for polymer electrolyte membrane fuel cell based on principal component analysis and Gaussian process regression. International Journal of Hydrogen Energy 98: 933–943. https://doi.org/10.1016/j.ijhydene.2024.12.067

Chen Y, Wang L, Tong L, Hao X, Wu X, Ding R, Kang S, Li S. 2023. Effects of biochar addition and deficit irrigation with brackish water on yield-scaled N_2O emissions under drip irrigation with mulching. Agricultural Water Management 277: 108129. https://doi.org/10.1016/j.agwat.2022.108129

Chrysanthopoulos E, Kallioras A. 2025. A framework for monitoring soil hydraulic properties of undisturbed soil samples and modeling unsaturated zone water flow. Environmental Processes 12 (3). https://doi.org/10.1007/s40710-025-00791-1

Comerma J, Torres S, Lobo D, Fernández R, Delgado R, Madero L. 1992. Aplicación del sistema de evaluación de tierras de la FAO, 1985, en la zona de Turén, Venezuela. Cuadernos de Agronomía 1 (1).

de Jong van Lier Q. 2017. Field capacity, a valid upper limit of crop available water? Agricultural Water Management 193: 214–220. https://doi.org/10.1016/j.agwat.2017.08.017

de Lima RP, da Silva AP, Giarola NFB, da Silva A, Rolim MM. 2017. Changes in soil compaction indicators in response to agricultural field traffic. Biosystems Engineering 162: 1–10. https://doi.org/10.1016/j.biosystemseng.2017.07.002

de Melo ML, Inforsato L, Pinheiro EAR, de Jong van Lier Q. 2023. Plant available water predicted by flux-based approach. Geoderma 429: 116253. https://doi.org/10.1016/j.geoderma.2022.116253

Dexter AR. 2004. Soil physical quality: Part I. Theory, effects of soil texture, density, and organic matter, and effects on root growth. Geoderma 120 (3–4): 201–214. https://doi.org/10.1016/j.geoderma.2003.09.004

Diamond J, Sills P. 2001. Soil water regimes, end of project reports. Teagasc, Johnstown Castle Research Centre. Wexford, Ireland. 31 p.

Dorji T, Field DJ, Odeh IOA. 2019. Soil aggregate stability and aggregate-associated organic carbon under different land use or land cover types. Soil Use and Management 36 (2): 308–319. https://doi.org/10.1111/sum.12549

Florentino AA. 1998. Guía para la evaluación y monitoreo de la degradación de suelo y de la sostenibilidad de uso de la tierra: indicadores físicos-valores críticos. Instituto de Edafología. Facultad de Agronomía. U.C.V. Maracay, Venezuela. 9 p.

García DY, Cárdenas JF, Silva A. 2018. Evaluación de sistemas de labranza sobre propiedades físicoquímicas y microbiológicas en un Inceptisol, Colombia. Revista de Ciencias Agrícolas 35 (1): 16–25. https://doi.org/10.22267/rcia.183501.79

Głąb T, Palmowska J, Zaleski T, Gondek K. 2016. Effect of biochar application on soil hydrological properties and physical quality of sandy soil. Geoderma 281: 11–20. https://doi.org/10.1016/j.geoderma.2016.06.028

González-Contreras Y, Cisneros-López HC, Montoya-Jasso VM, Cervantes-Ortiz F, Arreola-Tostado JM. 2023. Rendimiento, calidad física e índices de cosecha nutrimental de grano de maíz como respuesta a la aplicación de fertilización sintética y orgánica. Suelos Ecuatoriales 53 (1–2): 1–11. https://doi.org/10.47864/SE(53)2023p1-11_168

González-Contreras Y, Cisneros-López HC, Montoya-Jasso VM, Cervantes-Ortiz F, Arreola-Tostado JM. 2025. Evaluación de las emisiones de CO_2 en un Vertisol con historial productivo de maíz. Biotecnología en el Sector Agropecuario y Agroindustrial 23 (2). https://doi.org/10.18684/rbsaa.v23.n2.2025.2603

Lepore E, Schmidt O, Fenton O, Tracy S, Bondi G, Wall DP. 2024. Traffic induced compaction and physical quality of grassland soil under different soil moisture deficits. Soil and Tillage Research 244: 106205. https://doi.org/10.1016/j.still.2024.106205

Liu H, Rezanezhad F, Lennartz B. 2022. Impact of land management on available water capacity and water storage of peatlands. Geoderma 406: 115521. https://doi.org/10.1016/j.geoderma.2021.115521

Minasny B, McBratney AB. 2017. Limited effect of organic matter on soil available water capacity. European Journal of Soil Science 69 (1): 39–47. https://doi.org/10.1111/ejss.12475

Montoya-Jasso VM, Benedicto-Valdés GS, Ordaz-Chaparro VM, Ruiz-Bello A, Arreola-Tostado JM, Castillo-Valdez X. 2019. Evaluation of substrates mineralization by C-CO_2 flux under nitrogen fertilization. Suelos Ecuatoriales 49 (1–2): 19–28. https://doi.org/10.47864/se(49)2019p19-28_101

Montoya-Jasso VM, Ordaz-Chaparro VM, Benedicto-Valdés GS, Ojeda-Trejo E, Gutiérrez-Castorena EV. 2022. Atributos químicos para definir la aptitud agrícola de Vertisoles del Bajío mexicano. Suelos Ecuatoriales 52 (1–2): 130–136. https://doi.org/10.47864/SE(52)2022p130-139_163

Montoya-Jasso VM, Ordaz-Chaparro VM, Benedicto-Valdés GS, Ruiz-Bello A, Arreola-Tostado JM. 2021. Caracterización química y física de sustratos enriquecidos con minerales y composta. Terra Latinoamericana 39. https://doi.org/10.28940/terra.v39i0.601

Mori T, Rosinger C, Margenot AJ. 2023. Enzymatic C:N:P stoichiometry: Questionable assumptions and inconsistencies to infer soil microbial nutrient limitation. Geoderma 429: 116242. https://doi.org/10.1016/j.geoderma.2022.116242

Novillo-Espinoza ID, Carrillo-Zenteno MD, Cargua-Chavez JE, Nabel-Moreiral V, Albán-Solarte KE, Morales-Intriago FL. 2018. Propiedades físicas del suelo en diferentes sistemas agrícolas en la provincia de Los Ríos, Ecuador. Temas Agrarios 23 (2): 177–187. https://doi.org/10.21897/rta.v23i2.1301

Pinheiro EAR, de Jong van Lier Q, Inforsato L, Šimůnek J. 2019. Measuring full-range soil hydraulic properties for the prediction of crop water availability using gamma-ray attenuation and inverse modeling. Agricultural Water Management 216: 294–305. https://doi.org/10.1016/j.agwat.2019.01.029

Reyes RW. 2019. Comparación de tres métodos de evaluación de la calidad física en un suelo Vertisol de la llanura de Coro, Falcón-Venezuela. CIENCIAMATRIA 5 (8): 144–173. https://doi.org/10.35381/cm.v5i8.91

Reynolds WD, Elrick DE. 1986. A method for simultaneous in situ measurement in the vadose zone of field-saturated hydraulic conductivity, sorptivity and the conductivity pressure-head relationship. Ground Water Monitoring Review 6 (1): 84–95. https://doi.org/10.1111/j.1745-6592.1986.tb01229.x

Salazar PM, Lozano LA, Villareal R, Irizar AB, Barraco M, Polich NG, Soraco CG. 2022. Capacity and intensity indicators to evaluate the effect of different crop sequences and cover crops on soil physical quality of two different textured soils from Pampas Region. Soil and Tillage Research 217: 105268. https://doi.org/10.1016/j.still.2021.105268

Schjønning P, van der Akker JJ, Keller T, Greve MH, Lamandé M, Simojoki A, Stettler M, Arvidsson J, Breuning-Madsen H. 2015. Driver-Pressure-State-Impact-Response (DPSIR) analysis and risk assessment for soil compaction—A European perspective. Advances in Agronomy 133: 183–237. https://doi.org/10.1016/bs.agron.2015.06.001

Schulte RPO, Diamond J, Finkele K, Holden NM, Brereton AJ. 2005. Predicting the soil moisture conditions of Irish grasslands. Irish Journal of Agricultural and Food Research 44 (1): 95–110.

Walkley A, Black IA. 1947. An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Science 37 (1): 29–38. https://doi.org/10.1097/00010694-193401000-00003

Ward M, McDonnell K, Metzger K, Forristal PD. 2021. The effect of machine traffic zones associated with field headlands on soil structure in a survey of 41 tilled fields in a temperate maritime climate. Soil and Tillage Research 210: 104938. https://doi.org/10.1016/j.still.2021.104938

Xu W, Xu H, Delgado-Baquerizo M, Gundale M, Zou X, Ruan H. 2023. Global meta-analysis reveals positive effects of biochar on soil microbial diversity. Geoderma 436: 116528. https://doi.org/10.1016/j.geoderma.2023.116528

Agrociencia

SOIL QUALITY FORECASTING WITH OPTIMAL FEATURE SELECTION AND EXTENDED CROSS-STAGE PYRAMID NETWORK

Yamuna **Vilvanathan**[1], Jeevaa **Katiravan**[2], Visu **Pandurangan**[3]*, Smitha Sarojam **Ponnaiyan**[4]

[1]Sri Sairam College of Engineering. Department of Computer Science and Engineering. Anekal, Bangalore, Karnataka 562106, India.
[2]Velammal Engineering College. Department of Information Technology. Chennai, Tamil Nadu 600066, India.
[3]Velammal Engineering College. Department of Artificial Intelligence and Data Science. Chennai, Tamil Nadu 600066, India.
[4]Velammal Engineering College. Department of Computer Science and Engineering. Chennai, Tamil Nadu 600066, India.
* Author for correspondence: dr.visu.p@gmail.com

Citation: Vilvanathan Y, Katiravan J, Pandurangan V, Ponnaiyan SS. 2026. Soil quality forecasting with optimal feature selection and extended cross-stage pyramid network. **Agrociencia** 60(4): 565-585. https://doi.org/ 10.47163/agrociencia.v60i4.3465

Editor in Chief:
Dr. Fernando C. Gómez Merino

Received: September 13, 2025.
Approved: May 15, 2026.
Published in Agrociencia: June 01, 2026.

ABSTRACT

Soil quality is essential for sustainable agriculture. Nonetheless, inadequate irrigation methods, improper fertilizer use, and over-cultivation reduce soil quality, thereby decreasing soil fertility. Precise soil quality prediction is crucial for improving agricultural practices. Conventional deep learning models often encounter problems related to superfluous features, high computational demands, and inaccurate predictions. In this work, an innovative deep learning architecture integrating optimal feature selection was proposed to address these challenges. Initially, the Adaptive Parrot Optimization (AdPo) method was used to identify the most relevant features from pre-processed soil data. The Extended Cross Stage Pyramid Network (ExCSP_Net) was introduced to improve soil quality prediction. This network integrates a gated recurrent unit (GRU)-based attention module into the main pathway of the Cross Stage Partial (CSP) model to capture long-range dependencies and emphasize relevant information. In addition, a stacked autoencoder was incorporated before the feature-sharing stage in the short path of the CSP model to reduce dimensionality and generate meaningful representations. The AdPo+ExCSP_Net model demonstrated outstanding performance, achieving an accuracy of 98.58 %, recall of 98.09 %, precision of 98.32 %, F1-score of 98.15 %, Mean Absolute Error (MAE) of 0.53, Root Mean Square Error (RMSE) of 0.65, and coefficient of determination (R^2) of 0.99. These findings highlight the effectiveness of the proposed methodology for accurate soil quality prediction and the promotion of sustainable agricultural practices.

Keywords: adaptive weighting, missing data imputation, cross stage pyramid network, optimal feature selection, soil quality.

INTRODUCTION

Agricultural management is essential for ensuring the productivity and sustainability of farming systems. It involves the control, monitoring, and planning of agricultural activities, including resource allocation, irrigation, pest management, livestock care, and crop production (Li *et al*., 2025). Effective agricultural management is necessary to meet the growing demand for food, address environmental transformation, and protect natural resources (Song *et al*., 2024). Agricultural yields can be enhanced by optimizing farming practices while minimizing costs and environmental impacts (Safaie *et al*., 2023).

One of the most important aspects of agricultural management is soil quality management. Soil is a fundamental component of productive agriculture because it provides essential nutrients that support crop production (Barathkumar *et al*., 2025). Soil with higher organic matter content and an appropriate nutrient composition promotes the development of healthier plants, thereby improving crop productivity. In contrast, poor-quality soil limits crop productivity due to nutrient deficiencies and increases crop susceptibility to pests and diseases (Huang *et al*., 2023).

Soil quality is also critical for the long-term sustainability of farming. Poor irrigation management, excessive fertilizer use, and over-farming negatively affect soil quality through fertility loss, erosion, and soil degradation (Raza *et al*., 2023). Practices such as reduced tillage, cover cropping, organic farming, and crop rotation contribute to improved agricultural management by preserving and enhancing soil health (Omondiagbe *et al*., 2023). Effective agricultural management helps maintain soil fertility, enhance production, reduce the need for chemical inputs, and support sustainable agricultural productivity (Ghani *et al*., 2024).

Traditional soil evaluation methods involve the manual collection of soil samples from various locations, followed by chemical and physical analyses in laboratories to assess soil health based on the obtained results (Chaudhry *et al*., 2024). Accordingly, traditional soil quality forecasting relies on field observations, laboratory testing, physical soil sampling, historical data analysis, geographic surveys, and soil mapping to evaluate soil characteristics such as moisture levels, nutrient content, and pH levels, as well as to predict soil quality across larger areas (Suwardi *et al*., 2023; Pant *et al*., 2024). However, these methods face major challenges because they are expensive, labor-intensive, and time-consuming, particularly when large geographical areas are considered (Parewai and Köppen, 2024). In addition, soil properties can vary considerably over short distances, making it difficult to accurately capture soil characteristics across different regions (El Behairy *et al*., 2024). Spatial variability and the lack of real-time data further reduce the efficiency of traditional soil quality forecasting methods for continuous soil monitoring and agricultural management (Divya *et al*., 2024).

Deep learning-based methods for predicting soil quality provide an effective solution to many of the challenges associated with conventional techniques by using advanced algorithms and large datasets to achieve greater speed and accuracy (Chen *et al*., 2023). Using historical records, remote sensing data, and satellite imagery, deep learning

approaches can analyze complex soil characteristics without requiring extensive manual sampling (Selvanarayanan *et al*., 2024). In addition, these models are highly scalable, enabling more efficient soil quality assessment across large geographical regions (Shahzad *et al*., 2024). However, certain limitations remain. The accuracy of deep learning models depends heavily on the availability and quality of the training data (Shahare *et al*., 2024). Poor-quality or limited datasets may result in unreliable predictions, thereby increasing the complexity of agricultural management (Kalyani and Kolla, 2024). To address these limitations, a deep learning-based framework for forecasting soil quality with optimal feature selection was proposed in this study (Sumathi *et al*., 2023).

The primary objective of this research was to develop an efficient framework for forecasting soil quality that can improve prediction accuracy and convergence performance. For this purpose, the Adaptive Parrot Optimization (AdPo) algorithm was designed for feature selection to enhance the convergence rate in identifying the global optimal solution. In this approach, adaptive weighting was incorporated into the exploration stage of the parrot optimization algorithm to improve convergence efficiency during the search process. Furthermore, soil quality forecasting was performed using the proposed Extended Cross Stage Pyramid Network (ExCSP_Net) to improve forecasting accuracy. In the ExCSP_Net architecture, a gated recurrent unit (GRU)-based attention module was integrated into the main path of the Cross Stage Partial (CSP) model to capture important feature dependencies, while a stacked autoencoder was incorporated before the feature-sharing stage in the short path of the CSP model to reduce feature dimensionality and generate meaningful representations.

MATERIALS AND METHODS

Soil quality forecasting using optimal feature selection and deep learning-based prediction was introduced in this study (Figure 1). Initially, the soil data were acquired

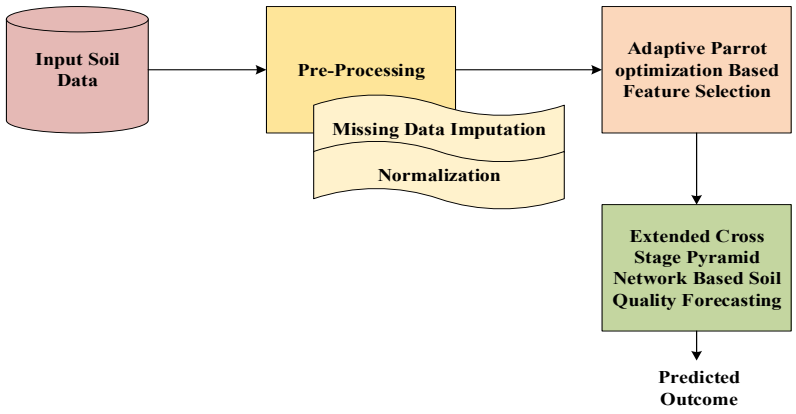

Figure 1. General workflow of the proposed soil quality forecasting framework.

from a publicly available dataset and pre-processed using missing data imputation and normalization techniques. Subsequently, significant features were selected from the preprocessed data using the Adaptive Parrot Optimization (AdPo) algorithm. Finally, soil quality prediction was performed using the Extended Cross Stage Pyramid Network (ExCSP_Net). For the development of ExCSP_Net, a gated recurrent unit (GRU)-based attention module was incorporated into the main path of the Cross Stage Partial (CSP) model, and a stacked autoencoder was integrated before the feature-sharing stage in the short path of the CSP model to minimize feature dimensionality.

Data acquisition

The input data for the proposed soil quality forecasting framework were obtained from the publicly available Soil Fertility Dataset available on Kaggle (https://www.kaggle.com/datasets/rahuljaiswalonkaggle/soil-fertility-dataset). The dataset contained 38 600 soil samples with multiple numerical attributes, including pH, nitrogen (N), phosphorus (P), potassium (K), organic carbon, moisture content, electrical conductivity, and other fertility-related indicators, resulting in a total of N predictor variables (Oukhattar *et al.,* 2025). The response variable corresponded to soil fertility status, representing the soil quality class indicating soil suitability for crop growth (Ziyadullaev *et al.,* 2024).

The proposed Adaptive Parrot Optimization with Extended Cross Stage Pyramid Network (AdPo+ExCSP_Net) model was implemented in the Python programming language on a Windows 10 operating system with 8 GB RAM. Conventional approaches, including Artificial Neural Network (ANN), Extreme Gradient Boosting (XGBoost), Random Forest (RF), Improved Soil Quality Prediction-Deep Learning (ISQP-DL), and the proposed model without AdPo, were used for comparative analysis. Model performance was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), coefficient of determination (R^2), processing time, F-score, recall, precision, and accuracy.

Pre-processing

Pre-processing based on missing data imputation and normalization was applied in the proposed soil quality forecasting model. Missing data imputation was performed to enable the effective use of incomplete datasets by estimating missing values while preserving the overall structure and richness of the data, preventing information loss and improving prediction accuracy. Missing values were estimated using the K-Nearest Neighbors (KNN)-based imputation method, in which predictions were generated from the values of the nearest neighbors in the dataset. The KNN-based approach considered the relationships among features and provided accurate imputation for soil data. The steps involved in KNN-based missing data imputation were as follows.

Step 1: The Euclidean distance between data points was calculated as:

$$D(X_u, X_v) = \sqrt{\sum_{s=1}^{S} \left(X_{u,s} - X_{v,s}\right)^2}$$

where the data points are represented as $X_{u,s}$ and $X_{v,s}$, respectively, the total number of data points is represented as S, and the distance between them is denoted as $D(X_u, X_v)$. The S nearest neighbors were selected based on the calculated distance.

Step 2: Missing data imputation was performed by estimating the mean of the selected S data points. Following missing data imputation, normalization was performed to transform the data into a standard form.

Normalization

Z-score-based normalization was applied to standardize the data and was expressed as:

$$I_z = \frac{x - Avg}{SD}$$

where I_z represents the normalized output, Avg is the average value, SD is the standard deviation, and x is the input value. Significant feature selection was subsequently performed from the preprocessed data using the proposed AdPo algorithm.

Feature selection

AdPo-based feature selection was used to identify the optimal feature subset from the pre-processed data. The proposed AdPo algorithm was designed to improve the convergence rate of the optimization process in identifying the global optimal solution. In this approach, adaptive weighting was incorporated into the exploration stage of the parrot optimization algorithm to enhance convergence efficiency during the search process. Forecasting accuracy was considered a fitness function for selecting the optimal feature subset. The population size, represented by L, denotes the number of parrots (solutions) in the swarm, whereas $Ymax$ is the maximum number of iterations used to update the positions of the solutions.

The initial position of the parrot was calculated as:

$$P_r^0 = D + u(0,1) \cdot (X - D)$$

where D and X are the lower and upper boundaries of the search space, respectively. The term $u(0,1)$ is a continuous uniform random variable within the interval (0,1),

which introduces randomness during initialization, and P_r^0 represents the initial position of the *r*th parrot.

The position updating process was expressed as:

$$P_r^{Y+1} = (P_r^Y - P_{best}) \cdot Q(n) + u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right)^{2 \cdot \frac{y}{Y_{max}}} \cdot P_{Avg}^Y$$

where P_r^Y represents the current position of the *r*th parrot, $P_r^{(Y+1)}$ is the updated position after iteration, and P_{best} indicates the best position identified by the swarm, corresponding to the optimal solution obtained thus far. The term *Q(n)* represents the Lévy flight distribution, which enables long exploratory movements during the search process and facilitates the identification of new solution regions. The variable $u(0,1)$ is a continuous uniform random variable within the interval (0,1), y represents the current iteration, and Y_{max} indicates the maximum number of iterations. In addition, P_{Avg}^Y is the average position of the swarm population, which guided the solutions according to the general movement trend of the swarm.

The solution updating mechanism of the proposed AdPo algorithm consisted of two major stages, namely exploration and exploitation, represented by $(P_r^Y - P_{best}) \cdot Q(n)$ and $u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right)^{2 \cdot \frac{y}{Y_{max}}} \cdot P_{Avg}^Y$, respectively.

The average position of the swarm at iteration Y was calculated as:

$$P_{Avg}^Y = \frac{1}{L} \sum_{d=1}^{L} P_{dY}$$

where *L* is the total population size and P_{dY} is the position of the *d*th parrot at iteration Y.

Lévy flight

This approach was used as a random walk mechanism in which most movements were small, while occasional large jumps allowed extensive exploration of the search space. The Lévy flight distribution was defined as:

$$Q(n) = \frac{\delta \cdot \varepsilon}{|s|}$$

where $\delta \sim N(0,n)$ and $s \sim N(0,n)$ represent normally distributed random variables, and ε was calculated as:

$$\varepsilon = \left(\frac{\Gamma(1+\lambda) \cdot sin\left(\frac{\pi \cdot \lambda}{2}\right)}{\Gamma\left(\frac{1+\lambda}{2}\right) \cdot \lambda \cdot 2^{\frac{1}{\lambda}}} \right)^{\frac{1}{\lambda}}$$

where λ represents the parameter controlling the shape of the distribution and was set to 1.5. Lévy flights were effective for both local and global search processes, thereby enabling the optimization algorithm to explore both nearby and distant regions of the solution space.

To improve the foraging behavior of the parrots, an adaptive weighting strategy was incorporated and expressed as:

$$Ad_w = P_r^{Y+1} + F\left((P_r^Y)_{good} - (P_r^Y)\right)$$

where Ad_w represents the adaptive weighting factor, $P_r^{(Y+1)}$ denotes the solution obtained in the current iteration, and P_r^Y is the solution obtained by the parrot search agent in the previous iteration. The term $(P_r^Y)_{good}$ indicates the best solution identified in the previous iteration, while F is the controlling parameter.

After incorporating adaptive weighting, the updated solution of the proposed AdPo algorithm was formulated as:

$$(P_r^{Y+1})_{AdPo} = Ad_w * (P_r^{Y+1})_{parrot}$$

Substituting the adaptive weighting expression into the position updating process yielded:

$$(P_r^{Y+1})_{AdPo} = \left[P_r^{Y+1} + F\left((P_r^Y)_{good} - (P_r^Y)\right)\right] *$$

$$\left[(P_r^Y - P_{best}) \cdot Q(n) + u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right)^{2 \cdot \frac{y}{Y_{max}}} \cdot P_{Avg}^Y\right]$$

The modified solution enhanced the exploration capability of the optimization process while improving convergence speed for identifying the global optimal solution.

Staying behavior. This represents the tendency of parrots to temporarily remain near their owner after movement. This behavior was modeled as a sudden movement toward the owner followed by a stationary phase and was expressed as:

$$P_r^{Y+1} = P_r^Y + P_{best} \cdot Q(n) + u(0,1) \cdot ones(1,n)$$

where $P_r^{(Y+1)}$ represents the updated position of the rth parrot, P_r^Y denotes the current position of the rth parrot, and $P_{best} \cdot Q(n)$ is the movement toward the owner modeled using the Lévy distribution. The term $u(0,1) \cdot ones(1,n)$ represents the random stopping behavior near the owner.

Communicating behavior. This approach was used to model the natural social interactions of parrots, which are highly social animals. This allows individual parrots to exchange information and update their positions according to interactions within the swarm, thereby improving optimization performance. The communication process, either through movement toward the flock or through remote interaction, was modeled as:

$$P_r^{Y+1} = \begin{cases} 0.2 \cdot u(0,1) \cdot \left(1 - \dfrac{y}{Y_{max}}\right) \cdot \left(P_r^Y - P_{Avg}\right) & if\ Q \leq 0.5 \\ 0.2 \cdot u(0,1) \cdot exp\left(\dfrac{y}{u(0,1) \cdot Y_{max}}\right) & if\ Q > 0.5 \end{cases}$$

where $P_r^{(Y+1)}$ represents the updated position of the rth parrot at iteration Y+1, and P_r^Y is the current position of the rth parrot at iteration Y. The term $u(0,1)$ generates a random number between 0 and 1 to introduce variability and randomness into the communication process. In addition, Q is a random value between 0 and 1 that determines whether the parrot moved toward the flock or communicated without movement.

Fear of strangers. This behavior was incorporated to simulate the natural tendency of parrots to avoid unfamiliar environments and remain close to safe regions. It ensures that the optimization process avoided unknown or less optimal regions of the search space and moved toward safer and more optimal solutions. The fear of strangers behavior was expressed as:

$$P_r^{Y+1} = P_r^Y + u(0,1) \cdot cos\left(0.5\pi \cdot \frac{y}{Y_{max}}\right) - cos(u(0,1) \cdot \pi) \cdot \left(\frac{y}{Y_{max}}\right)^2 \cdot (P_r^Y - P_{best})$$

where $u(0,1) \cdot cos\left(0.5\pi \cdot \dfrac{y}{Y_{max}}\right)$ describes the movement of the parrot toward a safe region. The term $cos(u(0,1) \cdot \pi) \cdot \left(\dfrac{y}{Y_{max}}\right)^2 \cdot (P_r^Y - P_{best})$ models the avoidance of unfamiliar or unfavorable regions in the search space. The variable $u(0,1)$ represents a random value within the interval [0,1], P_{best} is the best position identified by the swarm,

y is the current iteration, Y_{max} is the maximum number of iterations, and P_r^Y indicates the current position of the rth parrot.

Feasibility estimation

The feasibility of the proposed AdPo algorithm was evaluated based on forecasting accuracy and was calculated as:

$$Fit = \frac{TP + TN}{TP + TN + FP + FN}$$

where *Fit* represents the fitness value, *TP* denotes the true positives, *TN* the true negatives, *FP* the false positives, and *FN* the false negatives.

Termination

The algorithm terminated when the global optimal solution was achieved or when the maximum number of iterations was completed. Based on the solution obtained using the AdPo algorithm (Algorithm 1), the optimal feature subset was selected to eliminate redundant features.

Algorithm 1. Pseudo-code of the proposed Adaptive Parrot Optimization (AdPo) algorithm for optimal feature selection.

Start
Initialize the parameters
Initialize the position of parrot search agent
Estimate the fitness for parrot search agent
Sort and update the solutions
Choose the strategy randomly
{
Update the solution based on foraging behavior

$$(P_r^{Y+1})_{AdPo} = \left[P_r^{Y+1} + F\left((P_r^Y)_{good} - (P_r^Y)\right)\right] * \left[(P_r^Y - P_{best}) \cdot Q(n) + u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right)^{2 \cdot \frac{y}{Y_{max}}} \cdot P_{Avg}^Y\right]$$

Update the solution based on staying behavior

$$P_r^{Y+1} = P_r^Y + P_{best} \cdot Q(n) + u(0,1) \cdot ones(1,n)$$

Update the solution based on communicating behavior

$$P_r^{Y+1} = \begin{cases} 0.2 \cdot u(0,1) \cdot \left(1 - \frac{y}{Y_{max}}\right) \cdot \left(P_r^Y - P_{Avg}\right) & if\ Q \leq 0.5 \\ 0.2 \cdot u(0,1) \cdot exp\left(\frac{y}{u(0,1) \cdot Y_{max}}\right) & if\ Q > 0.5 \end{cases}$$

Algorithm 1. Continued

Update the solution based on fear of strangers

$$P_r^{Y+1} = P_r^Y + u(0,1) \cdot cos\left(0.5\pi \cdot \frac{y}{Y_{max}}\right) - cos(u(0,1) \cdot \pi) \cdot \left(\frac{y}{Y_{max}}\right)^2 \cdot (P_r^Y - P_{best})$$

}
Check the feasibility
Check $y > Y^{max}$
$Y = y + +$
Return the best solution
end

Extended cross stage pyramid network-based soil quality forecasting

Soil quality forecasting was performed using ExCSP_Net to improve forecasting accuracy (Figure 2). For the development of ExCSP_Net, a GRU-based attention module was incorporated into the main path of the CSP model, and a stacked autoencoder was integrated before the feature-sharing stage in the short path of the CSP model to minimize feature dimensionality. The CSP network improved the learning efficiency of deep neural networks by reducing computational complexity and enhancing gradient flow. The CSP model operated by dividing feature maps into two parts across different stages and subsequently merging them within the network to improve feature representation learning. In addition, it combined pyramid structures with cross-stage feature fusion to process multiple scales in the input data for soil quality forecasting.

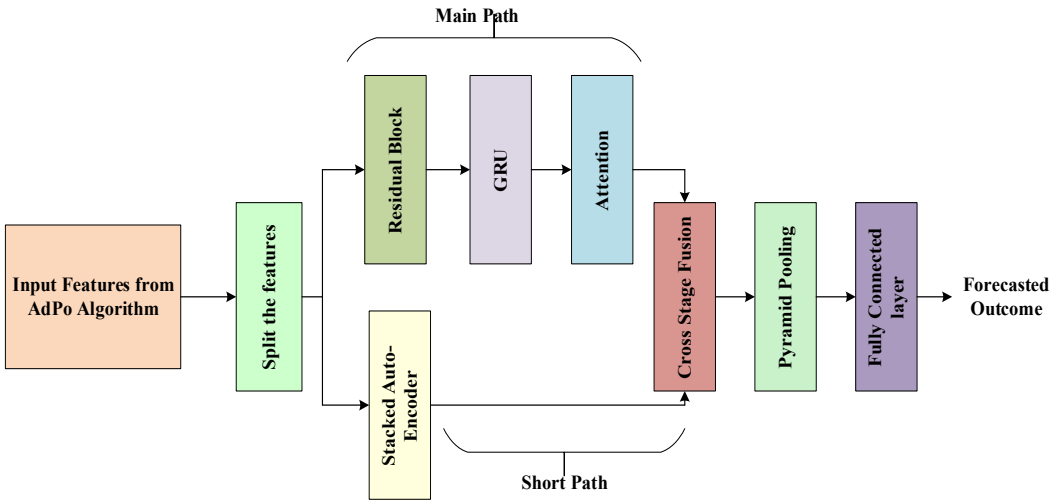

Figure 2. Architecture of the proposed Extended Cross Stage Pyramid Network (ExCSP_Net) for soil quality forecasting.

The features selected by the AdPo algorithm E were input into the ExCSP_Net model's input layer and the feature extraction module designed with both main and short paths. Initially, the features were split to facilitate their passage through these main and short paths. The splitting of features was expressed as:

$$E_M, E_S = Split(E)$$

where *EM* represents the features forwarded to the main path, and *ES* represents the features provided to the short path of the ExCSP_Net model.

The main path of the ExCSP_Net model consisted of multiple residual blocks for extracting complex features. In the proposed architecture, a GRU_Attn layer was integrated with the residual blocks to extract spatial and temporal features for improving forecasting accuracy. Initially, the features from the main path were provided to the residual block for spatial feature extraction. The residual block utilized residual connections to address the vanishing gradient problem and enabled effective training of very deep neural networks. The residual block was defined as:

$$B(e) = C(e, \{D_r\}) + e$$

where $B(e)$ represents the output of the residual block, $C(e,\{D_r\})$ is the output of the convolutional layers within the block corresponding to the learned residual representation, and e is the input feature map provided through the skip connection. The extracted features were then sent to the activation layer, which introduced non-linearity into the learned representations. The use of residual connections allowed these features to propagate through deeper layers without losing the original spatial information. As a result, the residual blocks enabled the network to maintain fine-grained spatial structures from earlier layers while also learning higher-level feature abstractions.

The deep features extracted through the residual block were subsequently provided to the GRU module for temporal feature extraction, enabling the acquisition of long-term dependent features. In soil quality prediction, temporal feature extraction using GRU facilitated the modeling of temporal relationships among measurements collected over time, making it suitable for applications in which soil quality changed dynamically (Babu *et al.,* 2024). Thus, efficient temporal features were extracted with minimal computational complexity using the proposed ExCSP_Net model.
The output generated from the residual block was provided to the input layer of the GRU. The input sequence was represented as

$$L = \{l_1, l_2, \ldots l_f\}$$

where l_f represents the input at time step f. During sequence processing, the GRU computed the hidden state Y_f at each time step. The hidden state estimation was performed through the processing of the reset gate, update gate, and candidate state. The reset gate controlled the retention of relevant information by eliminating unnecessary features and was expressed as:

$$G_f = \varepsilon\left(B_G \cdot \left[Y_{f-1}, l_G\right]\right)$$

where B_G represents the weight associated with the reset gate, l_G is the bias associated with the reset gate, and ε is the sigmoid activation function.
The output of the update gate was calculated as:

$$Z_f = \varepsilon\left(B_Z \cdot \left[Y_{f-1}, l_Z\right]\right)$$

where B_Z is the weight associated with the update gate and l_Z is the bias associated with the update gate.

Subsequently, the candidate state was estimated as:

$$\tilde{Y}_f = tanh\left(B_Y \cdot \left[G_Y \circ Y_{f-1}, l_Y\right] + Q_Y\right)$$

where B_Y is the weight associated with the candidate activation, ° denotes element-wise multiplication, l_Y is the bias associated with the candidate activation, and Q_Y indicates the additional learnable parameter (Genova *et al.*, 2024).

The hidden state of the GRU was then updated as:

$$Y_f = H_f \circ Y_{f-1} + \left(1 - H_f\right)\tilde{Y}_f$$

where Y_f represents the updated hidden state, H_f is the update control parameter, $Y_{(f-1)}$ is the previous hidden state, and $\tilde{Y}_f$ corresponds to the candidate hidden state.

The GRU allowed the proposed model to capture both short-term and long-term dependencies within soil quality data while maintaining low computational complexity. In addition, it mitigated the vanishing gradient problem and allowed the model to learn from long soil data sequences without losing important information.
Following temporal feature extraction, an attention mechanism was incorporated to assign weights to significant features for improving soil quality forecasting

performance. The attention score for the extracted features from the hidden layer of the GRU was calculated as:

$$p_f = q \cdot tanh\left(B_Y Y_f + Q_Y\right)$$

where q is the learnable vector, B_Y and Q_Y denote the weight and bias parameters, respectively, and p_f represent the attention score.

After computing the attention scores, the weights γ_f were assigned using the SoftMax function and expressed as:

$$\gamma_f = \frac{exp\left(p_f\right)}{\sum_{d=1}^{F} exp(p_d)}$$

The feature vector generated from the main path of the proposed ExCSP_Net model was defined as:

$$n = \sum_{f=1}^{F} \gamma_f Y_f$$

where n is the combined feature representation obtained from all input sequences. Thus, the output of the main path of the proposed ExCSP_Net model consisted of the most relevant features extracted with minimal computational complexity.

The short path of the ExCSP_Net model utilized the output of the stacked autoencoder for further processing to extract significant features with reduced dimensionality. The encoder compressed the input data into a lower-dimensional representation through multiple fully connected layers. The output of each encoder layer was expressed as:

$$K(r) = R\left(S(r) \cdot K(r-1) + v(r)\right)$$

where $K(r)$ is the output of the rth encoder layer, $S(r)$ denotes the weight matrix, $v(r)$ represents the bias term, R indicates the activation function, and $K(r-1)$ corresponds to the output of the previous layer.

At the center of the stacked autoencoder, a latent representation containing the most salient soil features was generated as:

$$x = R\left(S(x) \cdot K(r) + v(x)\right)$$

where x represents the latent feature vector, while $S(x)$ and $v(x)$ are the associated weights and bias parameters, respectively.

The decoder reconstructed the input data from the latent representation while preserving the significant features of the original input. The decoder output was expressed as:

$$K'(r) = R\big(S(r) \cdot K'(r-1) + v(r)\big)$$

where $K'(r)$ represents the output of the rth decoder layer, and $K'(r-1)$ denotes the output of the previous decoder layer.

The dimensionality-reduced features obtained using the stacked autoencoder were subsequently forwarded through the short path of the ExCSP_Net model. Cross-stage fusion was then performed on the features generated from both the main and short paths of the proposed ExCSP_Net model. The cross-stage fusion mechanism ensured that gradient information from both paths contributed to weight updating while reducing redundant gradients. The output of the fusion layer was expressed as:

$$E_{fused} = Concat(E_M, E_S)$$

where E_{fused} is the merged feature representation, and $Concat(\cdot)$ is the concatenation function.

Subsequently, pyramid pooling was performed to process multi-scale feature representations. The pyramid pooling operation was formulated as:

$$E_p = Concat\left(Pool\big(E_{fused}, y_1\big), Pool\big(E_{fused}, y_2\big), \ldots Pool\big(E_{fused}, y_z\big)\right)$$

where E_p represents the output of the pyramid pooling layer, and $y_1, y_2, \ldots, y_z$ are different pooling kernel sizes used for multi-scale feature extraction.
After the pooling operation, a fully connected network was used for soil quality prediction. The fully connected network consisted of multiple convolution layers, and its output corresponded to the predicted soil quality based on the input soil data. The output of the fully connected layer was formulated as:

$$Out = Conv\big(E_p, S\big)$$

where Out represents the predicted output, E_p is the output obtained from pyramid pooling, $Conv(\cdot)$ is the convolution operation, and S indicates the weight parameters. Thus, the proposed ExCSP_Net model enabled more accurate soil quality prediction.

RESULTS AND DISCUSSION

The pre-processing techniques, including missing data imputation and normalization, improved the quality of the input data by reducing noise and inconsistencies, contributing to higher prediction accuracy (Figure 3). The AdPo algorithm ensured that only the most relevant features were selected, which improved the decision-making capability of the model and further enhanced prediction accuracy. In addition, the ExCSP_Net architecture, incorporating GRU-based attention and stacked autoencoders, captured complex patterns in soil data more effectively, improving the overall prediction performance.

The stacked autoencoder incorporated within the ExCSP_Net model performed dimensionality reduction without losing critical information, minimizing overall prediction errors and reducing MAE and RMSE values. In addition, the combination

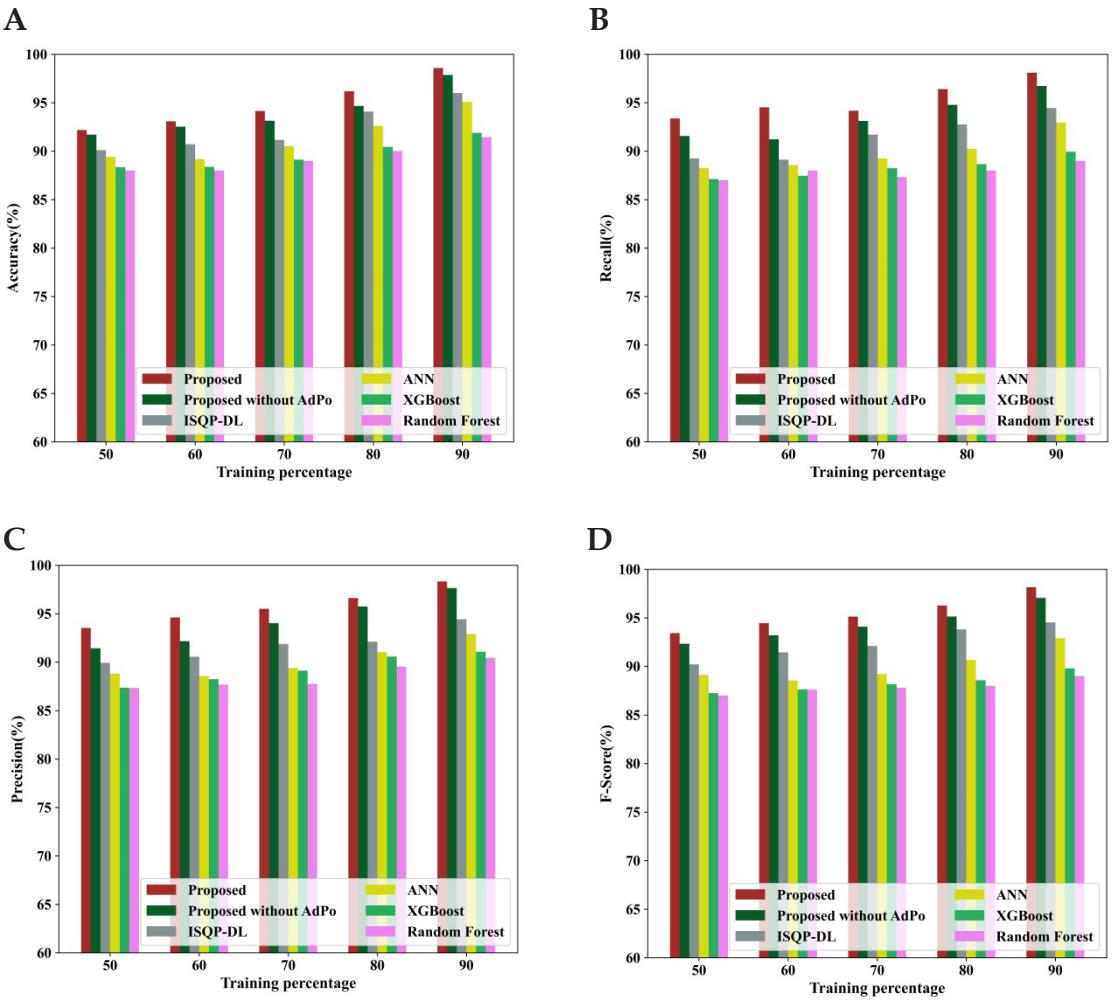

Figure 3. Performance evaluation of soil quality forecasting models based on A: accuracy; B: recall; C: precision; D: F-score.

of AdPo-based feature selection and ExCSP_Net-based soil quality forecasting enabled the model to capture complex relationships between input features and soil quality more effectively, resulting in a higher R^2 value (Figure 4).

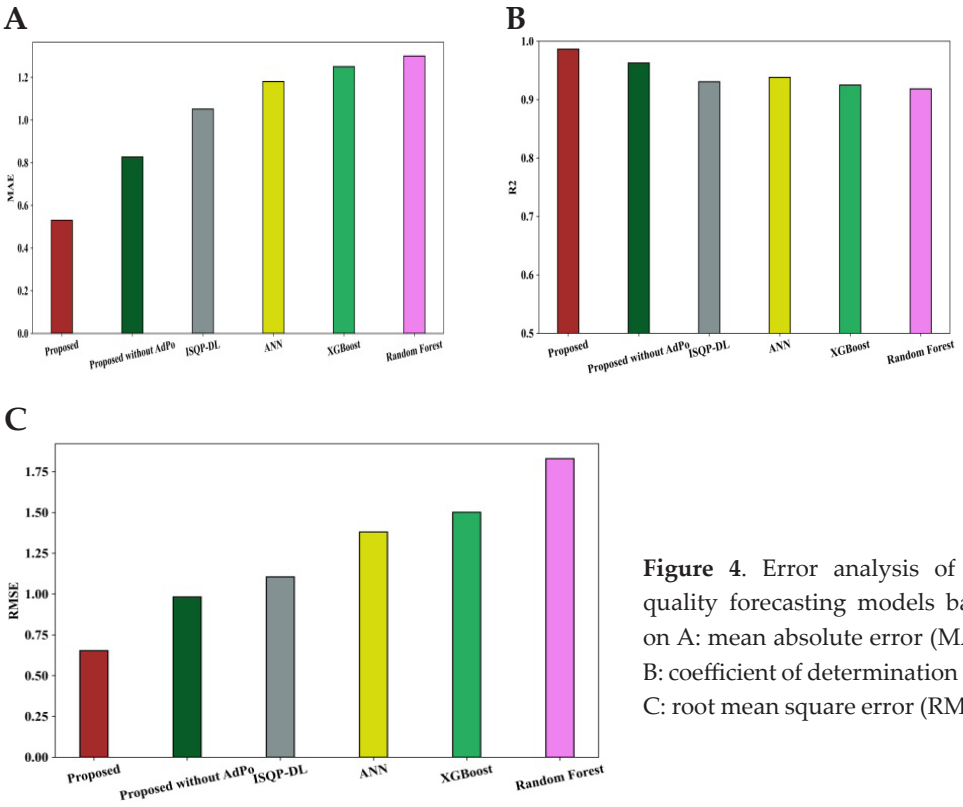

Figure 4. Error analysis of soil quality forecasting models based on A: mean absolute error (MAE); B: coefficient of determination (R^2); C: root mean square error (RMSE).

K-fold validation was performed to demonstrate the generalization capability of the proposed AdPo + ExCSP_Net model (Figure 5). In this approach, the dataset was partitioned into K mutually exclusive and equally sized folds. During each iteration, K−1 folds were used for training the model, while the remaining fold was utilized for validation. The proposed model demonstrated improved performance across all K-folds.

Confusion matrix analysis was performed to quantitatively evaluate the predictive performance of the proposed soil quality forecasting model (Figure 6). The confusion matrix metrics reflected the capability of the model to accurately identify soil quality classes while minimizing incorrect predictions, validating the effectiveness of the AdPo-based feature selection approach and the ExCSP_Net architecture.

SHapley Additive exPlanations (SHAP) analysis was performed (Figure 7) to demonstrate the interpretability of the model by quantifying the contribution of each soil feature selected by the AdPo algorithm to soil quality prediction.

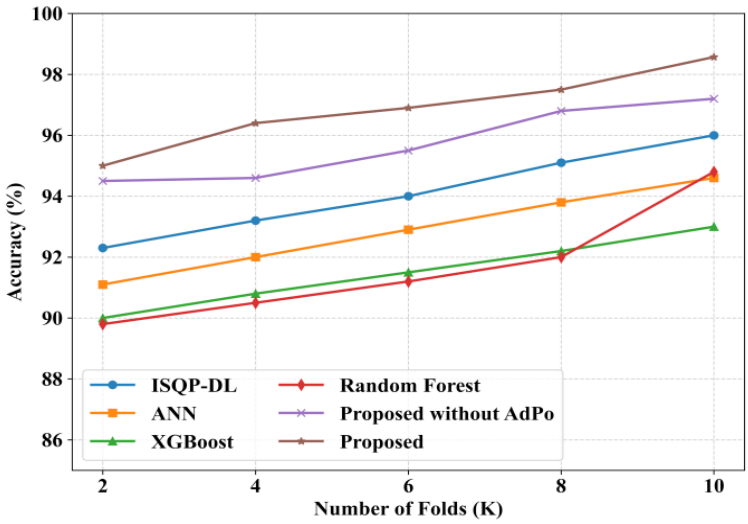

Figure 5. K-fold cross-validation performance of the proposed Adaptive Parrot Optimization with Extended Cross Stage Pyramid Network (AdPo + ExCSP_Net) soil quality forecasting model.

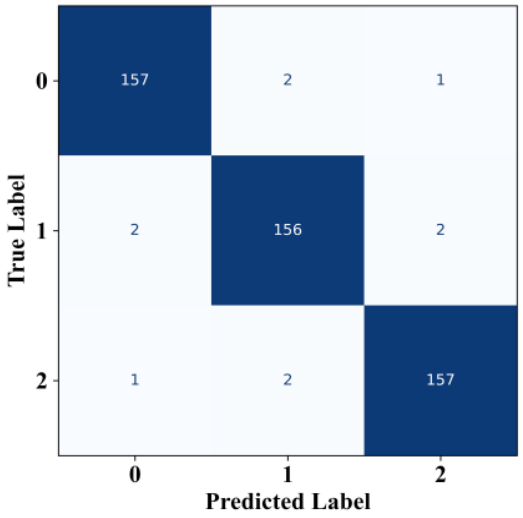

Figure 6. Confusion matrix of the proposed Adaptive Parrot Optimization with Extended Cross Stage Pyramid Network (AdPo+ExCSP_Net) model for soil quality forecasting.

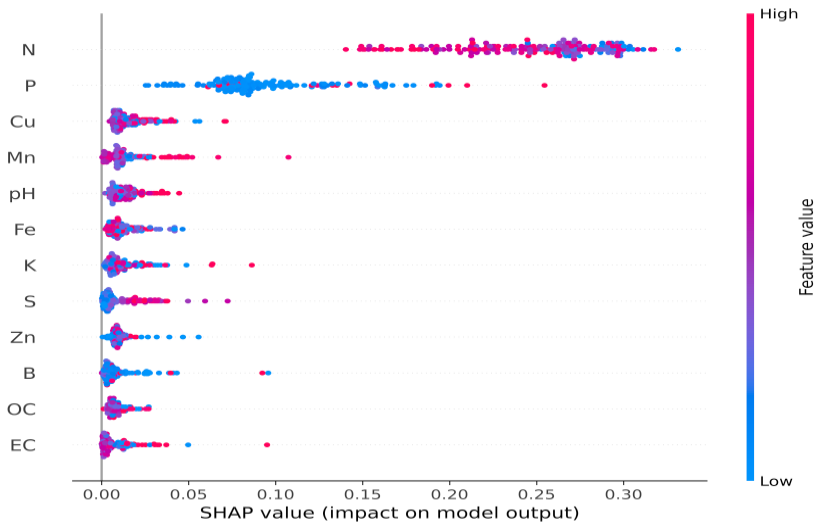

Figure 7. SHapley Additive exPlanations (SHAP)-based analysis of feature contributions in the proposed Adaptive Parrot Optimization with Extended Cross Stage Pyramid Network (AdPo+ExCSP_Net) model.

The proposed AdPo+ExCSP_Net model achieved an accuracy of 98.58 %, which was 2.61, 3.56, 6.78, 7.23, and 0.72 % higher than those obtained using the ISQP-DL, ANN, XGBoost, Random Forest, and proposed model without AdPo methods, respectively (Table 1). Similarly, the proposed method demonstrated improved performance across all evaluation metrics. Thus, the analysis portrays the superiority of the proposed AdPo-based feature selection and ExCSP_Net-based soil quality forecasting.

Table 1. Comparative performance evaluation of soil quality forecasting models based on the best obtained results.

Metrics/ Methods	ISQP-DL	ANN	XGBoost	Random Forest	Proposedmodel without AdPo	Proposed model
Accuracy	96.01	95.07	91.89	91.45	97.87	98.58
Recall	94.45	92.95	89.94	89.00	96.71	98.09
Precision	94.44	92.89	91.07	90.43	97.65	98.32
F-score	94.53	92.89	89.79	89.00	97.07	98.15
MAE	1.05	1.18	1.25	1.30	0.83	0.53
RMSE	1.11	1.38	1.50	1.83	0.98	0.65
R^2	0.93	0.94	0.93	0.92	0.96	0.99

ISQP-DL: improved soil quality prediction-deep learning; ANN: artificial neural network; XGBoost: extreme gradient boosting; MAE: mean absolute error; RMSE: root mean square error; R^2: coefficient of determination.

CONCLUSIONS

This research introduced a new model for forecasting soil quality based on optimal feature selection and deep learning. The proposed Adaptive Parrot Optimization (AdPo) algorithm was designed to perform optimal feature selection from the preprocessed data. The AdPo algorithm enhanced the convergence rate for identifying the global optimal solution through the incorporation of an adaptive weighting strategy, obtaining efficient and non-redundant features.

Soil quality forecasting was performed using the proposed Extended Cross Stage Pyramid Network (ExCSP_Net) model. The developed framework improved forecasting performance through enhanced short and main paths capable of extracting spatial-temporal features with reduced computational complexity. The integration of AdPo with ExCSP_Net demonstrated the effectiveness of combining optimal feature selection with deep learning for accurate soil quality forecasting and sustainable agricultural management. Despite the promising performance of the proposed framework, the computational burden remains a limitation. Therefore, future research will focus on the development of a lightweight model to further improve computational efficiency and practical applicability.

REFERENCES

Babu CS, Yadavamuthiah K. 2024. Soil quality prediction using deep learning. *In* Kumar R, Abdul Hamid AB, Ya'akub DNIB, Ya'akub NIB, Sharan HO, Kumar S. (eds.), Sustainable Development in AI, Blockchain, and E-Governance Applications. IGI Global Scientific Publishing: Hershey, PA, USA, pp: 171–188. https://doi.org/10.4018/979-8-3693-1722-8.ch010

Barathkumar S, Sellamuthu KM, Sathyabama K, Malathi P, Kumaraperumal R, Devagi P. 2025. Advancements in soil quality assessment: A comprehensive review of machine learning and AI-Driven approaches for nutrient deficiency Analysis. Communications in Soil Science and Plant Analysis 56 (2): 251–276. https://doi.org/10.1080/00103624.2024.2406484

Chaudhry H, Vasava HB, Chen S, Saurette D, Beri A, Gillespie A, Biswas A. 2024. Evaluating the soil quality index using three methods to assess soil fertility. Sensors 24 (3): 864. https://doi.org/10.3390/s24030864

Chen X, Zhang X, Wei Y, Zhang S, Cai C, Guo Z, Wang J. 2023. Assessment of soil quality in a heavily fragmented micro-landscape induced by gully erosion. Geoderma 431: 116369. https://doi.org/10.1016/j.geoderma.2023.116369

Divya A, Josphineleela R, Jaba Sheela L. 2024. A machine learning based approach for prediction and interpretation of soil properties from soil spectral data. Journal of Environmental Biology 45 (1): 96–105. https://doi.org/10.22438/jeb/45/1/mrn-5134

El Behairy RA, El Arwash HM, El Baroudy AA, Ibrahim MM, Mohamed ES, Rebouh NY, Shokr MS. 2024. An accurate approach for predicting soil quality based on machine learning in drylands. Agriculture 14 (4): 627. https://doi.org/10.3390/agriculture14040627

Genova G, Borruso L, Signorini M, Mitterer M, Niedrist G, Cesco S, Felderer B, Cavani L, Mimmo T. 2024. Analyzing soil enzymes to assess soil quality parameters in long-term copper accumulation through a machine learning approach. Applied Soil Ecology 195: 105261. https://doi.org/10.1016/j.apsoil.2023.105261

Ghani S, Thapa I, Adhikari DK, Waris KA. 2024. Soil categorization and liquefaction prediction using deep learning and ensemble learning algorithms. Transportation Infrastructure Geotechnology 12 (1). https://doi.org/10.1007/s40515-024-00474-8

Huang Y, Srivastava R, Ngo C, Gao J, Wu J, Chiao S. 2023. Data-driven soil analysis and evaluation for smart farming using machine learning approaches. Agriculture 13 (9): 1777. https://doi.org/10.3390/agriculture13091777

Kalyani NL, Kolla BP. 2024. Machine learning-based smart agricultural practices to assess soil fertility and nutrient dynamics. South Eastern European Journal of Public Health 24 (4): 1363–1378. https://doi.org/10.70135/seejph.vi.2873

Li Y, Yao G, Li S, Dong X. 2025. Predicting and mapping of soil organic matter with machine learning in the black soil region of the southern northeast plain of China. Agronomy 15 (3): 533. https://doi.org/10.3390/agronomy15030533

Omondiagbe OP, Lilburne L, Licorish SA, MacDonell SG. 2023. Soil texture prediction with automated deep convolutional neural networks and population-based learning. Geoderma 436: 116521. https://doi.org/10.1016/j.geoderma.2023.116521

Oukhattar M, Gadal S, Robert Y, Saby N, Houmma IH, Keller C. 2025. Variability analysis of soil organic carbon content across land use types and its digital mapping using machine learning and deep learning algorithms. Environmental Monitoring and Assessment 197 (5). https://doi.org/10.1007/s10661-025-13972-0

Pant J, Joshi MC, Singh D, Pant HK, Bhatt A, Pant D. 2024. An investigation into ensemble learning techniques for evaluating soil fertility through analytical approaches. Procedia Computer Science 235: 1998–2008. https://doi.org/10.1016/j.procs.2024.04.189

Parewai I, Köppen M. 2025. A digital twin approach for soil moisture measurement with physically based rendering simulations and machine learning. Electronics 14 (2): 395. https://doi.org/10.3390/electronics14020395

Raza I, Zubair M, Zaib M, Khalil MH, Haidar A, Sikandar A, Ashfaq MA. 2023. Precision nutrient application techniques to improve soil fertility and crop yield: A review with future prospect. International Research Journal of Educational and Technology 5 (8): 105–123.

Safaie M, Hosseinpour-Zarnaq M, Omid M, Sarmadian F, Ghasemi-Mobtaker H. 2023. Using deep neural networks for evaluation of soil quality based on VIS-NIR spectroscopy. Earth Science Informatics 17 (1): 271–281. https://doi.org/10.1007/s12145-023-01168-4

Selvanarayanan R, Rajendran S, Algburi S, Ibrahim Khalaf O, Hamam H. 2024. Empowering coffee farming using counterfactual recommendation based RNN driven IoT integrated soil quality command system. Scientific Reports 14 (1). https://doi.org/10.1038/s41598-024-56954-x

Shahare YR, Singh MP, Singh SP, Singh P, Diwakar M. 2024. ASUR: Agriculture soil fertility assessment using random forest classifier and regressor. Procedia Computer Science 235: 1732–1741. https://doi.org/10.1016/j.procs.2024.04.164

Shahzad A, Ibrahim M, Saleem RM, Zia S, Habib S, Mahmood M. 2024. Unveiling soil fertility patterns via image analysis and machine learning for accurate crop recommendations. Journal of Computing and Biomedical Informatics.

Song J, Shi X, Wang H, Lv X, Zhang W, Wang J, Li T, Li W. 2024. Improving soil quality index prediction by fusion of Vis-NIR and pXRF spectral data. Geoderma 447: 116938. https://doi.org/10.1016/j.geoderma.2024.116938

Sumathi P, Karthikeyan V, Kavitha M, Karthik S. 2023. Improved soil quality prediction model using deep learning for smart agriculture systems. Computer Systems Science and Engineering 45 (2): 1545–1559. https://doi.org/10.32604/csse.2023.027580

Suwardi, Sutiarso L, Wirianata H, Nugroho AP, Sukarman, Primananda S, Moch D, Hariadi B. 2023. Optimization of a soil type prediction method based on the deep learning model and vegetation characteristics. Plant Science Today 11 (1). https://doi.org/10.14719/pst.2926

Ziyadullaev D, Muhamediyeva D, Khujamkulova K, Abdurakhimov D, Maksumkhanova A, Ziyodullaeva G. 2024. Ensemble data mining methods for assessing soil fertility. E3S Web of Conferences 494: 02013. https://doi.org/10.1051/e3sconf/202449402013

Agrociencia

STANDARDIZED PRECIPITATION INDEX FORECASTING IN NORTH-CENTRAL MEXICO USING TRANSFORMER MODELS

Rafael **Magallanes-Quintanar**[1*], Carlos Eric **Galván-Tejada**[1], Jorge Issac **Galván-Tejada**[1], Santiago de Jesús **Méndez-Gallegos**[2], Antonio **García-Domínguez**[1]

[1]Universidad Autónoma de Zacatecas. Unidad Académica de Ingeniería Eléctrica. Avenida Ramón López Velarde 801, Colonia Centro, Zacatecas, Zacatecas, Mexico. C. P. 98000.
[2]Colegio de Posgraduados Campus San Luis Potosí. Iturbide 73, Salinas de Hidalgo, San Luis Potosí, Mexico. C. P. 78600.
* Author for correspondence: tiquis@uaz.edu.mx

Citation: Magallanes-Quintanar R, Galván-Tejada Ce, Galván-Tejada JI, Méndez-Gallegos SJ, García-Domínguez A. 2026. Standardized precipitation index forecasting in north-central mexico using Transformer models. **Agrociencia** 60(4): 586-608. https://doi.org/ 10.47163/agrociencia.v60i4.3641

Editor in Chief:
Dr. Fernando C. Gómez-Merino

Received: November 04, 2025.
Approved: June 05, 2026.
Published in Agrociencia: June 26, 2026.

ABSTRACT

Drought prediction is crucial for water resource management, agriculture, and climate adaptation in arid and semi-arid regions such as Zacatecas, Mexico. This study evaluates the advanced neural network architectures Long Short-Term Memory (LSTM), Vanilla Transformer, and Informer, for forecasting the Standardized Precipitation Index (SPI) using monthly precipitation data from 1964–2020 collected at 31 weather stations. SPI series were clustered into four regional climate zones. Models were implemented using the Nixtla *NeuralForecast* framework, and performance was assessed with Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Diebold-Mariano significance tests. The Informer model achieved the highest predictive accuracy, reducing average MSE by approximately 15 % relative to LSTM and consistently outper-forming the Vanilla Transformer in most regions. Statistical testing confirmed regional differences in model performance, suggesting that an adaptive, region-specific modeling approach is optimal for drought forecasting. These results demonstrate the robustness, efficiency, and transferability of Transformer-based models, particularly Informer, for operational drought monitoring under variable climatic conditions.

Keywords: Nixtla, Informer, Long Short-Term Memory, time-series, Standardized Precipitation Index, drought.

INTRODUCTION

Climate change poses a serious threat to global water resources. As the Earth's surface continues to warm, the future dynamics of precipitation and their effects on regional rainfall patterns remain uncertain (Ferreira *et al*., 2018). Furthermore, fluctuations in water availability, particularly the risk of future shortages, can adversely affect hydroelectric generation and agricultural operations (Karmalkar *et al*., 2011). Because historical rainfall and drought records are organized as time series, forecasting these variables represents an important challenge.

Time-series forecasting has been applied across numerous domains in which anticipating future trends based on historical data is essential. Key sectors benefiting from this approach include finance and economics (Siami-Namini and Namin, 2018), energy (Liu *et al*., 2023), healthcare (Kaushik *et al*., 2020), retail and e-commerce, manufacturing, and climate-related studies (Magallanes-Quintanar *et al*., 2022, 2024). In hydrology, for example, Long Short-Term Memory (LSTM) models have been successfully used for spatiotemporal forecasting of hydrological extremes, including drought events in river basins. Likewise, Transformer-based architectures have shown promise for national-scale drought forecasting across diverse climatic zones (Pathania and Gupta, 2025) and for hydrological drought prediction, with performance benchmarked against LSTM models (Amanambu *et al*., 2022).

In Mexico, droughts represent a significant threat to both water and food security. This concern is supported by the numerous droughts documented throughout the country's history (Florescano, 2000; García-Acosta *et al*., 2003), particularly in recent decades. For example, the most severe drought period of the last 70 years, in terms of its socio-economic impacts, occurred between 2011 and 2012. This event caused damage to approximately 1.5 million ha of cropland, the loss of more than 60 000 head of cattle, agricultural losses exceeding MXN 16 billion (approximately USD 1.3 billion), and widespread water shortages affecting up to 70 % of the national territory (Arreguín-Cortés *et al*., 2016).

Various methodologies have been developed to evaluate drought conditions, among which the Standardized Precipitation Index (SPI) is widely recognized for its simplicity and effectiveness in classifying climate regimes based on deviations in precipitation from long-term averages (McKee *et al*., 1993). Its exclusive reliance on precipitation time-series data contributes to its versatility and reliability for drought assessment across diverse temporal and geographic settings, particularly in vulnerable regions such as Mexico (Mahfouz *et al*., 2016). Building on this established utility, the present study applies SPI to evaluate drought conditions in Zacatecas, north-central Mexico, and provides a foundation for advanced forecasting using neural network architectures. The National Water Commission (CONAGUA) serves as the official body responsible for declaring drought conditions. This is accomplished through the Mexico Drought Monitor, which evaluates various drought parameters, including the SPI, the percentage of normal rainfall anomaly, and the vegetation health index, among others (Esquivel-Saenz *et al*., 2024).

In the context of climate change, the emergence of artificial intelligence (AI) models for drought prediction represents a substantial advance, reflecting their effectiveness and accuracy in drought assessment. In recent years, innovations in machine learning have led to remarkable improvements in operational efficiency, predictive accuracy, and accessibility. Notable examples include the development of attention-based architectures such as the Transformer (Vaswani *et al*., 2017) and its time-series-optimized variants (Su *et al*., 2025), the emergence of automated machine learning (AutoML) frameworks that reduce the need for manual hyperparameter tuning by

automating model selection and optimization processes (He *et al*., 2021), and the availability of open-source libraries such as NeuralForecast (Olivares *et al*., 2022), which lower barriers to the deployment of state-of-the-art neural forecasting models. These advances have enhanced the utility of AI methods for hydrological data analysis. Neural networks are widely recognized as an effective approach for data-driven learning and have demonstrated considerable success in modelling and forecasting nonlinear time series across multiple disciplines, particularly in water resources and hydrology (Ali *et al*., 2017). Models based on artificial neural networks (ANNs), particularly LSTM networks, have emerged as robust data-driven tools for forecasting monthly SPI values (Soh *et al*., 2018). Despite these advances, no studies have explored the integration of SPI and Transformer models in Mexico, particularly in arid and semi-arid regions such as Zacatecas, highlighting the novelty of the present research.

The Transformer model was originally developed to leverage attention mechanisms for the efficient modelling of sequential data, addressing key challenges in sequence learning for natural language processing tasks such as machine translation (Vaswani *et al*., 2017). More recently, Transformer-based architectures have demonstrated exceptional performance across a wide range of domains, including computer vision, speech processing, multimodal learning, reinforcement learning, and time-series forecasting (Su *et al*., 2025). Notably, Transformer variants, including the Informer model, have been successfully applied in hydrometeorological studies for streamflow prediction (Demiray and Demir, 2024) and drought assessment using precipitation indices (Ghobadi *et al*., 2025), highlighting their considerable potential for forecasting SPI time series.

This study applied LSTM and Transformer-based models (Vanilla Transformer and Informer) to develop and implement artificial neural network architectures for forecasting the regional SPI. Based on the above, the following hypothesis was proposed: Transformer-based architectures (Vanilla Transformer and Informer) will outperform LSTM networks in the regional forecasting of the SPI in Zacatecas, Mexico, owing to their superior ability to capture long-range temporal dependencies in hydroclimatic time series.

Accordingly, the objectives of this study were to develop artificial neural network architectures based on LSTM and Transformer methodologies; apply these models to forecast the regional SPI; perform a systematic comparative evaluation of the predictive performance of LSTM and Transformer models across all study regions; and assess the statistical significance of performance differences between the best-performing model and the alternative approaches.

MATERIALS AND METHODS

Study region and datasets

In this study, 31 monthly rainfall time series from weather stations across the state of Zacatecas, Mexico (Figure 1), were analyzed. Monthly precipitation data for the period

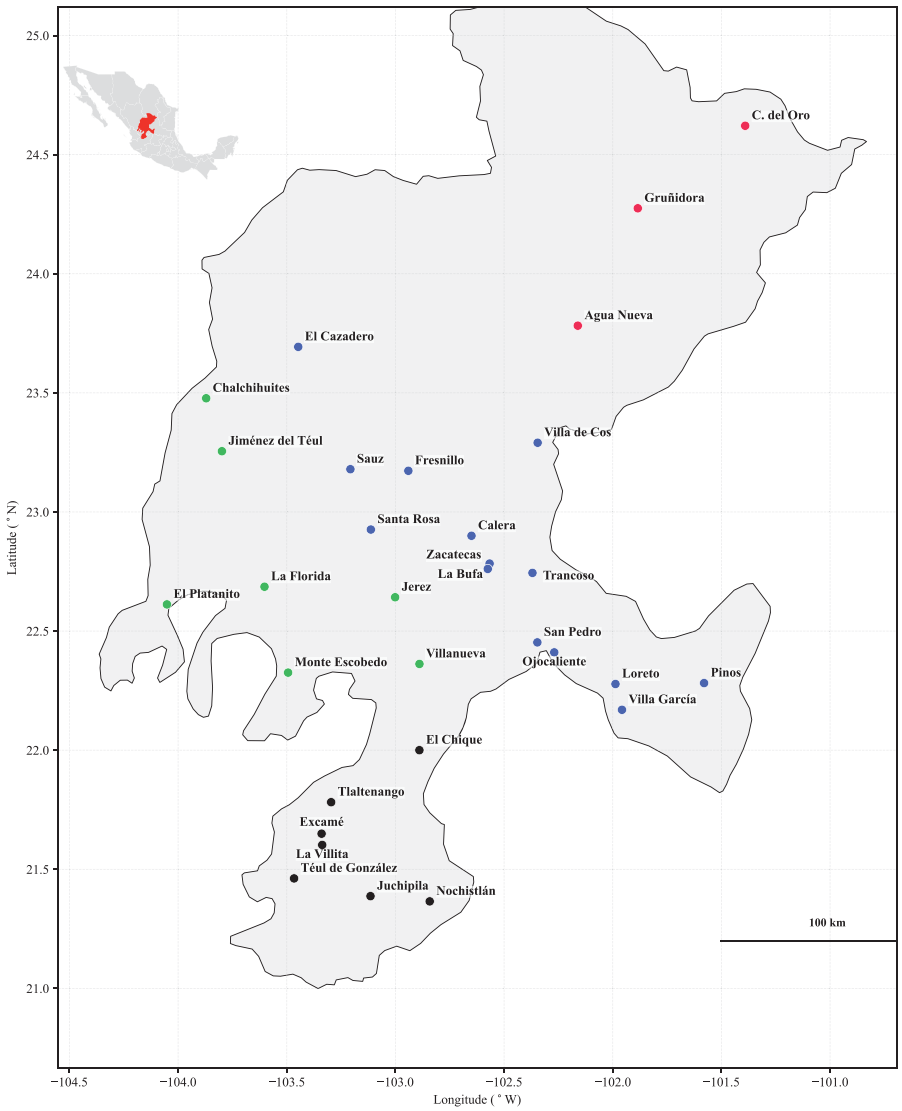

Figure 1. Geographic distribution of meteorological stations in Zacatecas, Mexico, and regional classification derived from cluster analysis. Colors represent Cluster I (Semi-arid, red), Cluster II (Highplain, blue), Cluster III (Mountains, green), and Cluster IV (Canyons, black).

1964–2020 were obtained from the network managed by the National Meteorological Service of Mexico. Prior to analysis, the datasets underwent a comprehensive quality-control process to remove outliers, correct missing or erroneous entries, and verify temporal consistency through homogeneity tests designed to detect shifts in seasonal variability. The Pettitt test was applied to identify abrupt change points in the mean

of each series, while the Standard Normal Homogeneity Test (SNHT) was used to detect step changes potentially associated with station relocations or instrument replacement. All 31 series were retained for the 1964–2020 period. In the few cases where isolated missing values were identified (representing less than 5 % of the records in any series), imputation was performed by linear interpolation using data from neighboring stations, following standard procedures for rainfall data reconstruction in hydrological studies (Navarro-Céspedes *et al*., 2023).

Standardized precipitation index (SPI)

In this study, the SPI was calculated using monthly rainfall data from 31 weather stations across Zacatecas, Mexico, spanning the period 1964–2020. The datasets were rigorously cleaned and validated prior to analysis. This methodology facilitates the assessment of precipitation anomalies by comparing observed values with long-term regional averages at specific locations and time scales. Following the methodology described by Koudahe *et al*. (2017), SPI computation proceeded through the following steps.

Monthly rainfall data were fitted to a gamma distribution, whose probability density function is defined as:

$$g(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{\frac{-x}{\beta}} \text{ for } x > 0$$

where $g(x)$ is the probability density function, α is the shape parameter ($\alpha > 0$), and β is the scale parameter ($\beta > 0$). The gamma function is defined as:

$$\Gamma(\alpha) = \int_0^{\alpha} y^{\alpha-1} e^{-y} dy$$

The parameters α and β are estimated as follows:

$$\alpha = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right)$$

$$\beta = \frac{\overline{x}}{\alpha}$$

$$A = \ln(\overline{x}) - \frac{\sum \ln(\overline{x})}{n}$$

where n is the number of precipitation observations and $\overline{x}$ is the arithmetic mean over the time scale of interest.

The cumulative probability $G(x)$ of an observed amount of rainfall for a given month and time scale is obtained by integrating the probability density function:

$$G(x) = \int_0^x g(x)dx = \frac{1}{\overline{\beta}^{\overline{\alpha}}\Gamma(\overline{\alpha})}\int_0^x x^{\overline{\alpha}} e^{\frac{\overline{x}}{\overline{\beta}}} dx$$

By substituting $t = \overline{x}/\beta$, the equation can be expressed as the incomplete gamma function:

$$G(x) = \frac{1}{\Gamma(\overline{\alpha})}\int_0^x t^{\overline{\alpha}-1} e^{-1}\, dt$$

Nevertheless, the gamma distribution function is undefined for $x = 0$ and $q = P(x = 0) > 0$; where $P(x = 0)$ is the probability of zero precipitation. Hence, the actual probability of non-exceedance $H(x)$ should be calculated as follows:

$$H(x) = q + (1 - q)G(x)$$

where $H(x)$ is the actual probability of non-exceedance and q the probability of $x = 0$. If m is zero in a sample of size n, then q is estimated as:

$$q = \frac{m}{n}$$

Finally, to calculate the SPI, the cumulative probability distribution $H(x)$ is transformed into a standard normal variable Z, with $\mu = 0$ and $\sigma = 1$. The interpretation of wet and drought periods based on SPI values was established by McKee *et al*. (1993). For the purposes of this study, a drought event was operationally defined as a period during which the SPI-12 value remained continuously at or below -1.0 for at least two consecutive months, consistent with the moderate-to-extreme drought categories proposed by McKee *et al*. (1993) (Table 1). Events with SPI-12 values ≤ -2.0 were classified as extreme drought.

SPI calculations incorporate multiple time scales because precipitation variability affects different components of the hydrological cycle (Caloiero, 2017). The 12-month SPI is particularly suitable for evaluating drought impacts on aquifer recharge and groundwater levels. In this study, SPI values were calculated using a 12-month accumulation window with R software version 4.4.1 (R Core Team, 2024) and the Standardized Precipitation-Evapotranspiration Index (SPEI) package version 1.8.1 (Beguería and Vicente-Serrano, 2017).

Table 1. Classification of wet and drought conditions based on the Standardized Precipitation Index (SPI) values.

SPI value	Class
≥2.0	Extremely wet
1.5 to 1.99	Severely wet
1.0 to 1.49	Moderately wet
-0.99 to 0.99	Near normal
-1.49 to -0.99	Moderately dry
-1.99 to -1.49	Severely dry
≤ -2.0	Extremely dry

Cluster analysis

Cluster analysis is a robust statistical technique widely used to identify homogeneous climate zones through the analysis of meteorological data. This method systematically categorizes observations according to their inherent characteristics and mutual relationships. Its primary objective is to group observations such that those within the same cluster exhibit high internal similarity while remaining distinct from those in other clusters. The effectiveness of the clustering process is enhanced by maximizing within-cluster homogeneity and between-cluster heterogeneity (Pampuch *et al*., 2023). Among hierarchical clustering algorithms, Ward's method was selected because of its proven efficiency and its tendency to produce compact, homogeneous clusters by minimizing total within-cluster variance. This characteristic makes it particularly suitable for identifying regional patterns in SPI data. In Ward's method, the distance between two clusters is defined as the increase in the total within-cluster sum of squares resulting from their merger.

In this study, a hierarchical tree-clustering algorithm was implemented to classify monthly SPI time series. The Canberra distance metric was selected as the linkage criterion because of its high sensitivity to relative variations near the origin, making it effective for identifying subtle differences in low-magnitude data (Lance and Williams, 1967). Unlike Euclidean distance, which is primarily influenced by absolute differences and may be biased by high-precipitation events, the Canberra metric ensures that variations during dry periods, which are critical for drought characterization, receive appropriate weight. This property makes it a robust choice for the regionalization of arid and semi-arid climates. Cluster analysis was performed using R version 4.4.1, with the ape package used for hierarchical cluster construction and visualization (Paradis and Schliep, 2018).

Neural time series forecasting

Evidence from previous studies indicates that artificial neural networks (ANNs) may provide a more efficient solution than conventional statistical and econometric methods for modelling nonlinear time-series data (Farajzadeh *et al*., 2014). Among the techniques used for time-series forecasting, LSTM networks and convolutional neural networks have received particular attention (Villegas-Vega *et al*., 2025).

Long Short-Term Memory Networks (LSTM)

LSTM networks were introduced by Hochreiter and Schmidhuber (1997) to address the difficulties encountered by Elman Recurrent Neural Networks in processing temporal patterns within datasets. One of the principal advantages of LSTM networks is their ability to capture long-term temporal dependencies while remaining effective at identifying short-term patterns. The architecture of an LSTM network and its information-processing pipeline (Figure 2) comprise three fundamental components: the input gate, the forget gate, and the output gate. These gates regulate the flow of information into, out of, and within the memory cell. Consequently, LSTM networks are capable of preserving temporal information across a wide range of time steps.

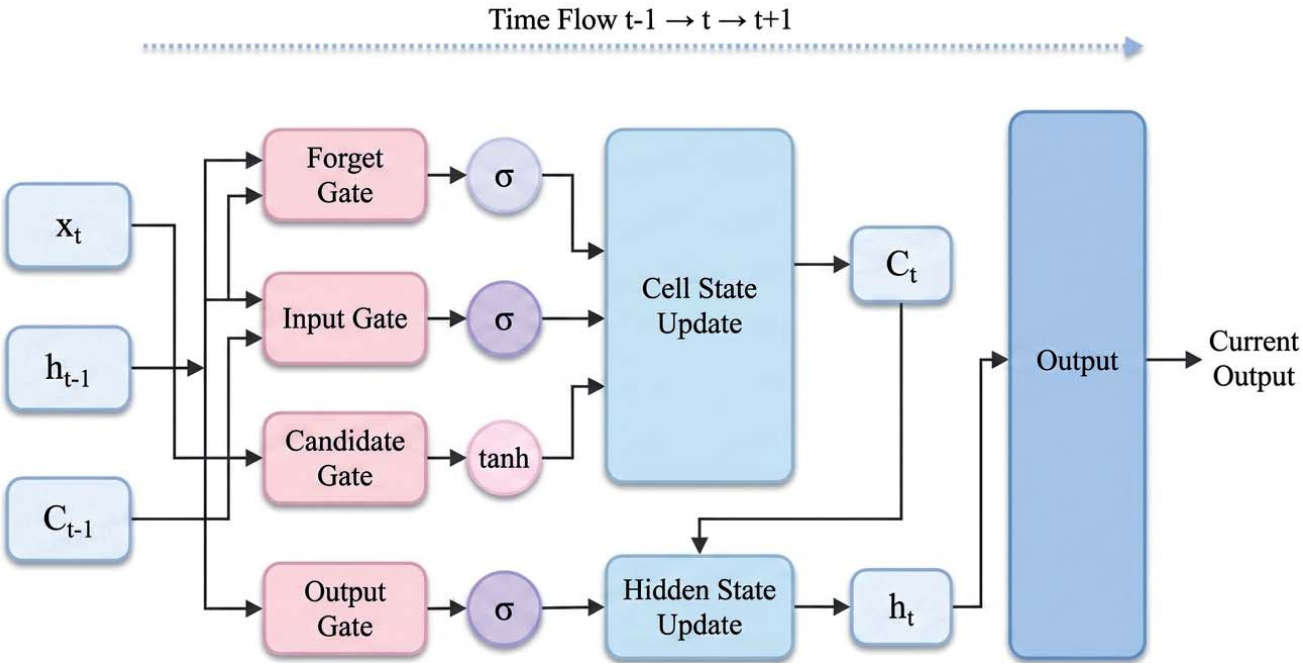

Figure 2. Schematic architecture and information flow of a Long Short-Term Memory (LSTM) unit.

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$\widetilde{C_t} = tanh(W_c[h_{t-1}, x_t] + b_c)$$

$$C_t = \sigma\left(f_t \times C_{t-1} + i_t \times \widetilde{C_t}\right)$$

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o)$$

where x_t is the input at time step t, h_{t-1} is the hidden state from preceding time step, and σ represents a logistic sigmoid function. The weight matrices *Wf*, *Wi*, *WC*, and *Wo*,

together with the bias terms b_f, b_i, b_c, and b_o are model parameters learned during the training phase.

Transformer model

In 2017, Google introduced the Transformer model (Vaswani *et al*., 2017), which uses attention mechanisms to process sequential data efficiently. The model was developed to address challenges associated with sequence-learning tasks in natural language processing, such as machine translation. It enables the transformation of an input sequence from one language domain into an output sequence in another.

The vanilla Transformer (Figure 3) follows the established architecture of neural sequence models and is based on an encoder-decoder framework (Vaswani *et al*., 2017). Both the encoder and decoder are composed of multiple identical blocks. Each encoder block contains a multi-head self-attention mechanism and a position-wise feed-forward network, whereas each decoder block incorporates cross-attention

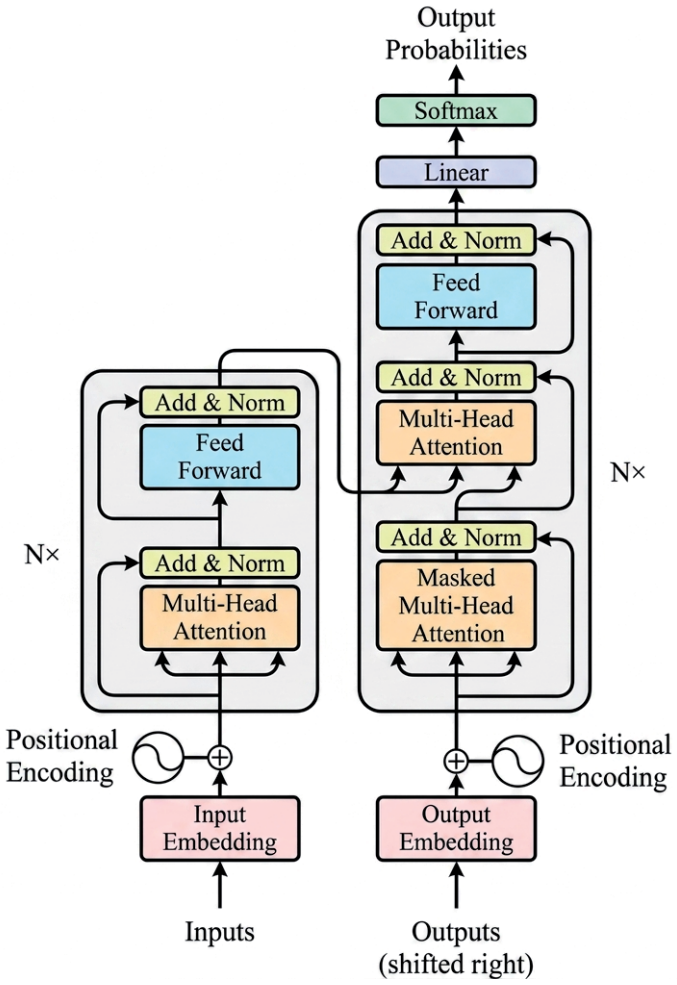

Figure 3. Architecture of the vanilla Transformer model (Vaswani *et al*., 2017).

mechanisms between the multi-head self-attention module and the position-wise feed-forward network.

If a segment of a time series is viewed as a sentence in one language and a subsequent segment as a sentence in another language, the multi-step forecasting problem can be reformulated as a sequence-learning task. Under this framework, the Transformer model is well suited for time-series forecasting. Following Vaswani *et al*. (2017), the core mathematical operations involved in processing time-series data with the Transformer model are described below.

Input representation. Time-series data are commonly represented as scalar- or vector-valued sequences. An input sequence X of length T is defined as:

$$X = [x_1 + x_2 + \cdots + x_T], x_T \in \mathbb{R}^d$$

where x_T is the input vector at time t, and d is the feature dimension.

To create an initial representation, input embeddings are combined with positional encoding to preserve temporal order:

$$Z^{(0)} = E(X) + P$$

where $E(X)$ maps x_T to a higher-dimensional space, and P is the positional encoding matrix.

Multi-head self-attention. For each layer l, the self-attention mechanism computes a weighted representation of the input:

$$Q = Z^{l-1}W_Q, K = Z^{l-1}W_K, V = Z^{l-1}W_V$$

where $W_Q, W_K, W_V \in \mathbb{R}^{d \times d_k}$ are learned projection matrices, and Q, K, and V represent the query, key, and value matrices, respectively.

The attention weights are computed as:

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)$$

and the output of the self-attention mechanism is:

$$O = AV$$

In multi-head attention, several attention heads are computed in parallel and subsequently concatenated and linearly projected:

$$O^{\text{multi-head}} = \text{Concat}(O_1, O_2, \ldots, O_h)W_O$$

where h is the number of heads and $W_O \in \mathbb{R}^{hd_k \times d}$

Feedforward layer. Each layer contains a position-wise feed-forward network that is applied independently at each time step:

$$Z^{(l)} = FFN\left(O^{\text{multi-head}}\right)$$

where

$$FFN(x) = \text{ReLU}(xW_1 + b_1)W_2 + b_2$$

and W_1, W_2, b_1, and b_2 are learned weights and biases.

Layer normalization and residual connections. Both the self-attention and feed-forward modules are surrounded by residual connections and layer normalization:

$$Z^{(l)} = \text{LayerNorm}\left(Z^{(l-1)} + \text{SelfAttention}\left(Z^{(l-1)}\right)\right)$$

$$Z^{(l)} = \text{LayerNorm}\left(Z^{(l)} + FFN\left(Z^{(l)}\right)\right)$$

Output representation. After L Transformer layers, the final output is:

$$Z = Z^{(L)}$$

Time series specific adjustments. For forecasting and regression tasks, a prediction head is used to map the output sequence to the target variable:

$$\hat{y}_t = Z_t W_{out} + b_{out}$$

For causal or autoregressive forecasting, a mask is applied to the attention weights to prevent access to future observations:

$$A_{ij} = \begin{cases} \text{softmax}\left(\dfrac{QK^T}{\sqrt{d_k}}\right) & \text{if } j \leq i \\ 0 & \text{otherwise} \end{cases}$$

Numerous variants of the Transformer architecture have been developed to address specific challenges in time-series modelling (Wen *et al*., 2023). These adaptations have demonstrated effectiveness in a wide range of applications, including classification (Li *et al*., 2021), anomaly detection (Tuli *et al*., 2022), and forecasting (Li *et al*., 2019).

In this study, LSTM (Hochreiter and Schmidhuber, 1997), Vanilla Transformer (Vaswani *et al*., 2017), and Informer (Zhou *et al*., 2021) models were implemented to forecast regional SPI time series derived from precipitation records collected across the state of Zacatecas, Mexico, during the period 1964–2020. Model development and implementation were carried out using the NeuralForecast library (Olivares *et al*., 2022), part of the Nixtlaverse ecosystem, which comprises a suite of open-source libraries designed to facilitate the development of accurate and computationally efficient neural forecasting models (Nixtla, 2025).

Forecasting performance

LSTM, Vanilla Transformer, and Informer models exhibit distinct strengths depending on the forecasting horizon. LSTM performs best in short- to medium-term forecasting and is particularly effective for one-step and recursive multi-step predictions. However, its ability to capture long-range temporal dependencies and perform direct multi-horizon forecasting is more limited. In contrast, the Vanilla Transformer is better suited to medium- and long-term forecasting because it can model global temporal dependencies and supports direct multi-horizon and sequence-to-sequence predictions. Although it often achieves greater accuracy over extended forecasting horizons, it generally requires larger datasets and greater computational resources.

The Informer model, a Transformer-based architecture optimized for efficiency, was specifically designed for long-term forecasting over extended sequences. By utilizing sparse attention mechanisms, Informer achieves faster computation and improved performance in multi-horizon prediction, although it may be less efficient for relatively simple or short-term forecasting tasks. To evaluate the predictive performance of the LSTM, Vanilla Transformer, and Informer models, three standard error metrics were used: Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\left(SPI_{p_i} - SPI_{o_i}\right)\right|$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(SPI_{p_i} - SPI_{o_i}\right)^2$$

$$RMSE = \sqrt{MSE}$$

where SPI_o represents the observed SPI values and SPI_p represents the corresponding predicted values. Lower values of MAE, MSE, and RMSE indicate better predictive performance.

The Mean Absolute Error (MAE) measures the average absolute difference between predicted and observed values, providing an easily interpretable measure of overall model accuracy. Lower MAE values indicate smaller average errors and, consequently, greater predictive precision. The Mean Squared Error (MSE) represents the average squared difference between observed and predicted values, placing greater emphasis on larger errors and serving as a sensitive measure of model fit. Lower MSE values indicate improved forecasting performance. The Root Mean Squared Error (RMSE), defined as the square root of MSE, expresses prediction errors in the original scale of the data. Like MSE, RMSE is sensitive to large deviations, while providing an interpretable measure of the typical magnitude of forecast errors when the model is unbiased (Su *et al*., 2025).

From a statistical perspective, the optimal model is the one that achieves the lowest MSE, RMSE, and MAE values among the candidate models. However, it is necessary to determine whether the superiority of this model is statistically significant rather than a result of random variation. To address this, the Diebold-Mariano (DM) test was applied to evaluate whether differences in forecast accuracy between the best-performing model and the alternative models were statistically significant at the 5 % significance level ($p < 0.05$).

When training multilayer neural networks, the available data are commonly divided into three subsets. The training set is used to compute gradients and update network weights and biases during the learning process. The validation set is used to monitor model performance and optimize hyperparameters throughout training. The test set serves as an independent benchmark for assessing the model's ability to generalize to previously unseen data, thereby providing an unbiased evaluation of predictive performance.

In this study, the LSTM, Vanilla Transformer, and Informer models were trained using an SPI dataset consisting of 518 monthly observations, representing 80 % of the total sample. The remaining 130 observations (20 % of the data) were used for validation. In addition, an independent set of 12 observed SPI values was reserved as the test sample for evaluating predictive accuracy. This procedure was applied to each SPI time series corresponding to the regions identified through the clustering process. The overall data workflow is illustrated (Figure 4).

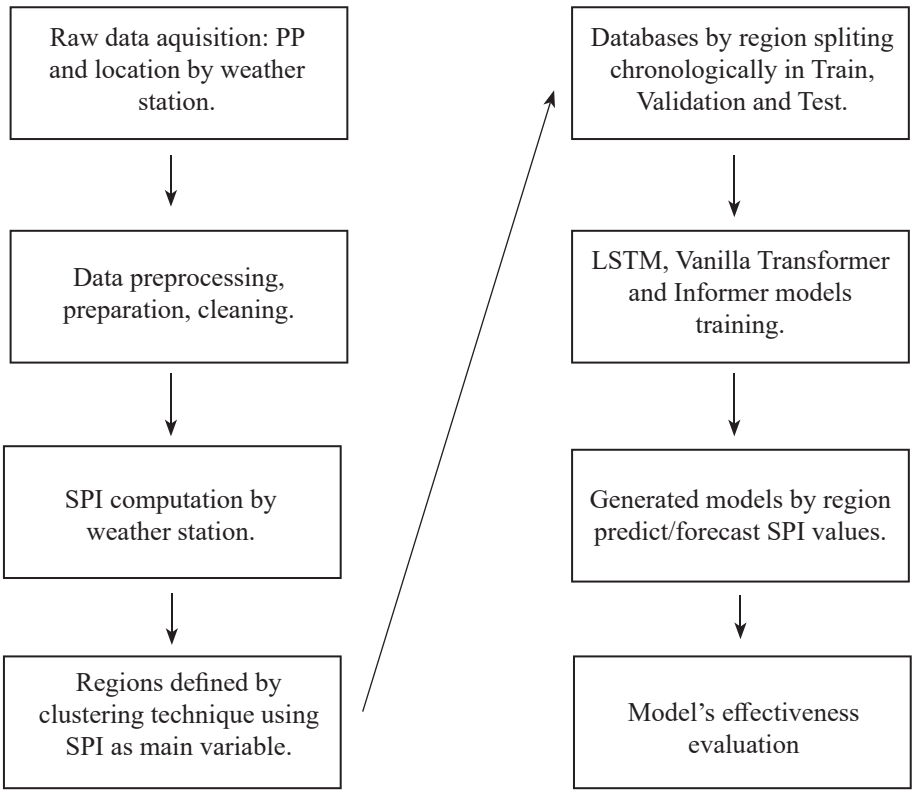

Figure 4. Overview of the analytical workflow from precipitation data processing for Standard Precipitation Index (SPI) forecasting and model evaluation.

RESULTS AND DISCUSSION

The dataset initially consisted of 31 individual rainfall time series, from which the corresponding SPI time series were derived. Cluster analysis was subsequently applied to the complete SPI dataset, resulting in the identification of four distinct regions (Semi-arid, Highplain, Mountains, and Canyons) characterized by similar SPI behavior and precipitation patterns (Figure 5). The regional SPI time series associated with these regions were then used as inputs for predictive modelling with the LSTM, Vanilla Transformer, and Informer architectures. SPI values were obtained from the training, validation, and testing phases, together with the SPI predictions generated by the three models for each region (Figure 6). Model evaluation based on MAE, MSE, and RMSE showed that both Transformer-based models consistently achieved lower error values than the LSTM model across all regions (Table 2).

Examination of the observed SPI-12 time series across the four regions revealed several historically documented drought events in Zacatecas. Using the operational criterion (SPI-12 ≤ -1.0 for at least two consecutive months), notable drought episodes were identified during 1994–1996, 2011–2012, and 2019–2020 in all regions, consistent

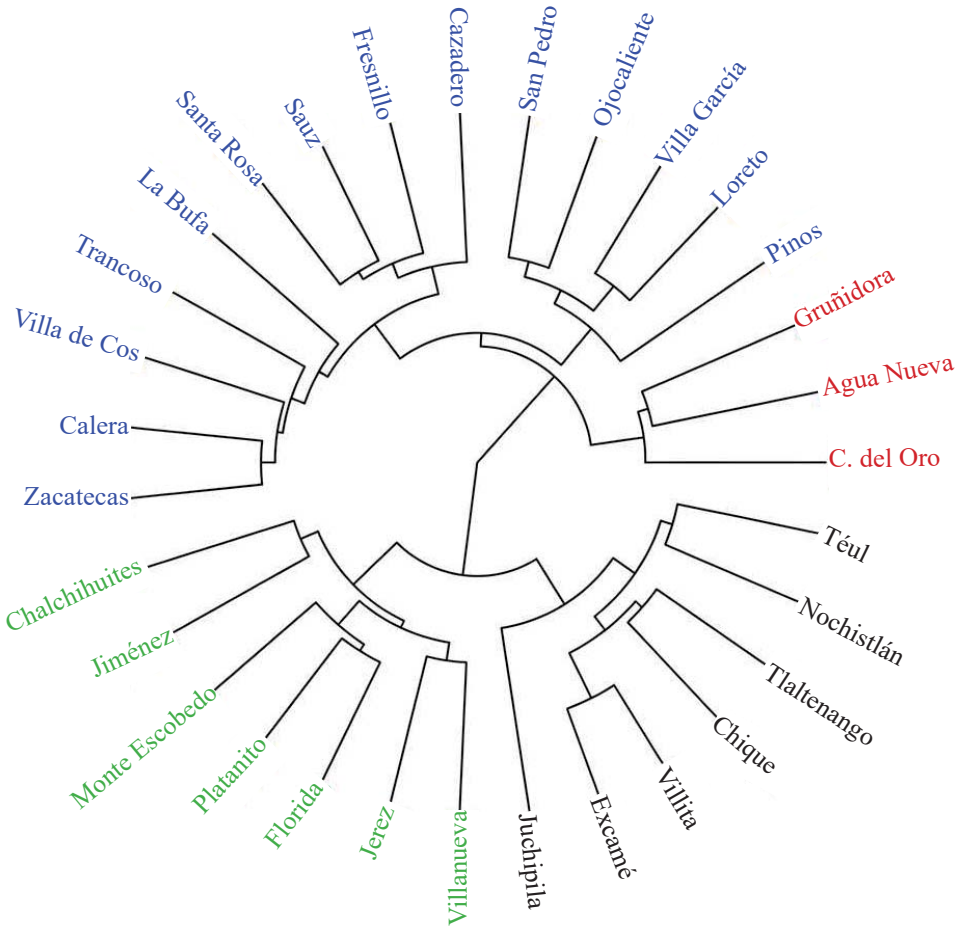

Figure 5. Dendrogram of 31 Standardized Precipitation Index (SPI) time series from Zacatecas, Mexico, generated using hierarchical clustering with Canberra distance. Four regions were identified: Cluster I, Semi-arid (red); Cluster II, Highplains (blue); Cluster III, Mountains (green); Cluster IV, Canyons (black).

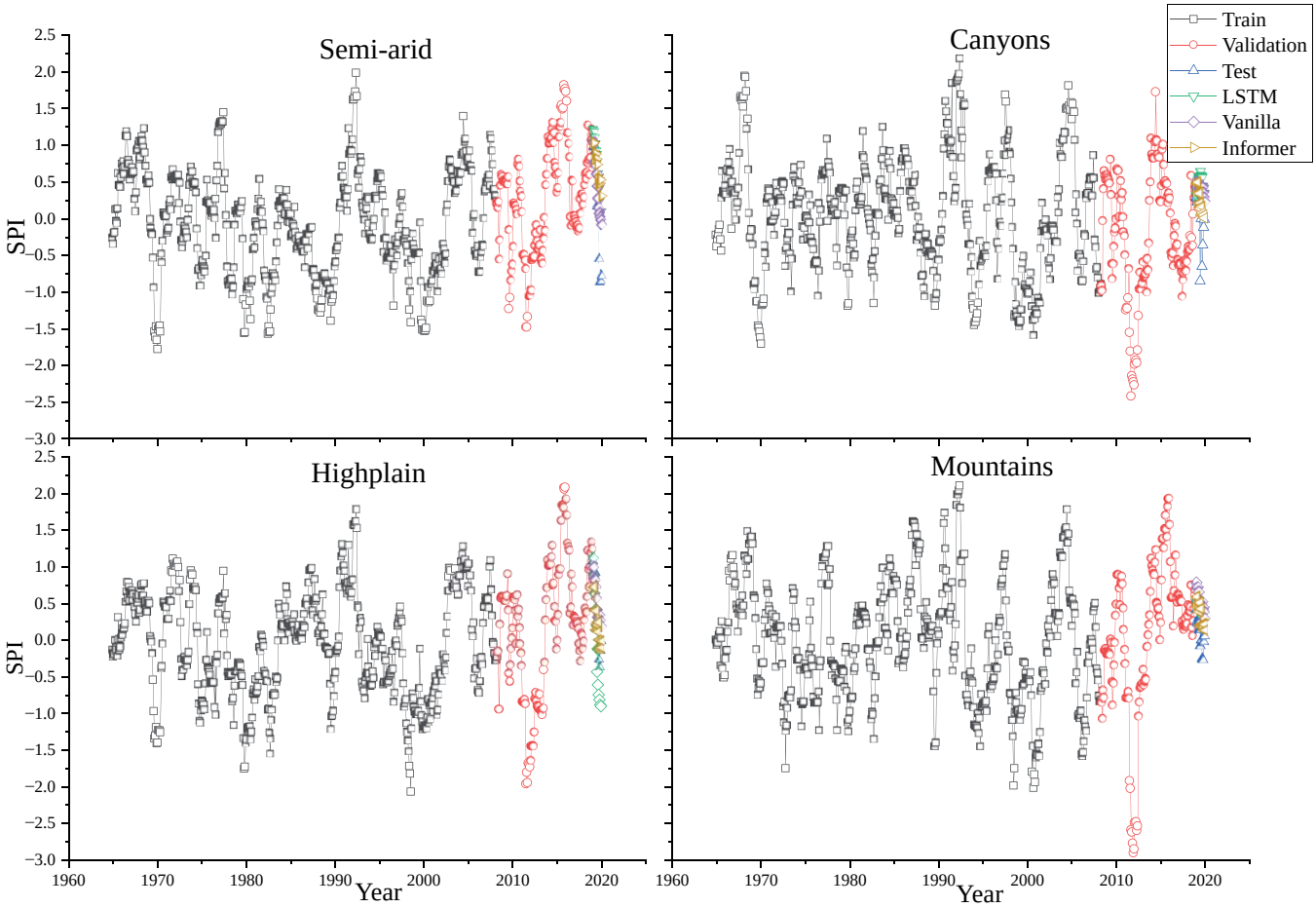

Figure 6. Time series of the regional Standardized Precipitation Index (SPI) for the territory of Zacatecas, encompassing both observed and projected data, in the period from 1964 to 2020.

Table 2. Performance metrics of the neural forecasting models during the validation phase for regional Standardized Precipitation Index (SPI) time series belonging to the state of Zacatecas, Mexico.

	Long Short-Term Memory			Vanilla Transformer			Informer		
Region	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
Semi-arid	0.692	0.832	0.747	0.612	0.783	0.705	0.289	0.538	0.466
Highplain	0.314	0.560	0.474	0.228	0.478	0.405	0.214	0.462	0.403
Mountains	0.694	0.833	0.691	0.332	0.576	0.458	0.252	0.502	0.409
Canyons	0.589	0.767	0.713	0.791	0.890	0.813	0.488	0.698	0.631

with nationally documented drought events reported by Arreguín-Cortés *et al*. (2016). The Semi-arid and Canyons regions exhibited the most prolonged negative SPI-12 episodes, indicating greater sensitivity to precipitation deficits. The extreme drought of 2011–2012 (SPI-12 ≤ −2.0) was particularly severe in the Semi-arid and Highplain regions, consistent with the documented impacts on agricultural production and livestock in north-central Mexico.

The Informer model outperformed both the Vanilla Transformer and LSTM models across all evaluation metrics. For Mean Squared Error (MSE), the Informer achieved an average value of 0.311 across all regions, whereas the LSTM model recorded a higher average MSE of 0.572. A similar pattern was observed for Root Mean Squared Error (RMSE), with the Informer attaining an average value of 0.538 compared with 0.832 for the LSTM model. Mean Absolute Error (MAE) also supported the superior performance of the Informer, which achieved an average value of 0.477, compared with 0.656 for the LSTM model. Among the analyzed regions, the Highplain region exhibited the lowest MSE, RMSE, and MAE values, indicating the strongest predictive performance. In contrast, the Canyons region showed the highest error values across all metrics and, therefore, the lowest predictive accuracy.

The Diebold-Mariano (DM) test results provide a more comprehensive assessment of model performance beyond conventional accuracy metrics. Although the Informer model demonstrated strong overall performance, its superiority was not consistent across all climatic regions of Zacatecas (Table 3). Specifically, the Informer showed statistically significant improvements over the Vanilla Transformer in the Semi-arid region, highlighting its effectiveness under conditions of low rainfall variability. However, in the Canyons and Mountains regions, the LSTM and Vanilla Transformer

Table 3. Diebold-Mariano (DM) test statistics of the neural forecasting models during testing phase for regional Standardized Precipitation Index (SPI) time series for the state of Zacatecas, Mexico.

Region	Informer *vs* Long Short-Term Memory DM statistic	*p*-value
Semi-arid	0.760	0.447
Highplains	-1.554	0.120
Mountains	-2.044	**0.041**
Canyons	-2.822	**0.005**
Region	Informer *vs* Vanilla Transformer DM statistic	*p*-value
Semi-arid	-2.546	**0.011**
Highplains	0.553	0.580
Mountains	-4.911	**0.000**
Canyons	-2.672	**0.008**

models achieved better performance than the Informer. These results suggest that forecasting accuracy is context-dependent and influenced by regional climatic variability.

Forecasting the SPI directly from historical SPI time series, rather than forecasting precipitation and subsequently calculating SPI, provides a streamlined and statistically robust approach that minimizes the error propagation associated with multi-step forecasting. Indirect approaches can accumulate uncertainty because precipitation forecast errors propagate through the gamma-distribution fitting and parameter-estimation stages required for SPI calculation. Previous studies have shown that these compounding errors can alter drought classification by more than 0.2 SPI units in semi-arid environments (Zuo *et al*., 2022). In the present study, direct SPI forecasting using the Informer model produced low error values, with MAE ranging from 0.214 to 0.487, demonstrating both high predictive accuracy and computational efficiency while avoiding the accumulation of biases associated with precipitation-based forecasting approaches.

Furthermore, machine learning models, including neural networks and LSTM architectures, have demonstrated superior performance and generalization when trained on SPI values rather than raw precipitation data. This advantage arises from the normalized and statistically stable properties of SPI, which reduce the noise and variability inherent in precipitation records, thereby facilitating more robust learning and improving forecast consistency (Docheshmeh Gorgij *et al*., 2021).

Direct SPI forecasting is increasingly being incorporated into operational applications, including agricultural advisory services, drought preparedness programs, and governmental policy frameworks, particularly in regions with established early warning systems. Prominent examples include the National Drought Mitigation Center, the U.S. Drought Monitor, and the Global Drought Preparedness Network. Although many operational platforms continue to rely on SPI values derived from observed or forecasted precipitation, there is a growing trend toward the integration of direct SPI forecasts, particularly those generated using machine learning and seasonal climate prediction models. In addition, SPI-based forecasting facilitates integration with large-scale climatic drivers, such as El Niño-Southern Oscillation, and seasonal prediction systems (Hao *et al*., 2018). Collectively, these advantages support the increasing preference for SPI-based drought forecasting over approaches based solely on precipitation modelling.

LSTM networks are widely used for time-series forecasting because of their ability to capture long-term temporal dependencies while preserving short-term patterns (Villegas-Vega *et al*., 2025). Nevertheless, this architecture presents certain limitations. Traditional Multilayer Perceptron (MLP) neural networks, for example, are unable to adequately represent the sequential structure inherent in time-series data, often leading to reduced predictive accuracy (Farajzadeh *et al*., 2014). In contrast, Transformer models use a sequence-to-sequence architecture that provides the flexibility required to address complex sequence-learning tasks. By incorporating attention mechanisms,

Transformer-based approaches can efficiently identify and retain long-range dependencies within sequences (Lezmi and Xu, 2023), including those present in SPI time series.

The results obtained indicate that both Transformer-based models, particularly the Informer, achieved high predictive accuracy in forecasting the SPI across the regions of Zacatecas. The Informer model outperformed the Vanilla Transformer, likely because its sparse attention mechanism more effectively captures long-range dependencies in extended SPI time series (1964–2020), a characteristic associated with the highly variable precipitation patterns of the region. In addition, its computational efficiency and ability to focus on the most relevant temporal features make it particularly suitable for representing the complex climatic dynamics reflected in SPI variability.

Overall, the Informer model demonstrated strong predictive accuracy for SPI forecasting based on the MSE, RMSE, and MAE evaluation metrics. Specifically, the Informer reduced the average MSE by approximately 15 % relative to the LSTM model, highlighting the magnitude of its performance improvement. This result reflects the model's ability to capture long-term dependencies and temporal variability within SPI data while maintaining computational efficiency.

The Informer model exhibited statistically significant advantages over the Vanilla Transformer in the Semi-arid region, as confirmed by significance tests evaluating whether performance differences could be attributed to random variation. In contrast, the results indicated that the LSTM and Vanilla Transformer models outperformed the Informer in the Canyons and Mountains regions. These findings suggest that regional climatic heterogeneity influences model performance and that a regionalized model-selection strategy may improve drought forecasting accuracy. Specifically, the Informer may be more suitable for Semi-arid regions, whereas LSTM or Vanilla Transformer architectures may be preferable in more topographically complex areas. Such an approach could improve forecasting performance while strengthening drought preparedness and water-resource management efforts across Zacatecas.

These findings are consistent with previous studies that successfully applied artificial neural networks to forecast the monthly SPI (Magallanes-Quintanar *et al*., 2022, 2024; Villegas-Vega *et al*., 2025). In addition, the present study extends the work of Giddings *et al*. (2005) by introducing a methodology capable of identifying smaller and more precise climatic regions in Mexico using SPI. This research contributes to the international literature by adapting the Informer model to capture complex, region-specific precipitation dynamics and by providing a scalable and transferable framework for drought forecasting in other regions of the world.

Additionally, the Informer model implemented through the NeuralForecast framework (Olivares *et al*., 2022) represents a valuable and contemporary tool for time-series forecasting, including drought-related applications. Its combination of predictive accuracy, computational efficiency, and scalability makes it a promising approach for assessing climatic and agricultural risks associated with drought, particularly in arid and semi-arid regions that are increasingly affected by climate change-driven extremes.

CONCLUSIONS

Drought prediction has become increasingly important in meteorology, hydrology, water resources management, and sustainable agriculture because of the growing dependence of human activities on reliable water supplies. To address this challenge, artificial intelligence models based on the Long Short-Term Memory (LSTM), Vanilla Transformer, and Informer architectures were developed to forecast four monthly Standardized Precipitation Index (SPI) time series representing regional conditions across Zacatecas, Mexico. According to the evaluation metrics, the Informer model achieved the highest predictive performance across all study regions, reducing the average MSE by approximately 15 % relative to the LSTM model and showing comparable improvements over the Vanilla Transformer. These results were supported by the Diebold-Mariano test. However, additional significance tests suggest that a regionalized model-selection strategy, using the Informer in semi-arid regions and LSTM or Transformer models in more complex terrains, may further improve forecasting accuracy and reliability.

In the context of climate change, accurate SPI forecasting with the Informer model has considerable potential to enhance water-resource management in Zacatecas. As precipitation patterns become increasingly variable, reliable SPI forecasts can support proactive drought mitigation, improve agricultural planning, and contribute to sustainable water management. The model's ability to capture long-term dependencies within the 1964–2020 dataset highlights its applicability for decision-making in semi-arid and mountainous regions affected by persistent climatic variability.

Future research should evaluate the incorporation of additional climatological variables, such as temperature, evapotranspiration, and large-scale climate indices (e.g., El Niño-Southern Oscillation), to further improve SPI forecasting performance. These studies should also be extended to other climatic settings, including extreme drought periods beyond 2020 and contrasting environments such as tropical and coastal regions, to assess the robustness, adaptability, and generalizability of the proposed approach across diverse climate regimes.

REFERENCES

Ali Z, Hussain I, Faisal M, Nazir HM, Hussain T, Shad MY, Mohamd Shoukry A, Hussain Gani S, 2017. Forecasting drought using multilayer perceptron artificial neural network model. Advances in Meteorology 2017. https://doi.org/10.1155/2017/5681308

Amanambu AC, Mossa J, Chen YH, 2022. Hydrological drought forecasting using a deep transformer model. Water 14 (22): 3611. https://doi.org/10.3390/w14223611

Arreguín-Cortés FI, López-Pérez M, Ortega-Gaucin D, Ibañez-Hernández Ó. 2016. La política pública contra la sequía en México: avances, necesidades y perspectivas. Tecnología y Ciencias del Agua 7 (5): 63–76.

Beguería S, Vicente-Serrano SM, 2017. SPEI Calculator. Digital.CSIC. http://doi.org/10.20350/digitalcsic/8997

Caloiero T. 2017. Drought analysis in New Zealand using the standardized precipitation index. Environmental Earth Sciences 76 (16): 1–13. https://doi.org/10.1007/s12665-017-6909-x

Demiray BZ, Demir I, 2024. Towards generalized hydrological forecasting using transformer models for 120-hour streamflow prediction. arXiv. https://doi.org/10.48550/arxiv.2406.07484

Docheshmeh Gorgij A, Alizamir M, Kisi O, Elshafie A. 2021. Drought modelling by standard precipitation index (SPI) in a semi-arid climate using deep learning method: Long short-term memory. Neural Computing and Applications 34 (3): 2425–2442. https://doi.org/10.1007/s00521-021-06505-6

Esquivel-Saenz PJ, Ortiz-Gómez R, Zavala M, Flowers-Cano RS. 2024. Artificial neural networks for drought forecasting in the central region of the state of Zacatecas, Mexico. Climate 12 (9): 131. https://doi.org/10.3390/cli12090131

Farajzadeh J, Fakheri Fard A, Lotfi S. 2014. Modeling of monthly rainfall and runoff of Urmia Lake basin using "feed-forward neural network" and "time series analysis" model. Water Resources and Industry 7–8: 38–48. https://doi.org/10.1016/j.wri.2014.10.003

Ferreira RN, Nissenbaum MR, Rickenbach TM. 2018. Climate change effects on summertime precipitation organization in the Southeast United States. Atmospheric Research 214: 348–363. https://doi.org/10.1016/j.atmosres.2018.08.012

Florescano E. 2000. Breve historia de la sequía en México (Segunda edición). Consejo Nacional para la Cultura y las Artes: Ciudad de México, México. 252 p.

García-Acosta V, Pérez-Zevallos JM, Molina-del Villar A. 2003. Desastres agrícolas en México: catálogo histórico. Centro de Investigaciones y Estudios Superiores en Antropología Social: Ciudad de México, México. 506 p.

Ghobadi F, Tayerani Charmchi AS, Kang D. 2025. Enhancing long-term flood forecasting with SageFormer: A cascaded dimensionality reduction approach based on satellite-derived data. Remote Sensing 17 (3): 365. https://doi.org/10.3390/rs17030365

Giddings L, Soto M, Rutherford BM, Maarouf A. 2005. Standardized precipitation index zones for Mexico. Atmósfera 18 (1): 33–56.

Hao Z, Singh VP, Xia Y. 2018. Seasonal drought prediction: Advances, challenges, and future prospects. Reviews of Geophysics 56 (1): 108–141. https://doi.org/10.1002/2016rg000549

He X, Zhao K, Chu X. 2021. AutoML: A survey of the state-of-the-art. Knowledge-Based Systems 212: 106622. https://doi.org/10.1016/j.knosys.2020.106622

Hochreiter S, Schmidhuber J. 1997. Long short-term memory. Neural Computation 9 (8): 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Karmalkar AV, Bradley RS, Diaz HF. 2011. Climate change in Central America and Mexico: Regional climate model validation and climate change projections. Climate Dynamics 37 (3–4): 605–629. https://doi.org/10.1007/s00382-011-1099-9

Kaushik S, Choudhury A, Sheron PK, Dasgupta N, Natarajan S, Pickett LA, Dutt V. 2020. AI in Healthcare: Time-series forecasting using statistical, neural, and ensemble architectures. Frontiers in Big Data 3. https://doi.org/10.3389/fdata.2020.00004

Koudahe K, Kayode AJ, Samson AO, Adebola AA, Djaman K. 2017. Trend analysis in standardized precipitation index and standardized anomaly index in the context of climate change in southern Togo. Atmospheric and Climate Sciences 7 (4): 401–423. https://doi.org/10.4236/acs.2017.74030

Lance GN, Williams WT. 1967. A general theory of classificatory sorting strategies: 1. Hierarchical systems. The Computer Journal 9 (4): 373–380. https://doi.org/10.1093/comjnl/9.4.373

Lezmi E, Xu J. 2023. Time series forecasting with transformer models and application to asset management. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4375798

Li C, Yang J, Zhang P, Gao M, Xiao B, Dai X, Yuan L, Gao J. 2021. Efficient self-supervised vision transformers for representation learning. arXiv. https://doi.org/10.48550/arxiv.2106.09785

Li S, Jin X, Xuan Y, Zhou X, Chen W, Wang YX, Yan X. 2019. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. arXiv. https://doi.org/10.48550/arxiv.1907.00235

Liu H, Liu Y, Guo X, Wu H, Wang H, Liu Y. 2023. An energy consumption prediction method for HVAC systems using energy storage based on time series shifting and deep learning. Energy and Buildings 298: 113508. https://doi.org/10.1016/j.enbuild.2023.113508

Magallanes-Quintanar R, Galván-Tejada CE, Galván-Tejada JI, Gamboa-Rosales H, Méndez-Gallegos SDJ, García-Domínguez A. 2024. Auto-machine-learning models for standardized precipitation index prediction in north-central Mexico. Climate 12 (7): 102. https://doi.org/10.3390/cli12070102

Magallanes-Quintanar R, Galván-Tejada CE, Galván-Tejada JI, Méndez-Gallegos SJ, García-Domínguez A, Gamboa-Rosales H. 2022. Narx neural networks models for prediction of standardized precipitation index in Central Mexico. Atmosphere 13 (8): 1254. https://doi.org/10.3390/atmos13081254

Mahfouz P, Mitri G, Jazi M, Karam F. 2016. Investigating the temporal variability of the standardized precipitation index in Lebanon. Climate 4 (2): 27. https://doi.org/10.3390/cli4020027

McKee TB, Doesken NJ, Kleist J. 1993. The relationship of drought frequency and duration to time scales. *In* Proceedings of the 8th Conference on Applied Climatology. Anaheim, CA, USA, pp: 179–183.

Navarro-Céspedes JM, Hernández JH, Alcántara-Concepción PC, Morales-Martínez JL, Carreño-Aguilera G, Padilla-Benítez F. 2023. A comparison of missing values imputation methods applied to precipitation of two semi-arid and humid regions of México. Atmósfera 37. https://doi.org/10.20937/atm.53095

Nixtla. 2025. Nixtlaverse. https://nixtlaverse.nixtla.io/ (Retrieved: May 2026).

Olivares KG, Challú C, Garza F, Canseco MM, Dubrawski A. 2022. NeuralForecast: User friendly state-of-the-art neural forecasting models. PyCon: Salt Lake City, UT, USA.

Pampuch LA, Negri RG, Loikith PC, Bortolozo CA. 2023. A review on clustering methods for climatology analysis and its application over South America. International Journal of Geosciences 14 (9): 877–894. https://doi.org/10.4236/ijg.2023.149047

Paradis E, Schliep K. 2018. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35 (3): 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pathania A, Gupta V. 2025. Interpretable transformer model for national scale drought forecasting: Attention-driven insights across India. Environmental Modelling and Software 187: 106394. https://doi.org/10.1016/j.envsoft.2025.106394

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Siami-Namini S, Namin AS. 2018. Forecasting economics and financial time series: ARIMA vs. LSTM. arXiv. https://doi.org/10.48550/arxiv.1803.06386

Soh YW, Koo CH, Huang YF, Fung KF. 2018. Application of artificial intelligence models for the prediction of standardized precipitation evapotranspiration index (SPEI) at Langat

River Basin, Malaysia. Computers and Electronics in Agriculture 144: 164–173. https://doi.org/10.1016/j.compag.2017.12.002

Su L, Zuo X, Li R, Wang X, Zhao H, Huang B. 2025. A systematic review for transformer-based long-term series forecasting. Artificial Intelligence Review 58 (3): 80. https://doi.org/10.1007/s10462-024-11044-2

Tuli S, Casale G, Jennings NR. 2022. TranAD: Deep transformer networks for anomaly detection in multivariate time series data. arXiv. https://doi.org/10.48550/ARXIV.2201.07284

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. 2017. Attention is all you need. Advances in Neural Information Processing Systems 30: 5998–6008.

Villegas-Vega R, Márquez-Grajales A, Mezura-Montes E, Salas-Martínez F, Ojeda-Misses MA, Romo-Gómez C. 2025. Optimization of LSTM networks through neuroevolution for drought forecasting in Mexico. Theoretical and Applied Climatology 156 (11): 562. https://doi.org/10.1007/s00704-025-05818-z

Wen Q, Zhou T, Zhang C, Chen W, Ma Z, Yan J, Sun L. 2023. Transformers in time series: A survey. *In* Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization: Macau, China, pp: 6778–6786. https://doi.org/10.24963/ijcai.2023/759

Zhou H, Zhang S, Peng J, Zhang S, Li J, Xiong H, Zhang W. 2020. Informer: Beyond efficient transformer for long sequence time-series forecasting. arXiv. https://doi.org/10.48550/ARXIV.2012.07436

Zuo DD, Hou W, Zhang Q, Yan PC. 2022. Sensitivity analysis of standardized precipitation index to climate state selection in China. Advances in Climate Change Research 13 (1): 42–50. https://doi.org/10.1016/j.accre.2021.11.004

Agrociencia

BIOMETRIC EVALUATION OF FRESHWATER TURTLES IN LOS RÍOS REGION IN SOUTHEASTERN MEXICO

Alfonso de Jesús **Sánchez-López**[1], Claudia Elena **Zenteno-Ruíz**[2], Raúl Enrique **Hernández-Gómez**[1*], Alfonso **Castillo-Domínguez**[1]

[1]Universidad Juárez Autónoma de Tabasco. División Académica Multidisciplinaria de Los Ríos. Carretera Tenosique-Estapilla km 1, Cocoyol, Tenosique de Pino Suárez, Tabasco, México. C. P. 86902.

[2]Universidad Juárez Autónoma de Tabasco. División Académica de Ciencias Biológicas. Carretera Villahermosa-Cárdenas km 0.5, entronque a Bosques de Saloya, Centro, Tabasco, México. C. P. 86039.

* Author for correspondence: raul.hernandez@ujat.mx

Citation: Sánchez-López AJ, Zenteno-Ruíz CE, Hernández-Gómez RE, Castillo-Domínguez A. 2026. Biometric evaluation of freshwater turtles in Los Ríos region in southeastern Mexico. **Agrociencia** 60(4): 609-621. https://doi.org/ 10.47163/agrociencia.v60i4.3317

Editor in Chief:
Dr. Fernando C. Gómez-Merino

Received: December 16, 2025.
Approved: May 26, 2026.
Published in Agrociencia: June 22, 2026.

ABSTRACT

In the state of Tabasco, there are nine species of freshwater turtles, all of which have a conservation status under NOM-059-SEMARNAT-2010. An alternative for the recovery of these populations is management for their conservation. This study characterized freshwater turtles in captivity at División Académica Multidisciplinaria de los Ríos (DAMR), in the municipality of Tenosique de Pino Suárez, Tabasco, Mexico, as part of the evaluation of their potential for reproduction and management within a Management Unit for Wildlife Conservation (UMA). The species under study were the Mesoamerican slider (*Trachemys venusta*), the furrowed wood turtle (*Rhinoclemmys areolata*), and the Tabasco mud turtle (*Kinosternon leucostomum*). Body measurements were taken, and physical characteristics were evaluated using a pre-established form. The turtles were marked using an adaptation of the Cagle (1939) method. A total of 113 individuals were analyzed: 76 furrowed wood turtles, 32 sliders, and five Tabasco mud turtles, with densities of 0.751, 0.316, and 0.049 individuals per m^2, respectively. The sliders presented a shell length-weight (SL-P) ratio of $P = 0.1935SL^{2.8195}$ and a female-to-male sex ratio of 1:1.8. The Mesoamerican sliders displayed a SL-P ratio of $P = 0.1325SL^{2.9222}$ and a sex ratio of 1.2:1. The Tabasco mud turtle presented a ratio of $P = 0.5159SL^{2.3777}$ and a sex ratio of 4:1. The prevalence of physical conditions was evaluated for each species and for the total number of turtles, where 42 % (n = 47) presented a good condition and 58 % (n = 66) displayed irregular physical characteristics. In the latter, the highest percentage of findings corresponded to erosion of the shell and plastron, while the least observed were malformation of the mouth and loss of eyes. Data analysis suggests that conditions are suitable for species conservation and for completing the necessary procedures to establish a UMA in the future.

Keywords: body biometrics, tagging, population density, UMA potential, captivity, *Trachemys venusta, Rhinoclemmys areolata, Kinosternon leucostomum*.

INTRODUCTION

Mexico, with its large abundance of species, marine and terrestrial ecosystems, and genetic resources, and harboring nearly 70 % of the variety of plants and animals on the planet, is ranked fifth among the 12 megadiverse countries in the world (SEMARNAT, 2018). However, a point of no return has been reached in the loss of terrestrial and marine biodiversity due to the absence of places in the Earth's biosphere without human footprints (Venter *et al*., 2016). In the search for solutions to minimize these effects, two correlated activities stand out: economic profitability based on the sustainable use of wildlife and the long-term conservation of species (Hernández-Silva *et al*., 2018).

By contrast, one of the most significant and growing problems in our country is the illegal trafficking of wildlife, which produces an incalculable environmental imbalance due to the lack of strategies to minimize this crime (Van Uhm, 2016; Arroyo-Quiroz and Wyatt, 2019). Wild animals are captured to be used as raw materials in the fashion industry, to be sold as pets, for consumption, and for alleged medicinal properties, beliefs, and aphrodisiacs (Quevans *et al*., 2013). Freshwater turtles in southeastern Mexico are intensely exploited as a food resource, as they are part of local culture and tradition, although their populations have been drastically reduced due to additional factors such as environmental contamination, deforestation, urbanization, and cattle breeding (Tellería, 2013).

Turtles are reptiles that belong to the order Testudines, anatomically unmistakable due to the presence of a dorsal shell and a bony ventral plastron, which provide protection against predators and climatic pressures. Continental turtles in Mexico are among the most diverse in the world (Cázares-Hernández, 2015). In Tabasco, they are part of the biotic heritage and have coexisted with local populations since ancient times. These communities value and appreciate them both for their nutritional importance and as part of their cultural heritage (Beauregard-Solís *et al*., 2010).

Tabasco has nine species of freshwater turtles, including the Central American river turtle (*Dermatemys mawii*), the Mesoamerican slider (*Trachemys venusta*), the giant musk turtle (*Staurotypus triporcatus*), the Central American snapping turtle (*Chelydra rossignoni*), the narrow-bridged musk turtle (*Claudius angustatus*), the furrowed wood turtle (*Rhinoclemmys areolata*), the white-lipped mud turtle (*Kinosternon leucostomum*), the Tabasco mud turtle (*Kinosternon acutum*), and the scorpion mud turtle (*Kinosternon scorpioide cruentatum*) (Beauregard-Solís *et al*., 2010). All these species have some conservation status under NOM-059-SEMARNAT-2010 (DOF, 2010).

Studies have been carried out in Tabasco on freshwater turtles, highlighting the evaluation of the physical condition of hickatees (*Dermatemys mawii*) and Mesoamerican sliders (*Trachemys venusta*) (Rangel-Mendoza and Weber, 2015). These investigations provide the basis for the implementation of wildlife conservation management in Mexico. In this context, four goals can be pursued: (1) maintain populations in a healthy state through monitoring, (2) increase populations through appropriate zootechnical practices, (3) enable sustainable use, and (4) control populations when they become harmful (Hernández-Silva *et al*., 2018).

The physical condition of individuals is an integrative indicator of health; its evaluation is non-invasive and enables systematic monitoring of prevalence (the proportion of the population presenting an ailment or condition), which is useful for management decision-making. However, further studies are still needed on the physical characteristics, habitat conditions, conservation, and management of other chelonian species.

Therefore, the main goal of this investigation was to identify, classify, and determine the prevalence of physical ailments in freshwater turtles kept in captivity in Los Ríos region of the state of Tabasco, Mexico, to generate baseline information for a management plan. The biometric assessment and evaluation of the physical condition of the initial populations were considered crucial for developing the management plan based on species requirements, particularly for infrastructure design and animal welfare considerations.

MATERIALS AND METHODS

Study site

This study was carried out between November 2021 and April 2022 at División Académica Multidisciplinaria de los Ríos (DAMR), in the municipality of Tenosique de Pino Suárez, Tabasco, Mexico (17° 29′ 15.63″ N, 91° 25′ 33.85″ W), at an altitude of 25 m (Figure 1). This area is located near the Aquaculture Production Unit (UPA), under RNPA number 27092543, property of DAMR.

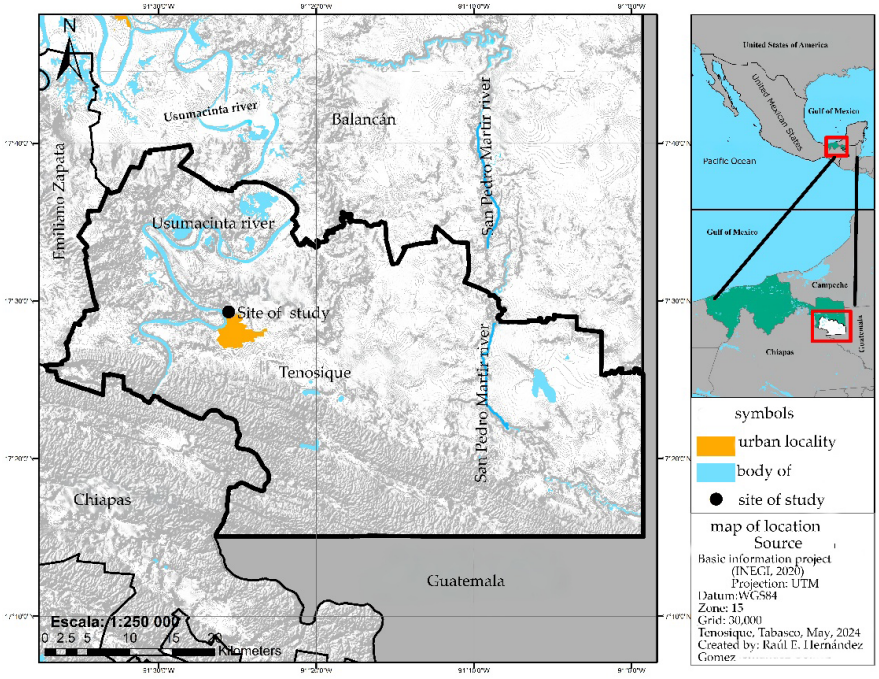

Figure 1. Geographic location of the study site in the municipality of Tenosique, Tabasco, Mexico.

Identification of species and capture

Species identification was carried out following the taxonomic classification of Casas-Andreu (1965) and the dichotomous keys described by Yvan *et al*. (1999) in the turtle identification manual of the Convention on International Trade in Endangered Species of Wild Flora and Fauna (CITES), as well as the updated identification guides for southeastern Mexico (Díaz-Gamboa *et al*., 2022). The physical characteristics of each species were determined based on color tones and the morphology of the shell, head, and size.

Turtles were captured manually in the area by picking them up from the underside of the shell, holding them with four fingers on the plastron, and the thumbs securing the carapace (Gallina and López-González, 2011).

Tagging and taking biometric data

Technical data sheets were prepared using predetermined forms that include marking criteria and records of physical characteristics similar to those used by the Management Units for Wildlife Conservation (UMAS). In the state of Tabasco, notable UMAs include the Conservation of Endangered Species (CICEA), within the Academic Division of Biological Sciences; the Tabasco State Government turtle farm; the Arca de Noé turtle farm; and the Arroyo Tabasquillo turtle farm. These data sheets are a fundamental part of the administrative process for the Ministry of Environment, Natural Resources and Fisheries (SEMARNAP) to develop a management plan aimed at the future establishment of a wildlife conservation and sustainable use management unit within the DAMR.

The turtles were marked using an adaptation of the Cagle (1939) method, which consists of marking each specimen with structured notches on the shell, enabling permanent identification of the individual. The notching was performed on the marginal scutes of each turtle's shell using compact frame saws.

To obtain biometric data, a digital CGOLDENWALL scale with an accuracy of 0.1 g to 10 kg was used, along with a 100 cm flexible measuring tape; a binary table 45 cm in width, 1.1 m in length, and 70 cm in height for handling; and plastic containers 60 cm in length, 40 cm in width, and 32 cm in height to place the turtles. A clipboard was also used to record data in the biometric forms. Minimum and maximum values, as well as averages and standard deviations, were calculated. The recorded variables included species, weight, sex, length, curved shell length (CSL), and curved shell width (CSW). Physical characteristics of the shell, plastron, head, eyes, limbs, and tail, such as hardness, malformation, curvature, presence of algae, wounds, mutilations, and edemas, as well as the behavior, were determined.

Proportion of sex by species

To assess the sex ratio of males to females within each species, the turtle tail and plastron characterization method was utilized (Eckert *et al*., 2000), following the next formula:

$$P = a / b$$

where a is the number of females and b is the number of males (Chavarría-Pérez *et al.*, 2020)

Evaluation of the optimal and abnormal physical conditions

The prevalence of physical conditions, defined as the ratio of affected individuals to the total population and expressed as a percentage, was calculated for nine external physical conditions. These conditions include erosion of the shell and plastron; depressions and curvature; grooves, cracks, and notches; wounds on the shell, plastron, tail, and head; malformation and wounds of the nose; edema in the limbs; malformation of the mouth; and loss of eyes, all of which are indicators of poor health status (Rangel-Mendoza and Weber, 2015). For this study, the total turtle population was analyzed, and totals were subsequently calculated by species.

Population density

Population density is an indicator that helps measure the level of satisfaction of organisms with the space available for interaction with the surroundings and with other individuals. This factor reflects their condition of well-being, indicating whether the population is clustered or dispersed. It is expressed in individuals per square meter and is calculated using the following formula:

$$D = N / S$$

where D is density, N is population abundance, and S is the surface area it occupies (expressed in m^2) (Mandujano-Rodríguez, 2011).

Length-weight ratio

The ratio between the curved shell length and the weight of the turtle was calculated using the formula proposed by Ricker (1975):

$$TW = aSL^b$$

where TW is total weight (in grams), a is a regression constant equivalent to the condition factor, SL is shell length (in centimeters), and b is the growth coefficient of the regression.

RESULTS AND DISCUSSION

Identification of species, capturing and tagging

Three species of tropical freshwater turtles were identified (*Rhinoclemmys areolata*, *Trachemys venusta*, and *Kinosternon leucostomum*), which are among the nine species

recorded for the state of Tabasco (Beauregard-Solís *et al*., 2010) and are endemic to the Neotropics (Ippi and Flores, 2001). Within the facilities of DAMR, a total of 113 individuals were recorded (Table 1).

Table 1. Composition and relative abundance of freshwater turtle species studied in Los Ríos region of Tabasco, Mexico.

Family	Scientific name	Common name	n	%
Emididae	*Trachemys venusta*	Central American slider	32	28
Geomydidae	*Rhinoclemmys areolata*	Furrowed wood turtle	76	67
Kinosternidae	*Kinosternon leucostomum*	Tabasco mud turtle	5	5

Analysis of biometric data of turtles

The *R. areolata* population presented a curved shell length (CSL) of 14 to 21 cm (18.65 ± 1.67 cm). In this species, the CSL with the highest frequency was 20 cm, and the lowest was 14 cm. The weight range for this species varied from 377 to 1248 g (761.17 ± 207.02 g). These data are similar to those obtained in another species of the same family (*R. pulcherrima*), with measurements in females of 17.1–23.6 cm and weights of 740–1870 g, while males presented lengths of 12.13–18.9 cm and weights of 240–690 g in captivity on the coast of Oaxaca (Rodríguez-Murcia *et al*., 2014).

Trachemys venusta presented CSL values ranging between 18 and 33 cm (23.12 ± 3.77 cm). The most frequent CSL was 23 cm, while the least frequent were 20, 26, 27, and 33 cm. The recorded weight ranged from 598 to 3242 g (1378.84 ± 686.57 g). This finding is consistent with descriptions reported in other studies, indicating shell lengths of up to 38 cm and weights between 1 and 3 kg (Guevara-Chumacero *et al*., 2017), a maximum shell length of 48 cm and a maximum weight of 5 kg (Lesher-Gordillo, 2019), and lengths between 20 and 60 cm (Gómez-Aguilar *et al*., 2018).

The species *K. leucostomum* recorded CSL values between 13 and 17 cm (14.77 ± 1.48 cm), with a weight interval of 226 to 478 g (318.22 ± 90.68 g). The highest CSL frequency recorded was 15 cm. The values for this species presented lower average length and weight (12.4 ± 1.52 cm and 273.3 ± 68.2 g) than those reported in another study conducted in Tabasco (Hernández-Guzmán *et al*., 2014). Similar records were reported in Ecuador and Colombia (Rodríguez-Murcia *et al*., 2014), with maximum shell lengths of 13.7 cm in females and 14.8 cm in males.

Sex ratio by species

The estimated sex ratio for *R. areolata* was 1:1.81 (female:male). These data are similar to those recorded in Mexico and Belize (Vogt *et al*., 2009), of 1:2 and 1:1.3, respectively. On the other hand, similar values have been reported for wild turtles in the Colombian Pacific region, as well as for *R. melanosterna*, with a ratio of 1:2.3 (Rengifo-Palacios *et al*., 2022).

For *T. venusta*, the sex ratio in this study was 1:1.28 (female:male). This differs from that reported for *Trachemys scripta elegans* in Spain, where the proportion is 9:1 (Patiño-Martínez and Marco, 2005). This suggests that sex ratios in turtles, both in the wild and in captivity, can vary.

In *K. leucostomum*, the sex ratio was 4:1 (female:male), differing from that recorded for wild populations in Colombia (1:1.75) (Rodríguez-Murcia *et al*., 2014). In another species of the same family (*Kinosternon scorpioides*), a ratio of 1:1.6 has been reported in Costa Rica (Acuña-Mesén and Márquez, 1993). This difference is probably due to the low abundance of this species in the captivity area of the DAMR.

Evaluation of the physical conditions of the turtles

A total of 42 % (n = 47) of the turtles were recorded as having optimal physical conditions, while 58 % (n = 66) presented irregular physical characteristics. The most prevalent condition was erosion of the shell and plastron, with 17 % (n = 19), whereas the least prevalent were malformations of the mouth and loss of eyes, with 1 % (n = 1) each (Figure 2).

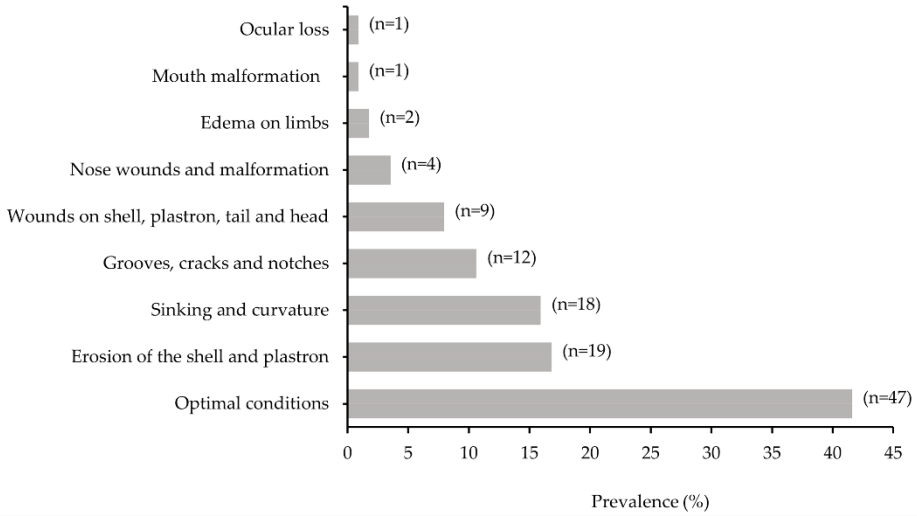

Figure 2. Prevalence of physical conditions in the total freshwater turtle population studied in Los Ríos region of Tabasco, Mexico.

The physical conditions observed in the analyzed individuals indicate that just over half present alterations related to confinement. Lesions on the shell and plastron showed a prevalence of 17 %, a recurring value in organisms under management and lower than that recorded for other species such as *Dermatemys mawii* in the Arca de Noé turtle farm (90 %), the Tabasco State Government turtle farm (55 %), and the Arroyo Tabasquillo turtle farm (75 %) (Rangel-Mendoza and Weber, 2015).

For the species *Rhinoclemmys areolata* (n = 76), 33 % (n = 25) presented optimal physical conditions, while 67 % (n = 51) showed irregular characteristics; the most prevalent were erosion of the shell and plastron, and the least frequent were malformations of the mouth and loss of eyes. This species presented the highest proportion of irregular physical conditions (Figure 3A). The limited available information on physical traits and the prevalence of abnormal conditions in *R. areolata* did not allow comparison with these results; however, there are studies on thermoregulatory behavior, temperament in captivity, and hematological parameters (Cassola *et al*., 2020) aimed at improving health status.

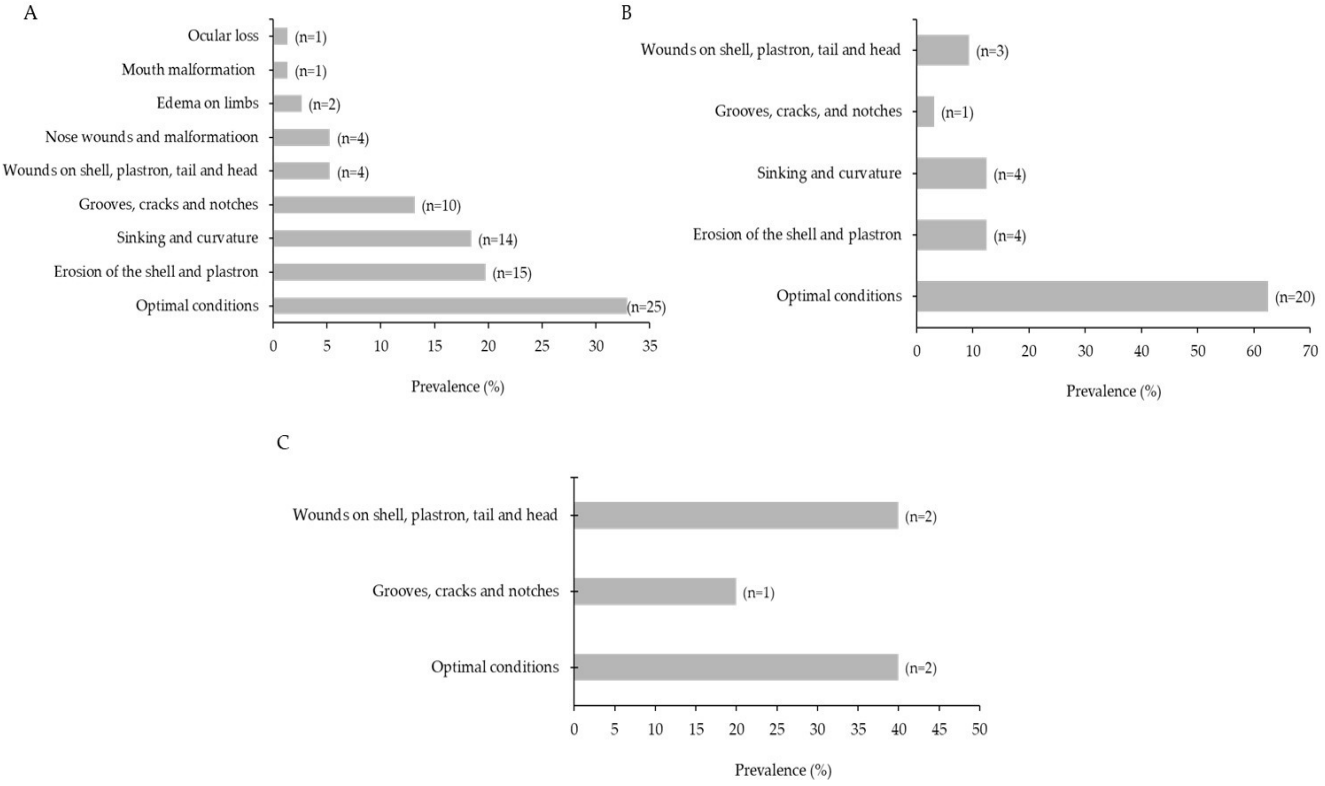

Figure 3. Species-specific prevalence of physical conditions in freshwater turtles from Los Ríos region of Tabasco, Mexico. A: *Rhinoclemmys areolata*; B: *Trachemys venusta*; C: *Kinosternon leucostomum*.

In *T. venusta* (n = 32), 62 % (n = 20) presented optimal conditions, while 38 % (n = 12) displayed irregular characteristics, highlighting erosion of the shell and plastron, as well as depression and curvature (Figure 3B). In *K. leucostomum* (n = 5), 40 % presented optimal physical conditions and 60 % showed irregular characteristics (Figure 3C). For this species, no studies were found on irregular physical characteristics and their prevalence that enable comparison of the observed degrees of impact.

Rhinoclemmys areolata presented the highest prevalence of irregular physical conditions, mainly related to erosion and curvature of the shell and plastron (67 %, n = 51), compared with 38 % (n = 12) in *T. venusta* and 60 % (n = 3) in *K. leucostomum*. This may be related to the greater population density of this species or to nutritional deficiencies, such as vitamin and iron deficiencies (SEMARNAT, 2009). In addition, its more docile behavior and more terrestrial habits promote mechanical damage to the shell through contact with hard or artificial surfaces, which increases its vulnerability relative to the other species with which it shares its habitat, as it is less efficient for swimming or for remaining protected for prolonged periods in bodies of water (Vogt *et al*., 2009).

The wounds observed in these species, kept in a natural area with structural modifications, can be attributed to handling in captivity, water quality, habitat conditions, and terrestrial activities typical of these reptiles (courtship, mating, reproduction, and nesting) (Rangel-Mendoza and Weber, 2015).

Population density

The population density of the three turtle species was dispersed (Table 2). Regarding the density of turtles of the genus *Rhinoclemmys* in captivity, insufficient data were found in the scientific literature to enable a comparative analysis; however, regional and international references derived from master's and doctoral theses are available. For comparison purposes, data for *Rhinoclemmys nasuta* in the wild indicate a density of 0.244 individuals per m^2 (Giraldo *et al*., 2012), which is lower than that recorded for *R. areolata* in this study, although it was assessed over a larger area.

Based on this, researchers consider the population density of the three studied species adequate for their distribution area. Nevertheless, additional areas will be required in the future, as the current space is limited for reproduction, population growth, and the maintenance of a stable population.

Table 2. Population density and dispersion area of freshwater turtle species in Los Ríos region of Tabasco, Mexico.

Species	Density (individuals per m^2)	Dispersion area (m^2)
Trachemys venusta	0.316	3.16
Rhinoclemmys areolata	0.751	1.33
Kinosternon leucostomum	0.049	20.24

Length-weight ratio by species

Rhinoclemmys areolata presented a shell-weight ratio of $P = 0.1935SL^{2.8195}$, with a coefficient of determination of $R^2 = 80.71$ % (Figure 4A). *Trachemys venusta* showed a ratio of $P = 0.1325SL^{2.9222}$, with $R^2 = 97.64$ % (Figure 4B). *Kinosternon leucostomum* presented a ratio of $P = 0.5159SL^{2.3777}$, with $R^2 = 83.95$ % (Figure 4C).

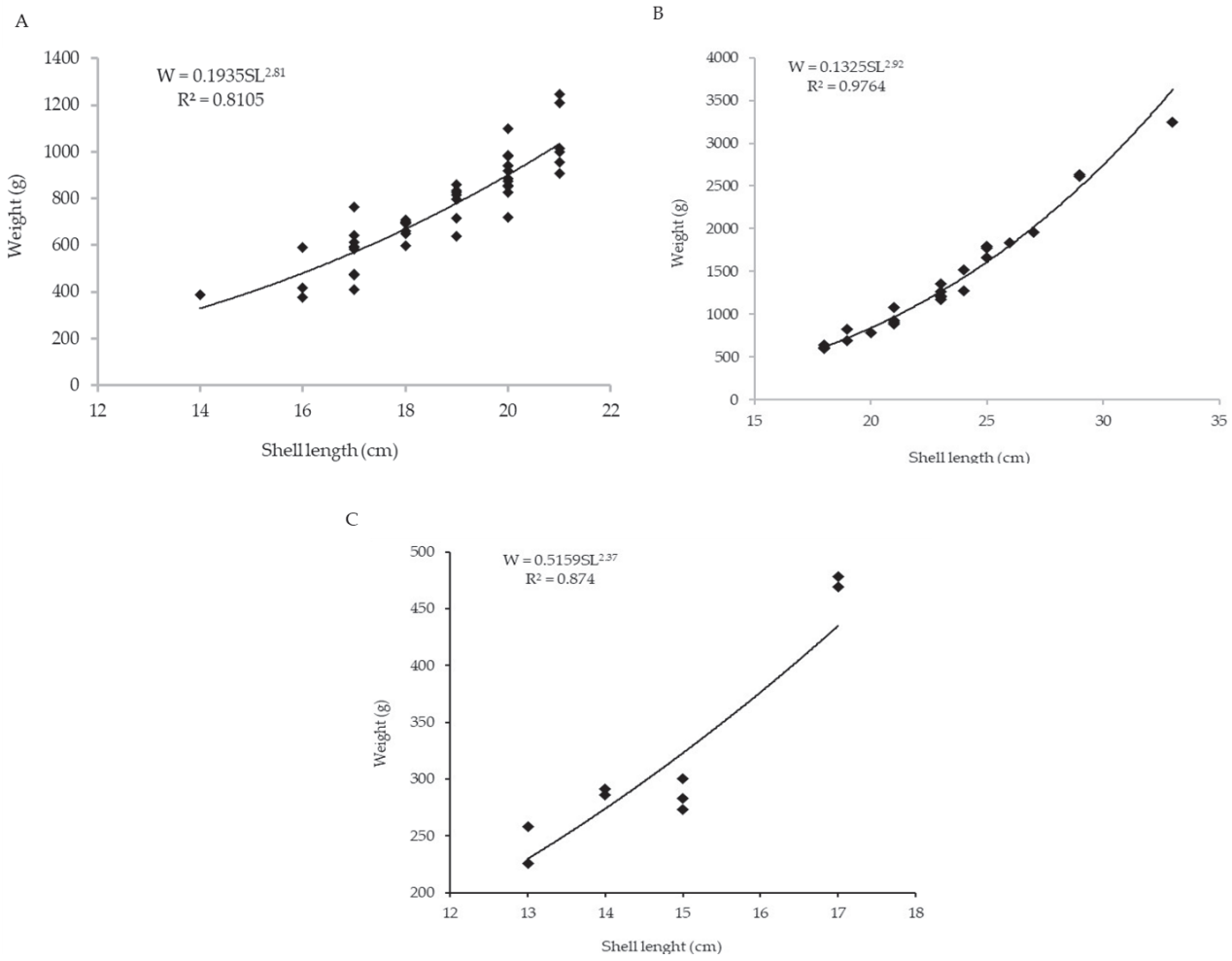

Figure 4. Species-specific shell length-weight (SL-W) ratio in captive freshwater turtles from Los Ríos region of Tabasco, Mexico. A: *Rhinoclemmys areolata*; B: *Trachemys venusta*; C: *Kinosternon leucostomum*.

The freshwater turtles in the DAMR presented an adequate length–weight ratio, indicating uniform growth in length and weight. This behavior is explained by the model in approximately 80 % for *R. areolata* and *K. leucostomum* and in 90 % for *T. venusta*. These values differ from those reported for *Rhinoclemmys nasuta* in the Colombian Pacific, with 96 % (Giraldo *et al*., 2012). For *T. venusta* and *K. leucostomum*, no comparative data were found; however, both species display similar behavior in the model, indicating adequate growth conditions.

CONCLUSIONS

The freshwater turtles in captivity at División Académica Multidisciplinaria de los Ríos (DAMR) generally present adequate physical conditions for their management, as they fall within the parameters established in similar management systems in Tabasco, which have proven effective in maintaining reproductive and long-lived populations. Based on the analyzed data and the review of applicable regulations (General Wildlife Law and the Standard Management Plan), it is concluded that the studied population can be integrated into a conservation scheme and the potential establishment of a Management Unit for Wildlife Conservation (UMA), which would contribute to fulfilling the objectives of reproduction, research, and environmental education within the DAMR.

ACKNOWLEDGEMENTS

The authors thank the División Académica Multidisciplinaria de los Ríos for providing permission, facilities, and support to conduct this work. They also acknowledge the collaborators in fieldwork, technical advisors, and the Secretariat of Science, Humanities, Technology and Innovation (SECIHTI) for granting the main author a scholarship to pursue a Master's degree (No. 790654).

REFERENCES

Acuña-Mesén RA, Márquez BC. 1993. El dimorfismo sexual de *Kinosternon scorpioides* (Testudines: Kinosternidae) en Palo Verde, Costa Rica. Revista de Biología Tropical 41 (2): 261–265.

Arroyo-Quiroz I, Wyatt T. 2019. Tráfico de vida silvestre entre la Unión Europea y México. Crítica Penal y Poder 16: 32–52.

Beauregard-Solís G, Zenteno-Ruiz CE, Armijo-Torres R, Guzmán-Juárez E. 2010. Las tortugas de agua dulce: patrimonio zoológico y cultural de Tabasco. Kuxulkab' 17 (31): 5–20. https://doi.org/10.19136/kuxulkab.a17n31.394

Cagle FR. 1939. A system of marking turtles for future identification. Copeia 1939 (3): 170–173. https://doi.org/10.2307/1436818

Casas-Andreu G. 1965. Estudio preliminar sobre las tortugas de agua dulce en México. Anales del Instituto Nacional de Investigaciones Biológico-Pesqueras 1: 363–401.

Cassola FM, Henaut Y, Cedeño-Vázquez JR, Méndez-de la Cruz FR, Morales-Vela B. 2020. Temperamento y comportamiento sexual en la tortuga de madera surcada *Rhinoclemmys areolata*. PLoS ONE 15 (12): e0244561. https://doi.org/10.1371/journal.pone.0244561

Cázares-Hernández E. 2015. Guía de las tortugas dulceacuícolas de Veracruz. Industria Gráfica Internacional: Xalapa, México. 65 p.

Chavarría-Pérez I, Melara-Soriano M, Liles MJ, Castro-Menjivar J. 2020. Determinación de la proporción sexual de tortugas Carey (*Eretmochelys imbricata*) inmaduras, mediante la detección de testosterona por medio de la prueba de ELISA. Revista Agrociencia 3 (16): 73–80. https://doi.org/10.5281/zenodo.10912976

Díaz-Gamboa LF, Gallardo A, Alemán MB, Chiappa-Carrara X, Arenas CY. 2022. Guía rápida de identificación tortugas terrestres y dulceacuícolas de la Península Mexicana de Yucatán. Universidad Nacional Autónoma de México. Ciudad de México, México. https://doi.org/10.13140/RG.2.2.29504.40967

DOF (Diario Oficial de la Federación). 2010. Norma Oficial Mexicana NOM-059-SEMARNAT-2001. Protección ambiental-Especies nativas de México de flora y fauna silvestres. Categorías de riesgo y especificaciones para su inclusión, exclusión o cambio. Lista de especies en riesgo. Gobierno de México. Secretaría del Medio Ambiente y Recursos Naturales. Ciudad de México, México.

Eckert KL, Bjorndal KA, Abreu-Grobois FA, Donnelly M. 2000. Técnicas de investigación y manejo para la conservación de las tortugas marinas. Grupo Especialista en Tortugas Marinas. Unión Internacional para la Conservación de la Naturaleza. Comisión de Supervivencia de Especies. Blanchard, PA, USA. 270 p.

Gallina S, López-González C. 2011. Manual de técnicas para el estudio de la fauna. Volumen I. Universidad Autónoma de Querétaro. Instituto de Ecología A.C. Querétaro, México. 377 p.

Giraldo A, Garces-Restrepo MF, Carr JL, Loaiza J. 2012. Population size and structure of the large-nosed wood turtle (*Rhinoclemmys nasuta*, Testudines: Geoemydidae) in an insular environment of the Colombian pacific region. Caldasia 34 (1): 109–125.

Gómez-Aguilar V, Ramos-Ferrer JA, Mendiola-Campuzano JVH. 2018. Eficacia de la aplicación de la normatividad protectora de las hicoteas (*Trachemys venusta*) en Tabasco. Kuxulkab' 24 (49): 19–30. https://doi.org/10.19136/kuxulkab.a24n49.2623

Guevara-Chumacero M, Pichardo-Fragoso A, Martínez-Cornelio M. 2017. La tortuga en Tabasco: comida, identidad y representación. Estudios de Cultura Maya 49: 97–122. https://doi.org/10.19130/iifl.ecm.2017.49.758

Hernández-Guzmán J, Rimber IJ, Shigueki YG, Arias-Rodríguez L. 2014. Los cromosomas de las tortugas tropicales: *Kinosternon leucostomum*, *Trachemys scripta* y *Staurotypus triporcatus* (Testudines: Kinosternidae/Emydidae). Revista de Biología Tropical 62 (2): 671–688. https://doi.org/10.15517/rbt.v62i2.10843

Hernández-Silva DA, Pulido SMT, Zuria I, Gallina TSA, Sánchez-Rojas G. 2018. El manejo como herramienta para la conservación y aprovechamiento de la fauna silvestre: acceso a la sustentabilidad en México. Acta Universitaria 28 (4): 31–41. https://doi.org/10.15174/au.2018.2171

Ippi S, Flores V. 2001. Las tortugas neotropicales y sus áreas de endemismo. Acta Zoológica Mexicana 84: 49–63.

Lesher-Gordillo JM. 2019. Establecimiento de un programa de manejo genético para las especies *Dermatemys mawii* (tortuga blanca) y *Trachemys venusta* (tortuga hicotea) en Unidades de Manejo de Vida Silvestre (UMA) para favorecer el flujo genético y la conectividad del Corredor Biológico Mesoamericano en Tabasco. Universidad Juárez Autónoma de Tabasco. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. Ciudad de México, México. 34 p.

Mandujano-Rodríguez S. 2011. Conceptos generales de ecología poblacional en el manejo de fauna silvestre. *In* Manual de Técnicas para el Estudio de la Fauna, Volumen I. Instituto de Ecología A.C. Universidad Autónoma de Querétaro. Querétaro, México, pp: 37–60.

Patiño-Martínez J, Marco A. 2005. Potencial invasor de los galápagos exóticos en el País Vasco. Munibe Ciencias Naturales 56: 97–112.

Quevans N, Falcón N, Elías R. 2013. Fauna silvestre y productos derivados decomisados durante el período 2000-2007, Lima–Perú. Salud y Tecnología Veterinaria 1 (1): 14–18. https://doi.org/10.20453/stv.v1i1.105

Rangel-Mendoza JA, Weber M. 2015. Evaluation of the physical condition of the white turtle, *Dermatemys mawii*, under captivity conditions in Tabasco, Mexico. Agrociencia 49 (5): 499–511.

Rengifo-Palacios MY, Cetré-Mosquera Z, Rengifo-Mosquera JT, Halaby-Guerrero JC. 2022. Estructura poblacional y estado de conservación de *Rhinoclemmys melanosterna* Gray, J.E, 1861 (Orden: Rhinoclemmys familia Geoemydidae) en ecosistemas perturbados por minería en el Pacífico colombiano. Boletín Científico Centro de Museos, Museo de Historia Natural 26 (1): 99–110. https://doi.org/10.17151/bccm.2022.26.1.7

Ricker WE. 1975. Cálculo e interpretación de estadísticas biológicas de poblaciones de peces. Boletín de la Junta de Investigación Pesquera de Canadá 191: 1–38.

Rodríguez-Murcia JD, Giraldo A, Garcés-Restrepo M, Sánchez F. 2014. Estructura poblacional y dimorfismo sexual de *Kinosternon leucostomum* (Testudines: Kinosternidae) en un sistema de charcas asociadas con el río Purnió, Caldas, Colombia. Investigación, Biodiversidad y Desarrollo 33 (2): 86–95.

SEMARNAT (Secretaría de Medio Ambiente y Recursos Naturales). 2009. Taller de capacitación para la conservación y aprovechamiento sustentable de tortugas dulceacuícolas del Sursureste de México. Secretaría de Medio Ambiente y Recursos Naturales. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Comisión Natural de Áreas Naturales Protegidas. Procuraduría Federal de Protección al Ambiente. Catemaco, México. https://biblioteca.semarnat.gob.mx/janium/Documentos/Ciga/libros2009/SF408T352009.pdf (Retrieved: June 2021).

SEMARNAT (Secretaría del Medio Ambiente y Recursos Naturales). 2018. México biodiversidad que asombra. Gobierno de México. Secretaría del Medio Ambiente y Recursos Naturales. Ciudad de México, México. https://www.gob.mx/semarnat/es/articulos/mexico-biodiversidad-que-asombra (Retrieved: May 2021).

Tellería JL. 2013. Loss of biodiversity: Causes and consequences of the species loss. Memorias de la Real Sociedad Española de Historia Natural 10: 13–25.

van Uhm DP. 2016. Monkey business. *In* The Illegal Wildlife Trade. Studies of Organized Crime (Volume 15). Springer: Cham, Switzerland. https://doi.org/10.1007/978-3-319-42129-2_8

Venter O, Sanderson EW, Magrach A, Allan JR, Beher J, Jones KR, Possingham HP, Laurance WF, Wood P, Fekete BM, *et al*. 2016. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. Nature Communications 7 (1). https://doi.org/10.1038/ncomms12558

Vogt RC, Platt SG, Rainwater TR. 2009. *Rhinoclemmys areolata* (Duméril and Bibron 1851): Furrowed wood turtle, black-bellied turtle, mojena. *In* Rhodin AGJ, Pritchard PCH, van Dijk PP, Saumure RA, Buhlmann KA, Iverson JB, Mittermeier RA. (eds.), Conservation Biology of Freshwater Turtles and Tortoises: A Compilation Project of the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group. Chelonian Research Foundation: Arlington, VT, USA. https://doi.org/10.3854/crm.5.022.areolata.v1.2009

Yvan L, Charette R, Leyva GFA. 1999. CITES identification guide – Turtles and tortoises. Guide to the identification of turtles and tortoises species controlled under the Convention on International Trade in Endangered Species of Wild Fauna and Flora. Ottawa, Canada. 232 p.

VOLUME 60, NUMBER 4 | May 16 - June 30 , 2026 | MEXICO