Artículo de Investigación. Revista Ciencia e Ingeniería. Vol.47, No.3 pp. 293−306, agosto-noviembre, 2026.
ISSN 1316-7081. ISSN Elect. 2244-8780 Universidad de los Andes (ULA)

Multivariate analysis of köppen-geiger climate similarity and discriminant factors across 70 major cities in Latin America and the Caribbean

Análisis Multivariante de la similitud climática köppen-geiger y factores discriminantes en 70 metrópolis de Latinoamérica y el Caribe

González-Calderón, Álvaro[1]*; Pacheco-Angulo, Carlos[1]; Peña-Guillen, Jesús[2]; Bianchi, Guillermo[3]; Yarbuh, Ismael[4]; González, Héctor[5]
[1] Faculty of Forestry and Environmental Sciences, University of The Andes, Venezuela.
[2] Faculty of Pharmacy and Bioanalysis. Chair of biostatistics, University of The Andes, Venezuela.
[3] Faculty of Science, Chair of biostatistics, University of The Andes, Venezuela.
[4] Faculty of Marine Sciences, Autonomous University of Baja California, Ensenada, Baja California, Mexico.
[5]Faculty of Engineer, University of The Andes, Venezuela.
***alvarof1971@gmail.com**

Abstract

The study assesses climatic similarity according to the Köppen-Geiger classification in 70 metropolitan areas in Latin America and the Caribbean (2014–2018), based on the integration of 25 urban, bioclimatic, geographic, socioeconomic, and sociopolitical indicators. Two datasets were constructed: one incorporating average elevation (metropolis) and another classifying elevation (low (1), intermediate (3), and high (3)). Multivariate statistical methods were employed, including cluster analysis (MC) and discriminant analysis (MD) based on the dependent variables (VDsc): Climate Zone, Population Stratification, and Geographic Location. Data validation was performed using Student’s t-test, analysis of variance (ANOVA) for independent samples, and Tukey’s post hoc test. Cluster structures were defined using the silhouette method, and Euclidean distance metrics were employed to assess similarity patterns.The results indicate that VDsc Climate Zone shows the highest similarity among urban groups (61%), suggesting relatively strong Köppen-Geiger climate consistency in the region. In turn, the MD (VDsc Climate Zone) yielded a low classification error (5.6%), highlighting the robustness of the clustering approach. Among the discriminating indicators that differentiate clusters are the Human Development Index and the Rule of Law Index. These results suggest that regional urban climate similarity does not depend solely on bioclimatic factors, but also on existing socioeconomic and governance interrelationships, allowing for a broader interpretation of climate patterns in complex metropolitan systems.

Keywords: Köppen-Geiger climate similarity, dependent variables, multivariate models, discriminant indicators.

Resumen

El estudio evalúa la similitud climática según la clasificación Köppen-Geiger en 70 áreas metropolitanas de América Latina y el Caribe (período 2014-2018), a partir de la integración de 25 indicadores urbanos, bioclimáticos, geográficos, socioeconómicos y sociopolíticos. Se construyeron dos conjuntos de datos: uno que incorpora la elevación media (metrópolis) y otro que discrimina la elevación baja (1), intermedia (3) y alta (3). Se emplearon métodos estadísticos multivariados, incluidos los análisis de conglomerados (MC) y discriminante (MD) a partir de las variables dependientes (VDsc): Zona Climática, Estratificación de la Población y Localización Geográfica. La validación de los datos se realizó con la prueba t de Student, análisis de varianza (ANOVA) para muestras independientes y prueba post hoc de Tukey. Las estructuras de los clústeres se definieron con el método de silhouette, y se emplearon métricas de distancia euclidiana para evaluar los patrones de similitud. Los resultados indican que VDsc Zona Climática presenta la mayor similitud entre los *grupos urbanos (61 %), lo que sugiere una coherencia climática Köppen-Geiger relativamente fuerte en la región. A su vez el MD (VDsc Zona climática) arrojó un bajo error de clasificación (5.6 %), lo que destaca la solidez del enfoque de agrupamiento. Entre los indicadores discriminantes que diferencian clústeres, tenemos los Índices de Desarrollo Humano y Estado de Derecho. Estos resultados sugieren que la similitud climática urbana regional no depende únicamente a factores bioclimáticos, sino a las interrelaciones socioeconómicas y de gobernanza existentes, lo que permite interpretar con mayor amplitud los patrones climáticos en sistemas metropolitanos complejos.*

Palabras clave**:** *Similitud climática Köppen Geiger, variables dependientes, modelos multivariantes, indicadores discriminantes.*

1. Introduction

The global population reached 8 billion people in 2022, of whom approximately 55% live in urban areas (United Nations Sustainable Development Goals, UN, 2023). In Latin America and the Caribbean (LAC), the urban areas of its metropolises are quite heterogeneous in terms of population strata, accounting for 80% of the regional population and 13% of the global population (UN-HABITAT, 2016). This population impact directly affects the regional metropolitan climate, according to scientific evidence (Hiroito et al., 2026; Wu et al., 2019). However, climate studies in metropolitan areas can be approached from different perspectives, depending on the urban, economic, political, and geographic realities of each specific region. Furthermore, these studies may have limitations in terms of the scales of analysis, which increases uncertainty and information gaps in climate projections (Lauwaet et al., 2015).

In studies on climate and socio-environmental dynamics in urban areas across Latin America and the Caribbean (LAC) over the past decade, regional and international experts have focused their efforts on geospatial analyses at the local metropolitan scale (Boccolini, 2025; Pacifici et al., 2019), the national scale (Casadei et al., 2021; Baumgartner, 2025; Henríquez et al., 2017), and regional (Wu et al., 2019; Dobbs et al., 2014) scales. In climate and cartographic analysis at various scales, the Köppen-Geiger climate classification has been used as a methodological framework due to its regular updates (Beck et al., 2018). However, various authors over the past 30 years (Grimmond and Souch, 1994) have recommended developing new cartographies based on Köppen-Geiger but better adapted to urban conditions in LAC, as seen in regional studies over the past decade (González-Calderón et al., 2025a; González-Calderón et al., 2025b).

However, to contribute new climate-related findings in metropolitan areas, it is necessary to include approaches that examine the relationships between Köppen-Geiger climate similarity and socioeconomic indicators (carbon, energy, population, income, pollution), urban indicators (heat islands, urban sprawl, population density), and sociopolitical indicators (climate risks, global risks, governance, rule of law)to name the most critical ones as well as new analytical frameworks that allow for the observation urban and climatic complexities in the region.

With regard to the analysis and assessment of Köppen-Geiger climate similarity in metropolitan areas, such analyses can be conducted at the regional metropolitan scale using multivariate models; these analyses offer several advantages: 1) it constitutes a non-parametric method that does not require specific probability distributions nor depend on strict statistical assumptions, such as linearity or symmetry, and allows for the inclusion of categorical indicators, such as climate codes; 2) it allows Köppen-Geiger subclassifications to be grouped into clusters that are internally homogeneous and distinct from one another; and 3) it facilitates the graphical visualization of the shortest distance between Köppen-Geiger subclassifications through probabilities of belonging to the same cluster, using estimation methods such as Euclidean distance matrices, specifically as a metric to evaluate similarity patterns among Köppen-Geiger metropolitan groups.

To conduct these analyses, the study was structured around two main criteria. The first criterion involved the selection of dependent variables based on: a) the Köppen-Geiger climate zone of the analyzed metropolitan areas, b) the population stratification of the metropolitan areas, and c) their geographic location. The second criterion was the application of multivariate cluster models (MC) to organize metropolitan groups based on climatic similarity, and the application of discriminant analysis (DA) for the analysis and validation of the cluster data generated by the MC. In accordance with these criteria, the main objectives of the study are: 1) To organize metropolitan areas according to low1, medium3, and high5 elevations. 2) To organize Köppen-Geiger clusters using MC based on climatic similarity and employing dependent variables (Climate Zone, Population Stratification, and Geographic Location). 3) Validate misclassification errors in the data using MD and identify the most discriminating indicators (D) in Köppen-Geiger clusters, which will allow for differentiation between these clusters and the climate clusters.

2 Methodology

2.1 Description of the study area

The Latin America and Caribbean (LAC) region comprises 33 countries, located in three subregions: South America, Mesoamerica, and the Caribbean. This region is also characterized by significant climatic and geographic variability according to the recent Köppen-Geiger climate classification (Beck et al., 2018).

In particular cities in Latin America and the Caribbean (LAC) exhibit significant population stratification (UN-Habitat, 2016) and a wide variety of urban landscapes depending on their Köppen-Geiger classification (see Figure. 1). Therefore, the study proposed including seventy major metropolises (Table 1) based on the high diversity of climates and the population significance of these metropolises (Economic Commission for Latin America, ECLAC 2017. As an example, Figure. 1 shows a random selection of 16 types of LAC metropolises based on their different Köppen-Geiger climates; furthermore, Figure. 1 illustrates a sample of the great diversity of urban landscapes existing in the region, indicating the high level of urban-climatic heterogeneity in LAC. These examples of metropolises and their respective climates are: 1) arid-desert (Maracaibo (Bsh); Tijuana (Bsk); Lima (Bwh); Arequipa (Bwk); 2) temperate (Buenos Aires (Cfb); k) Quito (Cfb); l) Greater Valparaíso (Csb); Santiago, Chile (Csc); Guatemala City (Cwb); La Paz (Cwc); 3) tropical (Rio de Janeiro (Af); b) Santo Domingo (Am); c) Cali (As); d) Panama City (Aw); 4) subtropical (Guadalajara (Cwa). Furthermore, Table 1 details each of the 70 metropolitan areas included in this study and organizes them in to the Köppen-Geiger climate groups to which they belong according to the updated regional classification (Beck et al., 2018).

Figure.1. Location of metropolises Köppen–Geiger climate **a**) Rio de Janeiro (Af); b) Santo Domingo (Am), c) Cali (As); d) Ciudad de Panamá (Aw); e) Maracaibo (Bsh); f) Tijuana (Bsk); g) Lima (Bwh); h) Arequipa (Bwk); i) Rosario (Cfa); j) Buenos Aires (Cfb); k) Quito (Cfb); l) Gran Valparaíso (Csb); m) Santiago de Chile (Csc), n) Guadalajara (Cwa); o) Ciudad de Guatemala (Cwb); p) La Paz (Cwc).

Table 1. Metrópolis according to their location Köppen-Geiger [(a)]

CLIMATE	KÖPPEN-GEIGER	SUB-GROUP	METROPOLISES
	Af	Ecuatorial	Salvador, Manaos, Belem, Rio de janeiro, Santos
Tropical	Am	Monsoon	Joao Pessoa, Managua, Santo Domingo, San Juan, Sao Luis, Recife, San Pedro Sula, Bucaramanga, Cali, Medellin
	As (Aw)	Tropical savanah	Caracas, Ciudad de Panamá, Fortaleza, La Habana, Natal, Brasilia, Santa-Cruz, Barranquilla, Guayaquil, Mérida, Goiania, Maracay, San Salvador, Tegucigalpa, Valencia, Vitoria.
Sub tropical	Cwa	Subtropical with dry winters	Belo horizonte, Campinas, Córdoba, Guadalajara, San Miguel de Tucumán, Cuernavaca, San José.
	Cfa	Humid subtropical	Asunción, Puerto Alegre, Rosario, Sao Paulo.
Arid	Bsh	Steppes (hot)	Maracaibo, Monterrey, León, Querétaro, Cartagena,
	Bsk	Steppes (cold)	Cochabamba, Tijuana, San Luis Potosí.
Desert	Bwh	Desertic (hot)	Lima, Torreon, Mexicali
	Bwk	Desertic (cold)	Ciudad Juarez, Mendoza, Arequipa
Temperate	Csb	Mediterranean with fresh summers	Concepción, Valparaíso
	Csc	Cold-summer Mediterranean	Santiago de Chile.
	Cwc	Cold subtropical	La Paz
	Cfb	West coast maritime (oceanic)	Buenos Aires, Curitiba, Montevideo, Bogotá, Quito,
	Cwb	Temperate with dry winters	Toluca, C.D México, Morelia, Puebla, C.D Guatemala

(a) 70 cities included in this study; the climate code colors correspond to the latest classification (Beck et al., 2018).

2.2 Data on selected indicators

A five-year period (2014–2018) was selected, using robust data on meteorological, bioclimatic, geographic, urban, socioeconomic, and sociopolitical indicators for LAC metropolises. This time frame was chosen because an updated Köppen-Geiger classification map was reported for the 2014–2018 period (Beck et al., 2018) and periodic data were reported for new sociopolitical indices (democracy, rule of law, global risks, climate risks, and social gaps) and socioeconomic indices (carbon emissions, particulate matter, human development, social inequality, and population). The indicator data were extracted from reports by nongovernmental organizations, scientific articles, and global platforms. In the first case, the indicators and abbreviations for bioclimatic variables (temperature (T), precipitation (P), humidity (H), and aridity (Ia)) were taken from the World Meteorological Organization (WMO, 2021), geographic–urban indicators (altitude (alt), urban area (UA), urban heat islands (UHI)) from Wu et al. (2019), and population (P), population density (Pd), and urban expansion (UG) (Cox, 2025).

Socioeconomic data including population (Pop), social inequality (GINI), income (GDP), and human development (HDI)—were drawn from data provided by the World Bank (2021) and the United Nations Development Programme (UNDP, 2018); pollution (PM2.5-PM10) from reports by the World Health Organization (WHO, 2018), and carbon emissions (Cf-CO2U) from Morán et al. (2018). Specifically, the energy data for the 70 metropolitan areas were recalculated for this study in accordance with the IPCC Guidelines (2006) and the protocol for national greenhouse gas inventories; the calculation focuses on quantifying fuel combustion and electricity consumption by sector. For this energy calculation in kilowatt-hours (kWh), data on carbon emissions in metropolitan areas (Moran et al., 2018) were used, and emissions were broken down by sector (residential, commercial, industrial, transportation) according to sector-specific emission factors for the local electricity grid, using Equation 1 (IPCC Guidelines, 2006).

$$KWh = Em\ (KG\ CO2\ e) / FE\ (Kg\ CO2e/Kwh). \qquad (1)$$

Where KWh is the kilowatt-hour of electricity consumption, Em is the carbon dioxide equivalent emissions per sector, and FE is the carbon dioxide equivalent emission factor per sector, expressed in kilowatt-hours per year.

Finally, the selected sociopolitical indicators were: the Democracy Index (DI) (Economic Intelligence Unit, EIU, 2015, 2019), the rule of law as measured by the World Justice Project (WJP, 2021), the Global Climate Risk Index (CRI) (Germanwatch, 2021), the Multidimensional Poverty Index (MPI) from the United Nations Development Programme (UNDP, 2023, 2024), and the Global Risks Index (WR) (World Economic Forum, 2024).

2.3 Statistical Tests and Multivariate Methods

To organize the data for the 70 metropolitan areas, a recoding scheme was applied in the database (APPENDIX 1) based on the following dependent variables: 1) Köppen-Geiger climate classification (Beck et al., 2018), population stratum (United Nations Habitat Organization (UN-Habitat), 2020) and Geographic Coordinates (Geographic Coordinate System (GCS)). Subsequently, two databases were organized for each dependent variable (DVcs), including or excluding the altitude (P) of the metropolises, based on: a) Köppen-Geiger Climate Zone (DVcs-Climate Zone, DVcs-Climate ZoneP); b) Population Stratification (DVscPopulationStratification,DVcsPopulationStratificatio nP); c) Geographic Location (DVcs-Geographic Location, DVcs-Geographic LocationP), resulting in a total of six analyses.

Using the recoded data, a Köppen-Geiger cluster analysis was performed using a conglomerate model (CM) with the statistical software R versión 4.0.2 **(**R Development Core Team, 2023) where the silhouette method was applied to measure the quality and appropriate number of clusters, and to indicate how well each observation unit fits within its own cluster; therefore, a Euclidean distance matrix was established for each case, which allowed for the initial definition of some climate groups. The maximum value of the silhouette curve will indicate the appropriate number of clusters, based on the coefficient that estimates the average distance between clusters. Subsequently, for the analysis and validation of the data, the following methods and procedures were employed:

1**. Application of the Student's t-test for the difference in means between independent samples.** This test was used to compare the means of two groups using 95% confidence intervals, which estimate the difference between population means. Confidence intervals consist of a lower limit (Li) and an upper limit (Ls), from which it can be seen that both limits may have positive or negative signs, or the lower limit may have a negative sign and the upper limit a positive sign; that is, the value zero is included between the two confidence limits as a possible value for estimating the difference in population means. If the 95% confidence interval does not contain the value zero, this implies that there are statistically significant differences between the means of the indicator being compared for both groups; consequently, this particular indicator has the ability to discriminate between the groups because it estimates that their means are different (one mean being greater or less than the other). Otherwise, it is concluded that the indicator does not discriminate between the groups, because their means are equal. For the comparison of means of three or more groups, one-way analysis of variance (ANOVA) for independent samples was applied, after verifying compliance with statistical assumptions, based on the following statistical hypotheses:

***Null hypothesis (H0)**:* the means of the three or more groups are equal.

Alternative hypothesis (H1): the means of at least two of the groups are different.

2. The decision criterion is as follows: if the p-value of the F-test statistic is less than or equal to 0.05, H0 is rejected; otherwise, it is not. In this way, the indicators that showed statistically significant differences were identified. If H0 is not rejected for a particular indicator, that indicator will not be included in the discriminant function because it lacks the ability to distinguish between groups. Conversely, if H0 is rejected for a particular indicator, it will be included in the discriminant function because it has the ability to distinguish between groups; however, the rejection of H0 does not imply that all the means of each group differ from one another.

3. After rejecting H0 and accepting H1, Tukey's test was applied to identify groups with differing means; however, for the purpose of constructing the discriminant function, the ultimate goal is to determine whether there are statistically significant differences between the group means, regardless of whether there are differences between pairs of means when making a comparison (means of three or more groups).

4. To validate the data and test the null hypothesis, a discriminant model (DM) was applied to assess statistically significant differences in the MC clusters, and one-way ANOVA was calculated to validate the discriminant indicators (D) within the dataset ($p < 0.0001$). Next, the data membership was evaluated using a random partition of the information, with 70% selected as the training set and the remaining 30% used to validate each group model. Finally, the selection of the D (clusters) was based on the highest mean values in the clusters (group means) according to the results of the linear discriminant equations (LD).

3. Results and Discussion

3.1 Cluster organization

The Euclidean distance matrix (see Figure. 2) allowed us to define climate groups based on the principal diagonal; the squares (bright red) represent a distance of zero, corresponding to the same statistical unit. As the color shifts from red to blue, the distance between climates increases, and consequently, they belong to different clusters. In this case, pure red represents dist (climate.i, climate.j) = 0 and pure blue represents dist (climate.i, climate.j) = 1. Furthermore, climates belonging to the same cluster are displayed in consecutive order.

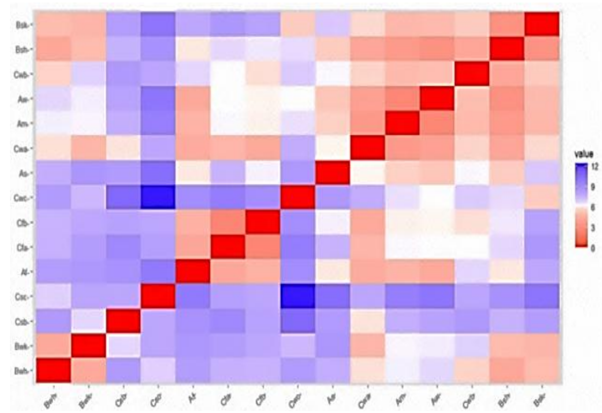


Figure. 2. Matrix of Euclidean distances (Köppen-Geiger climates).

Using R software version 4.0.2 (R Development), the percentage of climatic similarity within each cluster was evaluated based on the VDsc, with and without P. The silhouette method suggested a general configuration of between 2 and 8 clusters, as indicated by the peak of the silhouette curve (see Figure. 3): VDsc-Climate Zone (2), VDsc-Climate Zone P (3), VDsc-Population Stratification (5), VDsc-Population stratification P (5), VDsc-Geographic location (8), VDsc-Geographic Location P (6). However, for VDsc-Climate Zone and VDsc-Climate ZoneP, the silhouette curve recommends a smaller number of clusters (2 and 3, respectively). This implies greater climatic homogeneity in these VDsc categories.

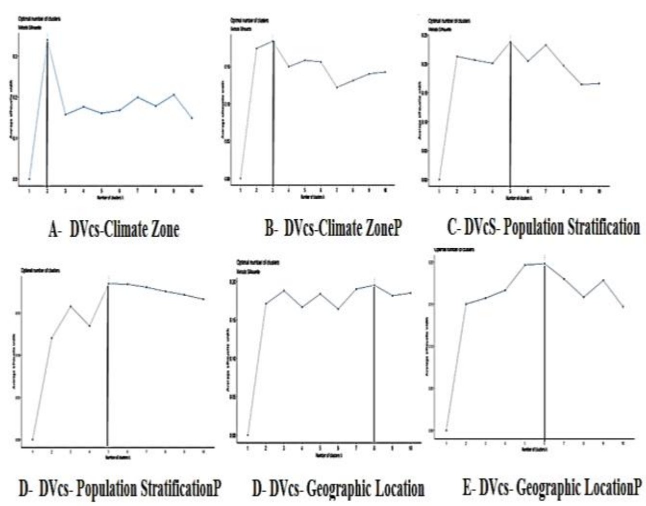


Figure. 3. Number of clusters recommended by the silhouette method

3.2 Köppen-Geiger similarity and number of clusters

The DVcs-Climate Zone yielded the highest similarity percentage (61%), followed by the DVcs -Climate Zone P (51.5%). In the other cases, lower percentages were obtained, such as DVcs -Population stratification (50.1%), DVcs-Population stratificationP (46.7%), DVcs-Geographic location (47.8%), and DVcs-Geographic locationP (50.1%).

The inclusion or exclusion of P resulted in differences in similarity, mainly in the climate zone, suggesting that P may influence the dispersion of cluster data. On the other hand, the clusters and dendrograms (detailed in Fig. 4) indicate that as similarity increases, the silhouette method suggests a smaller number of clusters, reflecting greater climatic homogeneity. In contrast, when similarity decreases, the silhouette method suggests a larger number of clusters (6–8), indicating more heterogeneous and diverse climates. The results of the decision-making process regarding the number of Köppen-Geiger clusters for each DVcs are described below.

DVcs-Climate Zone

The DVcs climate zone yielded two main clusters (Figure. 4a). Cluster 1 included desert, hot, cold, and temperate climates according to Bwh, Bwk, Csb, and Csc. Cluster 2 grouped tropical and subtropical climates (Af, Cfa, As, Am, Aw, and Cwa), as well as two subgroups: one temperate (Cfb, Cwc, and Cwb) and one semi-arid (Bsh, Bsk).

DVcs-Climate Zone P

The DVcs-climate zone P comprises three clusters (Figure. 4b). Cluster 1 grouped altitudinal zones 1, 3, and 5 of the tropical-subtropical type (1Af, 1Am, 1Aw, 1Cfa, 3Af, and 3Cwa), as well as tropical, subtropical, and temperate climates of altitudinal zone 5 (5As, 5CWa, and 5Cfb). Cluster 2 grouped temperate climates at elevation 5 (5Cwc and 5Cwb), arid-desert climates (5Bsk, 5Bwh, and 5Bsh), tropical high-altitude climates (5Aw), tropical low-altitude climates (3Am, 3Aw), and a warm arid low-altitude climate (1Bsh). Finally, cluster 3 grouped subtropical, temperate, and arid-desert climates of elevation 3 (3Cfa, 3Csc, 3Bwk, 3Bsh, and 3Bwh), as well as a temperate-arid subgroup of zone 1 (1Bsk, 1Cfb, and 1Bwh) and a single desert climate of zone 5 (5Bwk).

DVcs- Population Stratification

The DVcs-Population Stratification analysis identified five clusters (Figure. 4c). Cluster 1 grouped megacities with tropical, subtropical, and temperate climates (AfMC, CfaMC, CfbMC, and CwbMC). Cluster 2 grouped temperate climates with medium-to-large population strata (CsbMI and CscMG). Cluster 3 grouped an arid-semi-arid subgroup that includes a desert megacity (Bwh) and climates with medium-to-intermediate population strata (BskMM, BshMM, BwhMM, BwkMI, BwhMI, and BwkMM), as well as a temperate climate with a medium population stratum (CwbMM). Cluster 4 grouped intermediate-to-medium-stratum tropical and temperate climates (AmMI and CwcMM). Finally, Cluster 5 corresponded to a major subgroup of medium-intermediate population strata with tropical, subtropical, temperate, and arid climates (AfMM, AwMM, AmMM, AsMM, CwaMI, CwaMM, CfaMM, CfbMM, and CfbMG BshMI).

DVcs- Population Stratification

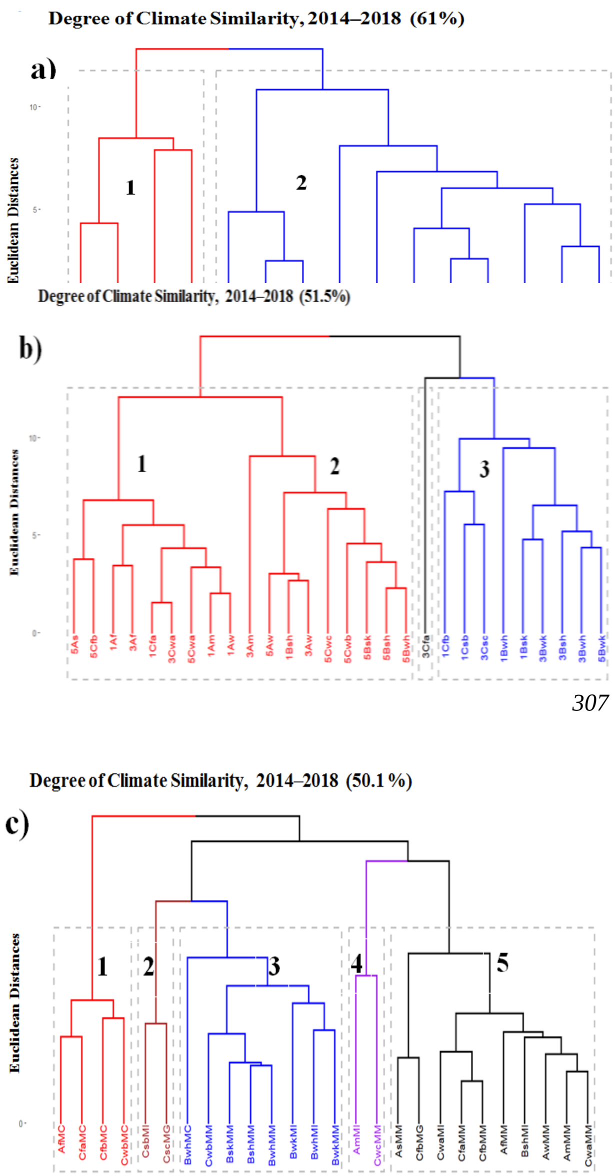
The DVcs -population stratification P analysis reveals five clusters (Figure. 4d). Cluster 1 comprises tropical, subtropical, and temperate megacities in altitudinal zones 1, 3, and 5 (1AfMC, 1CfbMC, 3CfaMC, 5CwbMC). Cluster 2 grouped temperate, subtropical, and desert climates with intermediate, medium, and large population strata at elevations 1, 3, and 5 (1CfbMM, 3BwkMM, 3CwaMI, 3CscMG, 1CsbMI, and 5CwaMM). Cluster 3 grouped arid and semi-arid climates at levels 1, 3, and 5 associated with medium-intermediate population strata (1BskMM, 5BwkMM, 5BwkMI, 5BskMM, 5BshMM, 5BwhMM, 3BshM, and 3BwhMI), as well as a desert megacity (1BwkMC). Cluster 4 grouped tropical climates at elevations 1 and 3 with medium-to-intermediate population strata (1AmMI, 3AmMM). Finally, cluster 5 grouped a main subgroup of arid, temperate, and tropical climates at elevations 1, 3, and 5 corresponding to medium, large, and intermediate population strata (3AfMM, 5CwaMI, 1AfMM, 1AmMM,1AwMM,5CfbMM,1CfaMM,3CwaMM,5AsMM, 5CfbMG,5CwcMM,1BshMM,1BshMI,5CwbMM,3AwMM and 5AwMM).

DVcs- Geographic Location

DVcs Geographic Location initially grouped eight clusters into dendrograms; however, based on homogeneity criteria, three main clusters were selected from within the dendrograms (Figure. 4e). Cluster 1 grouped arid, temperate, and subtropical climates located in ZI, ZTS, and ZTN (BskGCS-ZI, CwcGCS-ZI, CsbGCS-ZTS, BwkGCS-ZTS, CwaGCS-ZTS, BwkGCS-ZI, BskGCS-ZTN, BwhGCS-ZTN, and BwkGCS-ZTN. Cluster 2 grouped dry temperate, arid, and semi-arid climates in ZI, ZTS, and ZTN (CscGCS-ZTS, BwhGCS-ZI, BshGCS-ZTN, and CwbGCS-ZI). Cluster 3 mainly grouped tropical and subtropical climates ranging from ZI to ZTS (CfaGCS-ZTS, AsGCS-ZI, AfGCS-ZI, AfGCS-ZTS, AwGCS-ZI, AmGCS-ZI, and CwaGCS-ZI), as well as temperate and semi-arid climates ranging from ZTS to ZI (CfbGCS-ZTS, CfbGCS-ZI, and BshGCS-ZI).

DVcs- Geographic LocationP

Based on homogeneity criteria, three clusters were selected (Figure. 4f). Cluster 1 grouped arid, semi-arid, dry temperate, and subtropical climates at elevations 1, 3, and 5, distributed from the ZTS to the ZTN (3BwkGCS-ZTS, 1BshGCS-ZI,1BSkGCS-ZTN,5BwkGCS-ZTN,5BwkGCS-ZI,3BshGCS-ZI,3BwhGCS-ZTN,1CsbGCSZTS,3CscGCS-ZTS,1CfbGCS-ZTS and 3CwaGCS-ZTS). Cluster2 grouped tropical, subtropical, temperate, arid, and semi-arid climates at elevations 1,3and 5, mainly in the ZI (3AmGCS-ZI,5CwcGCS-ZI,5CwbGCS-ZI,5BskGCS-ZI,5BshGCS-ZI, 5AsGCSZI,5CfbGCSZI,5AwGCSZI,1BshGCSZI,3AwGCS-ZI, 5CwaGCS-ZI), 1AmGCS-ZI, 1AwGCS-ZI), as well as a warm desert subgroup at elevation 5 in the ZTN (5BwhGCS-ZTN). Finally cluster 3 grouped tropical, subtropical and temperate Mediterranean climates at elevations 1,3, and 5 from ZTS to ZI (3CfaGCS-ZTS, 3AfGCSZI,1AfGCSZI,1AfGCSZTS,3CwaGCSZI,1CfaGCS-ZTS,and5CfbGCS-ZTS).



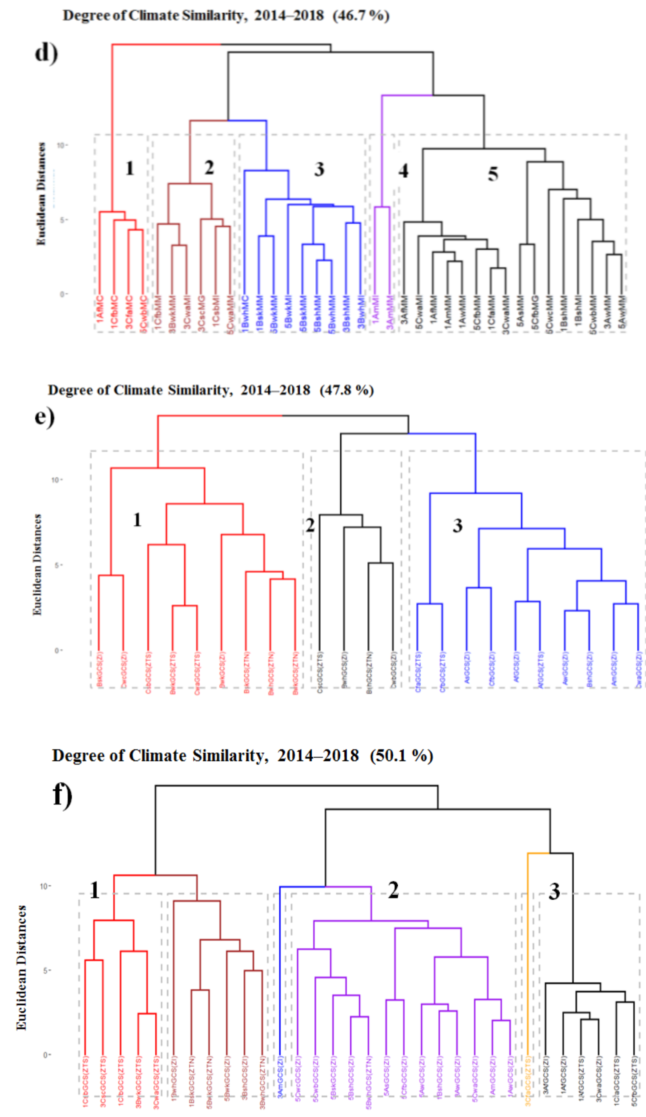


Figure. 4. Köppen-Geiger clusters based on climatic similarity: a) DVcs-Climate Zone; b) DVcs-Climate Zone P; c) DVcs-Population Stratification; d) DVcs-Population Stratification P; e) DVcs-Geographic Location; f) DVcs-Geographic Location P.

3.3 Cluster Validation with MD

Significant indicators were evaluated using the Student's t-test to assess the difference in means between indicators. In general, all indicators were statistically significant for the difference in means, with the exception of energy intensity (EI). The data for Fisher's Linear Functions (LD) and the discriminant indicators (D) for each VDcs are detailed in Figures. 5a, b, c, d, e, f.

DVcs- Climate Zone: In LD1, the values accounted for 87.4% of the total variability associated with the first eigen value. In the predicted data, 300 observations (98.3%) were correctly classified into the first cluster, and 37 observations (82.4%) into the second cluster. The model had a hit probability of 96.58% (overall error of 3.42%). With 70% of the test data, the model yielded a hit probability of 96.32% (overall error of 3.68%). The main discriminants (D), according to Fischer's Linear Functions (LD) shown in Fig. 5a, corresponded in LD1 to socioeconomic and sociopolitical indicators, such as the Human Development Index (HDI) and the Rule of Law Index (WJP), which explain 87% of the data in cluster 2, characterized by tropical and subtropical climates (Af, Cfa, As, Am, Aw, Cwa), as well as two subgroups: one temperate (Cfb, Cwc, Cwb) and another semi-arid (Bsh, Bsk). In the data validation (30%), eighty-nine observations (96.7%) were correctly classified in the first cluster, and eleven observations (78.5%) in the second, yielding a probability of correct classification of 94.34% (overall error of 5.66%). The analyzed group and the D data were as follows:

*(Group 1: T + UHI + P + H + Ai + UA + Pd + UG + Cf + PM25 + PM10 + E + GDP + **HDI** + GINI + DI + WR + **WJP,** data = todos).*

a)	
## Coefficients of linear discriminants:	
## LD1	
## Tm -2.556964e-01	## E 4.841893e-06
## UHI -5.589141e-01	## GDP -3.784208e-05
## Pp 5.534892e-03	**## HDI 3.513901e+00**
## H -3.594557e-02	## GINI -1.337285e-02
## AI -2.112587e-01	## DI -5.803131e-01
##UA -3.225837e-04	## WR 1.156023e-02
## Dp 8.216582e-05	**## WJP 9.420736e+00**
## UG -1.133921e-01	
## Cf 4.464681e-01	
## PM25 -6.593838e-03	**## LD1**
## PM10 2.330291e-02	**0.874**

Figure. 5a. Coefficients of Linear discriminants DVcs-Zona Climática

DVcs-Zona ClimáticaP: LD1 explained 68.82% and LD2 explained 31.11% of the total variability (100%) associated with the first eigenvalue. In the predicted data, fifty-nine observations (62.1%) were correctly classified in the first cluster, one hundred sixty observations (92%) in the second cluster, and sixty-seven observations (85%) in cluster 3. The model had a hit probability of 82.29% (overall error of 17.71%). At the 70% threshold, a hit probability of 81.68% was obtained (overall error of 18.03%). The main Ds (Fig. 5b) identified were: LD1, socioeconomic indicators, such as the Human Development Index (HDI), which explains 64% of the data in cluster 3, characterized by arid–desert climates of elevations 1, 3, and 5 (1Bsk, 1Bwh, 3Bwk, 3Bsh, 3Bwh, and 5Bwk), temperate climates at elevations 1 and 3 (1Cfb, 1Csb, and 3Csc), and subtropical climates (3Cfa). LD2 corresponded to multidimensional poverty (MPI), with a 36% contribution in cluster 1, characterized by tropical-subtropical climates at elevation 1 (1Af, 1Am, 1Aw, 1Cfa), elevations 3 (3Af, 3Cwa), tropical-subtropical climates at elevation 5 (5As, 5CWa), and temperate climates (5Cfb). In the validation set (30%), nineteen observations (65.5%) were correctly classified in the first cluster, forty-five (84.9%) in the second, and twenty-one (87.5%) in the third. The model yielded a probability of correct classification of 80.19% (overall error of 19.81%). The analyzed group and the D values were as follows:

*(Group 2: T + UHI + P + H + AI + UA + Pd + UG + Cf + PM25 + PM10 + E + GDP + **HDI** + GINI + DI + **MPI** + CRI + WR + WJP, data = todos).*

b)		
## Coefficients of linear discriminants:		
## LD1 LD2		**## LD1 LD2**
## Tm -1.963735e-01 1.266355e-01		## E 2.524385e-04 -2.494369e-04
## UHI -3.352328e-01 7.038552e-02		## GDP 1.272773e-05 1.852154e-05
## Pp 2.746760e-03 -5.962579e-03		**## HDI 3.735223e**+00 -3.817467e-01
## H -1.132694e-02 1.782206e-02		## GINI 2.189446e-02 -1.376801e-02
## AI -1.378548e-01 2.119149e-01		## DI -5.474027e-02 4.148062e-01
## UA -3.887148e-04 -6.873801e-05		**## MPI -1.827996e+01** 2.774986e+01
## Dp -7.301534e-05 -2.415177e-05		## IRC -1.305665e-04 2.303140e-03
## UG 4.104605e-01 3.650446e-01		## WR 2.096249e-01 2.393951e-02
## Cf -6.322513e-01 8.746610e-01		## WJP -3.195253e+00 -1.221432e+01
## PM25 -1.556801e-02 3.075582e-02		**## LD1 LD2**
## PM10 2.968480e-02 1.761032e-05		**0.6441 0.3559**

Figure. 5b. Coefficients of Linear discriminants DVcs-Climate Zone P

DVcs- Population Stratification: LD1 explained 64.6%, LD2 19.4%, LD3 11%, and LD4 4% of the total variability (100%). In the first cluster, twenty observations (100%) were correctly classified; in the second, two hundred six (92%); in the third, sixty-seven (78.8%); in the fourth, nine (90%); and in the fifth, twenty observations (100%). The model had a hit probability of 92.0% (overall error of 8.0%). At a 70% test set, a hit probability of 93.45% was obtained (overall error of 6.55%). The main Ds were (Fig. 5c): LD1, the socioeconomic indicator Cf, which explains 65% of the data in cluster 5, characterized by tropical, subtropical, and temperate climates (AfMM, AwMM, AmMM, AsMM, CwaMI, CwaMM, CfaMM, CfbMM, CfbMG BshMI). LD2 explained 19% of the data using socio-political discriminants associated with the MPI in cluster 4, characterized by tropical and temperate climates in the middle-to-upper strata (AmMI, CwcMM), as well as HDI-WJP socioeconomic-sociopolitical discriminants in cluster 5 with tropical, subtropical, and temperate climates (AfMM, AwMM, AmMM, AsMM, CwaMI, CwaMM, CfaMM, CfbMM, CfbMG, BshMI). LD3 explained 11% of the data using the socio-political discriminant (DI) in cluster 5, associated with tropical, subtropical, and temperate climates of the middle-intermediate stratum (AfMM, AwMM, AmMM, AsMM, CwaMI, Cwa, CfaMM, CfbMM, CfbMG, BshMI). LD4 explained 5% of the data using the MPI sociopolitical discriminant in cluster 4, characterized by intermediate-mid-stratum tropical and temperate climates (AmMI, CwcMM), and the discriminant D in cluster 5, associated with mid-stratum tropical-subtropical climates (AfMM, AwMM, AmMM, AsMM, CwaMI, CwaMM, CfaMM, CfbMM, CfbMG, BshMI). At the 30% level, six observations were correctly classified in cluster 1 (100%), fifty-nine in the second, twenty-three in the third (88.46%), three in the fourth (100%), and six in the fifth (100%). The model had a probability of correct classification of 91.5% (overall error of 8.49%). The analyzed group and the Ds were as follows:

*(Group 3: T + UHI + P + H + Ai + UA + Pd + UG + Pob + **Cf** + CO2U + PM25 + PM10 + E + GDP + Ec + **HDI** + GINI + **DI** + **MPI** + CRI + WR + **WJP**, data = EDT).*

c)				
## Coefficients of linear discriminants:				
##	**LD1**	**LD2**	**LD3**	**LD4**
## Tm	-1.031481e-01	-2.511564e-01	1.893354e-01	-3.025492e-02
## UHI	-1.094328e-01	-5.312495e-02	-5.726338e-01	2.820398e-01
## Pp	9.379955e-04	2.420910e-03	-3.010313e-03	4.466051e-03
## H	-1.314303e-02	-2.213944e-02	-2.821690e-02	2.185923e-02
## AI	-3.493348e-02	-1.103735e-01	9.312495e-02	-1.631734e-01
## AU	-6.476825e-04	-4.736709e-05	2.835232e-04	5.229903e-04
## Pd	2.341185e-05	-8.455955e-06	-3.279583e-05	3.133422e-05
## UG	6.725911e-02	4.074391e-01	1.358000e-01	2.970179e-01
## Pob	-1.153870e-07	1.691227e-07	-9.353967e-08	2.799051e-07
## Cf	**4.826588e-01**	-9.929316e-02	-2.124822e+00	-7.007743e-01
## CO2U	1.964870e-09	-5.272576e-08	3.069093e-09	-2.090217e-08
## PM25	6.060888e-02	-1.589199e-02	1.716485e-02	-2.030105e-02
## PM10	-1.100215e-02	5.410402e-02	-8.088687e-03	1.904079e-02
## E	-1.718553e-04	1.312024e-04	5.083545e-04	2.984005e-04
## GDP	9.684098e-05	-4.696003e-05	5.930511e-05	1.084296e-04
## PIBU	-3.209622e-11	5.946692e-12	7.267208e-12	-1.756860e-11
## DHI	1.911195e-01	**5.664848e+00**	-6.601283e+00	-1.891247e-01
## GINI	1.912382e-02	3.210715e-02	1.116507e-02	2.888097e-02
## DI	-9.818861e-02	-2.370415e-01	**1.039366e+00**	**1.139351e+00**
## MPI	-2.330643e+01	**9.001050e+00**	1.075668e+01	**1.890509e+01**
## CRI	-1.742389e-03	-5.339699e-03	-6.420046e-03	1.958136e-02
## WR	5.995910e-02	6.367696e-02	-1.233945e-01	-1.316368e-01
## WJP	-4.754784e+00	**3.854008e+00**	-1.565761e+01	-4.375040e+00
##Proportion of trace:				
##	**LD1**	**LD2**	**LD3**	**LD4**
	0.6546	0.1903	0.1096	0.0456

Figure.5c. Coefficients of Linear Discrimination (DVcs) – Population Stratification.

DVcs-Population StratificationP: LD1 explained 60.1%, LD2 23.1%, LD3 12%, and LD4 0.4% of the total variability (100%) associated with the first eigenvalue. In the predicted data, thirty-five observations (89.7%) were correctly classified in the first cluster, twenty (100%) in the second, fifty-five observations (83.2%) in the third, thirty observations (75%) in the fourth, and one hundred fifty-nine (91.3%) in the fifth. The model had a prediction accuracy of 85.43% (overall error of 14.57%). At 70%, the prediction accuracy was 88.12% (overall error of 11.88%). The main Ds identified (Fig. 5d) were: LD1, the sociopolitical WJP, which explains 53.95% of the data in cluster 4, characterized by tropical climates at elevations 1 and 3 with medium-intermediate population sizes (1AmMI, 3AmMM), and the MPI in cluster 3, arid and desert climates at elevations 1, 3, and 5 with intermediate, medium, and megacity population strata (1BskMM, 5BwkMM, 5BwkMI, 5BskMM, 5BshMM, 5BwhMM, 3BshMM, 3BwhMI, 1BwkMC). LD2 explained 25% of the data through the socioeconomic variable HDI in cluster 4, characterized by tropical climates at elevations 1 and 3 with an intermediate-to-medium population stratum (1AmMI, 3AmMM). LD3 explained 14% of the data using the socioeconomic-sociopolitical HDI-WJP in cluster 4, associated with tropical climates at elevations 1.3 and an intermediate-middle socioeconomic stratum (1AmMI, 3AmMM). LD4 explained 7% of the data using the HDI in cluster 4, characterized by tropical climates at elevations 1.3 with an intermediate-to-medium population stratum (1AmMI, 3AmMM). At 30%, eleven observations (91.6%) were correctly classified in the first cluster, six (100%) in the second, nineteen (90.47%) in the third, eight observations (66.6%) in the fourth, and fifty-two observations (94.54%) in the fifth. The model had a probability of correct classification of 90.57% (overall error of 9.43%). The analyzed group and the D values were:

*(Group 4: T + UHI + P + H + Ai + UA + Pd + UG + Pob + Cf + CO2U + PM25 + PM10 + E + GDP + Ec + **HDI** + GINI + Di + **MPI** + CRI + WR + **WJP**, data = EDT).*

d)				
## Coefficients of linear discriminants:				
##	LD1	LD2	LD3	LD4
## T	9.825061e-02	-5.980919e-02	6.065041e-02	-5.561179e-02
## UHI	2.552795e-01	1.567675e-01	-3.198137e-01	-1.192029e-01
## P	-5.944077e-04	-1.155341e-04	-1.003070e-03	2.487387e-03
## H	9.861419e-04	3.760102e-02	-5.139314e-02	-5.404051e-03
## Ai	1.630903e-02	2.997191e-03	1.876556e-03	-8.736291e-02
## UA	1.277564e-03	-7.113342e-04	-5.686717e-04	4.902985e-04
## Pd	-6.764504e-05	-4.432701e-05	-5.193006e-05	6.818490e-05
## UG	-2.869688e-01	-1.984622e-01	2.853995e-01	-8.599047e-02
## Pob	1.497875e-07	3.092173e-08	2.517365e-07	3.971040e-07
## Cf	-1.938650e+00	-1.939484e+00	-5.535766e-03	-1.419028e+00
## CO2U	-5.502999e-08	2.744848e-08	-2.671824e-08	1.204599e-08
## PM25	-2.239285e-02	-2.640296e-02	-3.888357e-02	1.712715e-02
## PM10	4.899685e-03	4.951951e-03	4.277051e-02	-1.337017e-02
## E	5.498675e-04	5.262279e-04	1.844643e-04	3.852822e-04
## GDP	-6.567096e-05	-6.408081e-06	-1.038503e-06	7.349103e-05
## PIB	3.206638e-11	-7.403884e-12	-3.609347e-12	-3.239431e-11
## HDI	-2.391473e+00	**7.272044e+00**	**2.668520e+00**	**1.702363e+00**
## GINI	-3.327577e-03	-1.182164e-02	-1.637774e-02	8.766614e-02
## DI	4.015723e-01	-7.359410e-01	-1.092568e-01	1.999380e-01
## MPI	**3.216869e+01**	1.567768e+01	-3.494943e+00	-6.700180e+01
## CRI	-7.492727e-04	3.937064e-03	-4.204126e-03	-9.618873e-03
## WR	-1.141543e-01	-1.512331e-02	7.670284e-02	-7.048551e-02
## WJP	**4.171974e+00**	2.517836e+01	**1.911230e+00**	-7.331396e+00
## Proportion of trace:				
##	**LD1**	**LD2**	**LD3**	**LD4**
##	0.5395	0.2462	0.1424	0.0719

Figure.5d..Coefficients of Linear discriminants DVcs-estratificación poblacionalP.

DVcs-Geographic Location: The results were: LD1 accounted for 68.34% and LD2 for 31.66% of the total variability (100%). In the predicted data, forty-three observations (46%) were correctly classified in the first cluster, forty-three observations (86%) in the second, and two hundred forty observations (96%) in the third. The model had a hit probability of 87.43% (overall error of 12.57%). At 70%, a hit probability of 86.94% was obtained (overall error of 13.06%). The main Ds in LD (Fig. 5e) were as follows: LD1 corresponded to the socio-political DI, which explains 66% of the data in cluster 1, characterized by arid, desert, and subtropical climates from ZI to ZTN (BskGCS-ZI, CwcGCS-ZI, CsbGCS-ZTS, BwkGCS-ZTS, CwaGCS-ZTS, BwkGCS-ZI, BskGCS-ZTN,BwhGCS-ZTN,BwkGCS-ZTN). LD2 corresponded to MPI (33% of the data in cluster 3), in tropical, subtropical, temperate, and arid climates from ZI to ZTS (CfaGCS-ZTS, CfbGCS-ZTS, AsGCS-ZI, CfbGCS-ZI, AfGCS-ZI, AfGCS-ZTS, AwGCS-ZI, BshGCS-ZI, AmGCS-ZI and Cwa GCS-ZI). The WJP model explained 33% of the data (cluster 2) in temperate Mediterranean and arid-desert climates from ZTN to ZTS (CscGCS-ZTS, BwhGCS-ZI, BshGCS-ZTN, CwbGCS-ZI). In the first cluster, 30% of the data were correctly classified (33.3% of observations); in the second, 12% (80%); and in the third, 69% (92%). The model had a probability of correct classification of 81.91% (overall error of 18.09%). The analyzed group and the D values were as follows:

*(Group 5: T + UHI + P + H + Ai + UA + Pd + UG + Pob + CO2U + E + GDP + Ec + HDI + GINI + **DI** + **MPI** + WR + **WJP**, data = EDT).*

e)	
## Coefficients of linear discriminants:	
## LD1 LD2	## LD1 LD2
## T 3.249188e-02 1.440877e-0	## GINI 9.155145e-02 -2.267800e-02
## UHI 1.651492e-01 -5.818917e	**## DI 6.536410e-01** -8.003617e-01
## P -3.115379e-04 -5.254654e-0	**## MPI** -2.784947e+01 **1.201859e+00**
## H 5.670070e-02 2.110833e-(	## WR 1.357036e-01 -5.686446e-02
## Ai 2.185356e-02 1.926973e-(	**## WJP** -1.398587e+01 **1.005793e+01**
## UA 6.587021e-04 4.414466e-	**## Proportion of trace:**
## Pd -2.104734e-06 5.662389e-	**## LD1 LD2**
## UG 6.896203e-02 7.492398e-	**## 0.66 0.33**
## Pob -1.842953e-07 -4.011257e-07	
## Cf -3.817282e-01 -1.886603e-01	
## CO2U 3.278228e-08 1.618322e-07	
## E -8.053558e-05 -1.329463e-04	
## GDP -2.187200e-05 7.592688e-05	
## Ec -6.702968e-13 -3.789834e-12	
## HDI -3.210427e+00 9.113197e-02	

Figure 5e. Coefficients of Linear discriminants DVcs- Geographic Location.

VDsc- Geographic LocationP: LD1 explained 65.14% of the total variability, and LD2 explained 34.86% (100%). In the predicted data, 176 observations (88%) were correctly classified in the first cluster, 59 observations (59.78%) in the second, and 60 observations (80%) in the third. The model has a probability of correct classification of 84.29% (overall error of 15.71%). At the 70% threshold, a hit probability of 86.94% was obtained (overall error of 15.57%). The Ds (Fig. 5f) were: LD1, based on urban-bioclimatic (UHI), socioeconomic (Cf), and sociopolitical (MPI and WJP) indicators, accounting for 66% of the data. In cluster 1, the Ds were UHI-MPI in arid, semi-arid, and temperate climates at elevations 1, 3, and 5 from ZTN-ZTS (3BwkGCSZTS, 1BshGCSZI, 1BSkGCSZTN, 5BwkGTN, 5BwkGZI, 3BshGCSZI, 3BwhGCSZTN, 1CsbGCSZTS, 3scGCSZTS, 1CfbGCSZTS, 3CwaGCS-ZTS). In cluster 2, Cf-WJP included tropical, subtropical, arid, and semi-arid climates at elevations 1, 3, 5, mainly in ZI (3AmGCSZI, 5CwcGCSZI,5CwbGCSZI,5BskGCSZI,5BsCSZI,5AsGCS ZI,5CfbGCSZI,5AwGCSZI,1BshGCSZI,3AwGCSZI,5Cwa GCS-ZI,1AmGCS-ZI,1AwGCSZI,and5BwhGCSZTN).LD2 corresponded the urban-bioclimatic UHI, which explained 33% of the data (cluster 1), associated with arid, semi-arid, dry temperate, and subtropical climates ranging from ZTN toZTS(3BwkGCSZTS,1BshGCSZI,1BSkGCSZTN,5BwkG CSZN,5BwkGCSZI,3BshGCSZI,3BwhGCSZTN,1CsbGCS ZT, 3CscGCSZTS, 1CfbGCS-ZTS, and 3CwaGCS-ZTS). At 30%, fifty-three observations (83.3%) were correctly classified in the first cluster, sixteen (69.56%) in the second, and eighteen in the third. The model has a hit probability of 82.08% (overall error of 17.92%). The analyzed group and the Ds were as follows:

*(Group.6 ~ T + **UHI** + P + H + Ai + UA + Pd + UG + Pob + **Cf** + CO2U + PM25 + PM10 + E + GDP + Ec + HDI + GINI + DI + **MPI** + CRI + WR + **WJP**, data = EDT).*

f)	
Coefficients of linear discriminants:	
## LD1 LD2	## LD1 LD2
## T 2.458773e-01 -1.575990e-02	## Ec 1.219103e-11 2.406801e-11
## UHI 2.944345e-01 2.42644(	## HDI -5.124125e+00 -1.863713e+00
## P -5.595659e-03 -5.203170e-0	## GINI -9.759080e-03 4.739755e-02
## H 5.017177e-03 8.038625e-(	## DI -4.968645e-02 5.773422e-01
## AI 2.350867e-01 6.602790e-(	**## MPI 1.552463e+01** -2.866513e+01
## UA 1.701694e-03 -9.420953e-	## CRI 4.389778e-03 -1.441390e-03
## Pd 8.920549e-05 2.987594e-(	## WR -1.540690e-01 2.134234e-01
## UG 1.088632e-01 1.711240e-	**## WJP 2.838096e+00** -1.820523e+01
## Pob -2.609060e-07 -1.604487e-	**## Proportion of trace:**
## Cf 2.064441e+00 -6.675510e-	**## LD1 LD2**
## CO2U -2.134826e-08 -4.56579	## 0.6702 0.3298
## PM25 -4.208021e-02 2.513465e-02	
## PM10 3.468920e-02 -2.349374e-02	
## E -6.626427e-04 1.383731e-04	
## GDP -4.628690e-05 -9.146982e-06	

Figure .5f. Coefficients of Linear discriminants DVcs- Geographic LocationP

3.4. Classification errors and D in groups

A key step was to verify compliance with the statistical assumptions of normality and equal variances. To this end, the Shapiro-Wilk test (normality) and Levene's test (equal variances) were applied. However, these statistical assumptions were not met, a situation that frequently occurs when there is no experimental control over the indicators being analyzed. To address this limitation, the Central Limit Theorem (CLT) was applied, which states that as the sample size increases (in this case, 350 observations), the probability distribution of the test statistic tends to approximate a normal distribution. This criterion allows for the approximation of multiple commonly used distributions, such as the binomial, Poisson, chi-square, Student's t, and gamma distributions, among others, when their parameters increase and the analytical calculation becomes more complex (Devore, 2001). Overall analyses revealed varying classification errors across the groups (Tables 2,3). The most reliable results are DVsc-Climate Zone (5.56%) and DVsc-Population Stratification

(8.49%), and the least reliable are DVsc-Geographic Location (18.9%), DVsc-Geographic LocationP (17.92%), and DVsc-Population StratificationP (9.4%).The generally recognized features were sociopolitical (MPI, WJP, DI) and socioeconomic (HDI,Cf). In other cases, the Köppen-Geiger bioclimatic patterns (T,H,P) did not constitute features within the clusters. In (DVsc-Climate Zone, lowest classification error), 87% of the data (cluster 2) corresponds to human development and the rule of law (HDI and WJP) in tropical, subtropical, temperate, and arid environments. In DVsc-Climate Zone, the Koppen-Geiger similarity percentage (61%) contributed to the lower classification error; however, this degree of similarity was not sufficient for the bioclimatic patterns to be classified as *D*, nor was the climate-related socio-political index (CRI). This finding regarding the CRI differs from studies, reports, and maps (United Nations, 2016; Germanwatch,2021) that link heterogeneity, population density, poverty, and inequality in metropolitan areas to climate change risks and to urban vulnerability in hotspots across the Caribbean and Central America. However, it is advisable to continue monitoring the CRI, given global changes in vulnerable geographies (Pugh et al., 2016) and the differential susceptibility of certain regions (world) to variations in precipitation (Giorgi, 2006).

In DVsc-Population Stratification, Cf, as a critical indicator (IPCC, 2021), explained a significant percentage of the data (65%) but was associated with cluster 5 (tropical and subtropical climates with medium-to-large population strata, temperate climates with medium-to-large strata, and hot arid climates with intermediate strata). Other socioeconomic and sociopolitical variables, such as human development, rule of law, and multidimensional poverty (HDI-MPI-WJP), contributed only 19% of the information in cluster 4, in tropical and temperate climates with medium-to-large populations (AmMI and CwcMM). In this regard, the influence of the MPI is unprecedented, as it not only affects the quality of life of the regional metropolitan population but also influences the climatic and socio-environmental dynamics of tropical monsoon metropolises (AmMI), such as Managua (high temperatures 27–28 °C), and cold subtropical metropolises at high altitudes (3,600 meters above sea level, such as La Paz (Cwc). In another case, DVsc-GeographicLocationP, 67% of the data (clusters 1 and 2) identified urban-bioclimatic, socioeconomic, and sociopolitical factors associated with UHI, Cf, MPI, and WJP. The results (Table 3) show that UHI is strongly related to WJP and MPI in arid, desert, subtropical, tropical, and temperate environments at elevations 1, 3, and 5, from ZTN to ZTS. This result is expected regarding the relationship between UHI and Cf; however, the influence of WJP on heat islands is unprecedented, as previous studies mainly associate UHI with social vulnerability (Klein-Rosenthal et al., 2014; Harlan et al., 2007). Nevertheless, this result should be interpreted as a relative approximation of reality, due to the high classification error (17.92%) of DVsc-GeographicLocationP.

.

Table.2. DVcs, LD, Classification error and climatic similarity.

VDsc	LD *D*	E (%)	S (%)
VDcs Climate Zone	LD1 (87%) (IDH-WJP)	**5.66**	**61**
VDcs Climate Zone P	LD1 (64%) (HDI) LD2 (36%) (IPM)	19.81	51.5
VDcs - Population stratification	LD1 (65%) (Cf) LD2 (19%)(DIH,MPI,WJP) LD3 (11%) (DI) LD4 (5%) (DI-MPI)	**8.49**	**50.1**
VDcs - Population stratificationP	LD1 (54%) (MPI-WJP) LD2 (25%) (HDI) LD3 (14%) (HDI-WJP) LD4 (7%) (HDI)	9.43	46.7
VDcs - Geographic Location	LD1 (67%) (DI) LD2 (33%) (MPI-WJP)	18.9	47.8
VDcs - Geographic LocationP	LD1 (67%) (Cf-UHI-MPI-WJP) LD2 (33%) (UHI)	17.92	50.1

(a) The VDcs with the lowest classification error are highlighted in gray**.** Where: ***D.*** discriminant **E**, Error, **S**: Similarity.

Table 3. Discriminants (D) in Koopen-Geiger clusters in the VDsc with the lowest classification error. (highlighted in gray).

VDsc	FEATURES OF KOOPEN GEIGER GROUPS
VDcs Climate Zone (5.66%) LD1 (87%) (HDI-WJP)	**A 61% climate similarity** across temperate, arid, semi-arid, tropical, and subtropical climates. The D-scores for human development and the rule of law **characterize 87% of the metropolises.**
	Tropical and subtropical (Af, Cfa, As, Am, Aw, Cwa), temperate (Cfb, Cwc, Cwb), semi-arid (Bsh, Bsk).
VDcs - Population stratification (8.49%) LD1 (65%) (Cf)	**A 50.1% climate similarity** in tropical, subtropical, and arid climates across medium-to-intermediate population size and megacities. Socioeconomic indicators such as carbon (Hcpc) influence climate-socio-environmental dynamics **(62% of metropolises).**
	Tropical-subtropical(AfMM, AwMM, AmMM, AsMM, CwaMI,CwaMM,CfaMM)Temperate(CfbMM,CfbMG), Arid (BshMI).
LD2 (19/%) (MPI)	A small group comprising 3% of tropical and temperate regions, characterized by sociopolitical factors such as multidimensional poverty (MPI); however, this dimension accounts for only 19% of the data.
	Tropical (AmMI), Temperate (CwcMM).
LD2 (19/%) (HDI-WJP)	A diverse group (63%) tropical, subtropical, temperate, and semi-arid cities of medium, intermediate, and large size, characterized by D socioeconomic and sociopolitical (HDI,WJP).
	Tropical-subtropical (AfMM, AwMM, AmMM, AsMM, CwaMI, CwaMM, CfaMM), Temperate (CfbMM, CfbMG), Hot semi-arid (BshMI).
LD3 (11%) (DI)	A group comprising 63% of tropical, subtropical, temperate, and semi-arid metropolises characterized by the sociopolitical variable democracy (ID), which characterizes the same groups as LD2 (HDI-WJP), but with a share of only 11% of the data.
	Tropical-subtropical (AfMM, AwMM, AmMM, AsMM, CwaMI, CwaMM, CfaMM); Temperate (CfbMM, CfbMG); Hot semi-arid (BshMI).

(a) The VDcs climate zone (5.66%) successfully groups climates with high similarity (61%) and thus allows for the formation of a more climatically homogeneous cluster (87% of metropolitan areas).

4. Conclusions

The Köppen-Geiger meteorological characteristics (T, P,H) did not constitute metropolitan districts in Latin America and the Caribbean (LAC) based on climatic similarity criteria. On the other hand, the carbon footprint indicator (Cf) was classified as D based on the population stratum and in linear combination with the ICU and the WJP. This result suggests that the carbon impact in metropolitan areas has a limited scope depending on the population's sociopolitical-climatic dynamics and due to hotspots affecting metropolitan environments with social segregation. In DVsc-PopulationStratification, it was verified that social gaps (MPI) are related to human development (HDI), democracy (DI), and the rule of law (WJP) in a small group (3%) of climates with extreme monsoon and cold subtropical temperatures or geographies (altitude).

The Köppen-Geiger similarity index (61%) in the DVsc Climate Zone, along with the lower classification error (5.66%), made it possible to reliably group 87% of regional metropolises based on socioeconomic and sociopolitical factors (HDI-WJP) into tropical, subtropical, and semi-arid climates. This finding implies that human development and the political rights of the population influence the climatic and socio-environmental morphology of regional metropolises (see Table 3). Furthermore, the Köppen-Geiger patterns present in this cluster (87%) do not alter the D status of the aforementioned urban socioeconomic and sociopolitical factors. In this regard, governance and socioeconomic conditions are better able to identify urban characteristics than meteorological indicators traditionally used in regional climate studies.

Finally, population density (Pd), metropolitan area size (UA), energy (E), and climate risk (CRI) did not emerge as significant factors in any of the multivariate analyses conducted. These results allow us to rule out these indicators, which are considered critical in the scientific literature (IPCC, 2021), but which do not provide sufficient evidence to characterize metropolitan groups under the methodological approach of climate similarity.

References

Baumgartner, W. H. (2025). Gentrificación verde en América Latina. Universalidades, especificidades y particularidades en Brasil. Revista de Geografía Norte Grande, (93), 1-17. https://doi.org/10.4067/S0718-34022026000100110.

Beck, H., Zimmermann, N., McVicar, T.R., Vergopolan, N.,Berg,A., Wood,E.F.(2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, (5), *180-214.* https://doi.org/10.1038/sdata2018.214.

Boccolini, S. M. (2025). Metrópolis más que urbanas: reconceptualizando los territorios de soporte de la vida urbana. Revista de Investigaciones Geográficas: una mirada desde el Sur, (69), 3-23. https://doi.org/10.5354/0719-5370.2025.76776.

Casadei, P., Semmartin, M. y Garbulsky, M.F. (2021). Análisis regional de las islas de calor urbano en la Argentina. *Ecología Austral*, *31*(1), 190–203. https://doi.org/10.25260/EA.21.31.1.0.970.

Comisión Económica Para La America Latina (CEPAL) (2017). Observatorio Demográfico de América Latina y el Caribe 2017: Tablas de mortalidad = Demographic Observatory of Latin America and the Caribbean 2017: Life tables. https://www.cepal.org/es/publicaciones/42361-observatorio-demografico-america-latina-caribe-2017-tablas-mortalidad.

Cox, W. (2025). *Demographia World Urban Areas: 2025 Edition*.Demographia/ChapmanUniversity.https://blogs.chapman.edu/wpcontent/uploads/sites/56/2025/06/Demographia-World-Urban-2025.pdf.

Devore, J. (2001). *Probabilidad y Estadística para Ingeniería y Ciencias.* Séptima edición. CengageLearning.https://intranetua.uantof.cl/facultades/csbasicas/matematicas/academicos/jreyes/DOCENCIA/APUNTES/APUNTES%20PDF/Probabilidad%20y%20Estadistica%20para%20Ingenieria%20y%20Ciencias%20%20Jay%20Devore%20%20Septima%20Edicion.pdf

Dobbs, C., Nitschke CR, y Kendal D. (2014). Global Drivers and Tradeoffs of Three Urban Vegetation Ecosystem Services. *PLoS ONE*, (9), 1-9. https://doi:10. 1371/journal.pone.0113000.

Economist Intelligence Unit (2019). *Democracy index 2019*: *A year of democratic setbacks and popular protest*.https://www.eiu.com/n/campaigns/democracy-index-2019/.

Economist Intelligence Unit. (2015). *Democracy index 2015: Democracy in an age of anxiety*.https://www.eiu.com/public/topical_report.aspx?campaignid=DemocracyIndex201

Germanwatch. (2021). *Global Climate Risk Index 2021: Who suffers most from extreme weather events? Climate-related loss events in 2019 and 2000 to 2019*.https://www.germanwatch.org/sites/default/files/Global%20Climate%20Risk%20Index%202021_2.pdf.

Giorgi F. 2006. Climate change hot-spots. Geophysical Research Letters, 33 (17), 1-4. https://doi.org/10.1029/2006GL025734.

González Á., Pacheco, C., Henao.,A., Monjardin, S.,, Plata, W., y Peña, J. (2025a). Identification of discriminating indicators and criteria for sustainability, resilience, and liveability in Latin American and. Revista Ciencia e Ingeniería. 46(1), 87102.https://www.researchgate.net/publication/391425684_Identification_of_discriminating_indicators_a nd_criteria_forsustainability_resilience_and_liveability_in_Latin_American_and_Caribbean_metropolises.

González-Calderón, Á., Pacheco-Angulo, C., y Peña-Guillén, J. (2025b). Caracterización socio-ambiental y urbana de algunos emplazamientos metropolitanos de Latinoamérica y el Caribe mediante técnicas multivariantes, 2014-2018. *Revista Geográfica Venezolana*, 66(1), 106-131. https://doi.org/10.53766/RGV/2025.66.1.07.

Grimmond, C.S.B., Souch, C. (1994). Surface description for urban climate studies: A GIS based methodology. *Geocarto International* 1, (47), 59. https://doi.org/10.1080/10106049409354439

Harlan, S., Brazel, A.J.,Jenerette, G.D., Jones, N.S., Larsen, L, Prashad, L ., Stefanov, W.L. (2007). In the shade of affluence: the inequitable distribution of the urban heat island. *Equity and the Enviro*nment, (15), 173–202. https://doi: 10.1016/S0196-1152(07)15005-5.

Henríquez, C., Mallea, C., Henríquez-Dole, L y Samaniego, H. (2017). Dispersión y Escalamiento Urbano en el Sistema de Ciudades Chileno Urban Sprawland Scaling in the Chilean Cities System. *Investig. Geogr*, (54),5-22.https://doi:10.5354/0719-5370.2017.48039.

Hiroito.,M., Rodrigues Pereira,M.M y Pereira de Freitas Neto, R. (2026). Machine Learning and Remote Sensing Applications in Urban Heat Island research: a bibliometric analysis. revista regeo, São José dos Pinhais,17,(2),124.https://doi.org/10.56238/revgeov17n2-098.

Intergovernmental Panel on Climate Change (IPCC) (2006). 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Volume 2, Energy. https://www.ipccnggip.iges.or.jp/public/2006gl/pdf/2_Volume2/V2_1_Ch1_Introduction.pdf.

Intergovernmental Panel on Climate Change. (IPCC) (2021): Summary for PolicymakersTechnical Summary Frequently Asked Questions Glossary. Part of the Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on ClimateChange.https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SummaryVolume.pdf.

Klein-Rosenthal, J., Kinney, P. L., and Metzger, K. B. (2014). Intra-urban vulnerability to heat-related mortality in New York City, 1997-2006. *Health & Place*,(*30)*,4560.https://doi:10.1016/j.healthplace.2014.07.014.

Lauwaet, D., Hooyberghs, H., Maiheu, B., Lefebvre, W., Driesen, G., Van Looy, S., De Ridder, K., (2015). Detailed Urban Heat Island Projections for Cities Worldwide: Dynamical Downscaling CMIP5 Global Climate Models. Climate (3), 391–415. https://doi.org/doi: 10.3390/cli3020391.

Moran, D., Kanemoto, K., Jiborn, M., Wood, R., Tobben, J., & Seto, K. (2018). Carbon footprints of 13 000 cities. *Environ. Res. Lett*, (13), 06404, 1-10. https://doi.org/10.1088/1748-9326/aac72a.

Pacifici, M., Rama, F y de Castro Marins, K. (2019). Analysis of temperature variability within outdoor urban spaces at multiple scales. Urban Climate, (27), 90-104. https://doi.org/10.1016/j.uclim.2018.11.003.

Pugh, T., Muller C, Deryng, J.E.D., Folberth, C., Olin, S., Schmid, E., Arneth, A. (2016). Climate analogues suggest limited potential for intensification of production on current croplands under climate change. Nature Communications, (7), 1-8. https://doi.org/10.1038/ncomms12608.

R Core Team. (2023). *R: A language and environment for statistical computing* (Versión 4.3) [Software]. R Foundation for Statistical Computing. https://www.R-project.org/.

United Nations. 2016. Habitat iii regional report Latin America and the Caribbean.Sustainable Cities with Equality. Mexico. 978-92-1-133393-0, 978-92-1-1327595.https://unhabitat.org/sites/default/files/documents/2019-05/habitatiii-regional-report-lac.pdf.

United Nations (UN). (2023). Sustainable Development Goals Report. Special Edition. United Nations.https://unstats.un.org/sdgs/report/2023/TheSustainableDevelopmentGoalsReport2023_Spanish.pdf.

United Nations Development Programme (UNDP). (2018). Human Development Indicators and Indices: 2018StatisticalUpdateTeam.https://hdr.undp.org/system/files/documents/2018humandevelopmentstatisticalupdate.pdf.

United Nations Development Programme (UNDP). (2024). The 2023/2024 Human Development Report. Breaking the gridlock. Reimagining cooperation in a polarizedworld.https://hdr.undp.org/system/files/documents/global-reportdocument/hdr202324reporten.pdf.

United Nations Development Programme and Oxford Poverty and Human Development Initiative. (UNDP). (2023). Global multidimensional poverty index 2023. Unstacking global poverty: Data for high impactaction.https://hdr.undp.org/system/files/documents/hdp-document/2023mpireporten.pdf.

United Nations Habitat Organization (UN-Habitat).(2020). Global State of Metropolises. Population Data Booklet.https://unhabitat.org/globalstateofmetropolis2020%E2%80%93populationdatabooklet.

World Bank Group. (2021). Free and open access to global developmentdata.https://data.worldbank.org/country/latin-america-and-caribbean.

World Economic Forum (WEF) (2024). *The Global Risks Report2024*.19th Edition.https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf.

World Health Organization (WHO) (2018). *Guías de calidad del aire relativas al material particulado.*: https://www3.paho.org/hq/index.php?option=com_to pics&view=rdmore&cid=9833&Itemid=40799&lang=es.

World Justice Project. (WJP). (2021). Rule of Law Index. Current and Historical Data: https://worldjusticeproject.org/ourwork/research-and-data/wjp-rule-law-index2021/current-historical-data.

World Meteorological Organization (WMO) 2021. World Meteorological Information Service.https://worldweather.wmo.int/es/home.html.

Wu, X., Wang, G., Yao, R., Wang, L., Yu, D., Gui, X. (2019). Investigating Surface Urban Heat Islands in South America Based on MODIS Data from 2003–2016. *Remote Sens*, (11), 1212, 1-16. https://doi: 10.3390/rs1110.

Received: *april 30, 2026*

Accepted: *july 10, 2026*

González-Calderón, Álvaro*. Architect. Master of Science in Renewable Natural Resource Management and Environment, CIDIAT-ULA. Dr. Candidate in Forestry and Environmental Sciences, University of The Andes.*
https://orcid.org/0000-0002-1997-7351

Pacheco-Angulo, Carlos. *Forestry Engineer University of The Andes, Ph.D Geographic Information Technologies. Madrid,Spain.Email:pachecocar@gmail.com****.***
https://orcid.org/0000-0001-8724-9287

Peña-Guillen,Jesús. *Bachelor in Mathematics, PhD Statistics, Faculty of Economics and Administrative Sciences, University of The Andes.Venezuela.Email: penaguillenjesusalberto@gmail.com.*
https://orcid.org/0000-0003-2942-7086

Bianchi, Guillermo. *Biologist, Master of Science in Applied Statistics, University of The Andes, Venezuela.Email: gbianchip@gmail.com.****https://orcid.org/0000-0002-8307-4472***

Yarbuh, Ismael. *Master of Science in Renewable Natural Resource Management and Environment, CIDIAT-ULA. PhD in Earth Sciences (Geology), CICESE, Mexico. Postdoctoral Fellow, Mexican Center for Innovation in Geothermal Energy (CEMIEGEO-CICESE), Ensenada, Baja-California.E-mail:uyarbuh@uabc.edu.mx*
https://orcid.org/0000-0002-7146-3841

González, Héctor. *Electrical Engineer, University of the Andes Venezuela Email:hectogonza@gmail.com*
https://https://orcid.org/0009-0000-9098-6552

Apendix 1. Criteria for recoding Metropolis samples according to DVcs

DVcs	CATEGORIZATION OF METROPOLIS	RECODING OF METROPOLIS
	Köppen-Geiger regional climates	Average elevation
DVcs - Climate Zone	Af, Am, Aw, AsBsh, Bsk, Bwh, Bwk, Cfb, Cfa, Csc, Csb, Cwc, Cfa, Cwa: Cwb	(KGBsh, Bsk, Cwb),
	Regional population ranges	Population segment
	500,000–999,000 (Intermediate Metropolises (MI))	(BskMI, CscMM, BwkMG, AwMC)
DVcs - Population Stratification	>1 million–4,999,000 (medium-sized cities (MM))	
	>5,000 million–9,999,000 million (MG) =>10 million inhabitants Megacities (MC)	
	Regional geographical parallels	Geographic location
DVcs Geographic Location	(ZTN) Arctic Circle (Latitude 66.33° N) – Tropic of Cancer (Latitude 23.5° N) (ZI) Tropic of Cancer GCS (Latitude 23.5° N) – Tropic of Capricorn (Latitude -23.5°) (ZTS) Capricorn-Capricorn Tropic GSC (Latitude -23.5) – Antarctic Circle (Latitude -66.33).	(CwaGCS (ZTN), AmGCS (ZI), BwhGCS (ZTS).
DVcs – Climate Zone P	Köppen-Geiger regional climates	Climate - elevation discrimination (P)
	Af, Am, Aw, AsBsh, Bsk, Bwh, Bwk, Cfb, Cfa, Csc, Csb, Cwc, Cfa, Cwa: Cwb	Includes suffixes 1, 3, and 5
	Regional population ranges	Population stratum—quota
	500,000–999,000 (Intermediate Metropolises (MI))	Includes suffixes 1, 3, and 5
DVcs - Population StratificationP	>1 million–4,999,000 (medium-sized cities (MM))	
	>5,000 million–9,999,000 million (MG) =>10 million inhabitants Megacities (MC)	
DVcs - Geographic LocationP	Regional geographical parallels	Geographic location, elevation (P)
	(ZTN) Arctic Circle (Latitude 66.33° N) – Tropic of Cancer (Latitude 23.5° N) (ZI) Tropic of Cancer GCS (Latitude 23.5° N)–Tropic of Capricorn (Latitude -23.5) (ZTS) Capricorn-Capricorn Tropic GSC (Latitude -23.5) – Antarctic Circle (Latitude -66.33).	Prefixes 1, 3, and 5 are included