United Nations Educational, Scientific and Cultural Organization
Geothermal Training Programme Under the auspices of UNESCO

GRÓ GTP, Grensásvegur 9,
IS-108 Reykjavík, Iceland

Reports 2021
Number 15

ESTIMATION OF THE GEOTHERMAL GRADIENT OF TWO AREAS IN COLOMBIA, USING GRAVITY ANOMALIES AND SUPPORT VECTOR MACHINE

J. Camilo Matiz-León
Geothermal Research Group
Servicio Geológico Colombiano – SGC
Diagonal 53 No. 34-53, Bogotá D.C.
COLOMBIA
jmatiz@sgc.gov.co; cama@grogtp.is

ABSTRACT

Based on shallow temperature surveys at 1.5 m depth in the Azufral volcano geothermal area and bottom hole temperature values from oil wells with depths between 515 m and 6273 m in the Eastern Llanos sedimentary basin, the geothermal gradient in the two study areas was estimated from the first vertical derivative of the total Bouguer gravity anomaly and through the classification algorithm of support vector machines. Taking as training samples the variations in rock density in the vertical component and as test samples the geothermal gradient values calculated by conventional methodology, a trained correlation was generated between gravimetry and the variation in temperature as a function of depth. With the intention to present clearly the units used in each area of interest for the geothermal gradient, for Azufral Volcano Geothermal Area used °C/m due to surface soundings measured at 1.5 meters. For reference in this document, the equivalence of 1°C/m equals 1,000°C/km was taken. However, these values for Azufral volcano are not real, they are used in a practical sense to demonstrate the method while waiting for more realistic values in the results. For the Llanos Orientales Sedimentary Basin due to the oil wells used with depths between 0.5 km and 6 km, the gradient values are represented in °C/km.

From the 11 classes assigned to the Azufral volcano geothermal area, gradients with values between -6°C/m and 2°C/m were found. For the Eastern Llanos sedimentary basin, with 53 attributed classes, gradients with values between 1°C/km and 72°C/km were obtained. The training errors were approximately 42.0% (Azufral volcano geothermal area) and 76% (Eastern Llanos sedimentary basin), respectively. With the resulting confusion matrix for each estimate, it was possible to relate that the evaluation of thematic accuracy, for both study areas, had a very high kappa value, which infers that the condition number of the calculated matrix alters the output values due to small changes in the input arguments of the classification.

Through the analysis of the final maps obtained (qualitative analysis), a decrease in the range of values obtained for the geothermal gradient was observed. For the Azufral volcano geothermal area, the variation of the intermediate gradients went from -1.8°C/m and -1.0°C/m through the conventional methodology to -1.0°C/m and -0.4°C/m with the classification. For Eastern Llanos sedimentary basin, the qualitative analysis showed that the range of positive geothermal gradients (from 34°C/km to 58°C/km – assuming these values as anomalies) in the SW-NE direction for the conventional methodology, is located in the same position as the intermediate geothermal gradient values, with two groups of variations, between 21°C/km and 25°C/km, and 38°C/km and 72°C/km.

1. INTRODUCTION

In geothermal exploration it is important to have information on the geothermal gradient to identify the potential of the geothermal resource. In countries like Colombia, where the geothermal industry is still incipient, it is beneficial to reinforce exploration methodologies with techniques and methods derived from other geosciences different from 3G (geology, geophysics, and geochemistry), which contribute significantly to the conceptualization of the generated models. The geothermal potential of all Colombia, estimated at 1170 MWe ± 30 MWe (Alfaro et al., 2021), is calculated from the chemical geothermometers of the hydrothermal systems located in the volcanic complexes of the Central Cordillera of the Colombian Andes. It is of great importance to extend the generation of knowledge to the estimation of the potential in the sedimentary basins of the national territory.

The conventional methodology to calculate the geothermal gradient is to use temperature logs from drilled wells, or to use bottom hole temperature (BHT) if the logs do not exist. There are many oil wells in the sedimentary basins in Colombia. Their BHT has been used to provide the official version of the geothermal gradient map of the country (Alfaro et al., 2009). However, this information is not sufficient to make a complete study of the areas of interest for exploration. Other types of data such as shallow temperature surveys (STS) at depths close to the surface in volcanic systems can provide geothermal gradient data. There are no geothermal wells or oil wells within the hydrothermal systems of Colombia. Multipurpose and multidisciplinary geophysical inputs such as the total Bouguer gravity anomaly indicate specifically variations in rock density, from which a direct relationship with temperature variations with depth can be generated. Therefore, it is possible to approach the treatment and processing of the physical properties of the Earth in an innovative way from a data management point of view. This is where automated learning plays a major role. Different approaches such as neural networks have previously been applied to geothermal gradient solutions with gravimetry data (Mohamed et al., 2014; Mohamed et al., 2015), hidden systems (Brown et al., 2020), Play Fairway Analysis (Faulds et al., 2020), heat flow (Lösing and Ebbing, 2021) and district heating (Sretenović et al., 2018).

The present investigation describes the application of classification algorithms to the first vertical derivative of the total Bouguer gravity anomaly as training data, including geothermal gradient values calculated from STS (for a high-temperature volcanic system) and BHT (for a low-temperature sedimentary basin). The corresponding validation was executed by means of the confusion matrix for each classification, as well as a quantitative comparison of the correlation between the results found and the products already generated by conventional methodologies. An integration of the estimated geothermal gradients and the geothermal component of each area was performed and compared with previous models of each study area (conceptual model of the volcanic system and estimation of geothermal gradients and heat flow for the sedimentary basin).

2. STUDY AREAS

2.1 Geological settings of Colombia

The tectonic structure of Colombia, in particular the Andean zone, is a result of the interaction of the Cocos, Nazca, Caribbean and South American plates as well as the Coiba microplate (Figure 1). Regionally, this interaction results in a great variation of the stress tensors that mark out separate seismotectonic provinces (see e.g., Alfaro and Rodríguez-Rodríguez, 2020).

An important consequence of this phenomenon is the movement of the North Andean block towards the north-northeast (Figure 2). The North Andean block is bounded to the east by the fault system that runs from Guayaquil in Ecuador to the foothills of the Eastern Cordillera following the Boconó Fault in Venezuela. In the north, the movement of the North Andean block manifests offshore in the Caribbean

Deformed Belt, a later and more outer compressional phase of the accretion prism that includes the Sinú and San Jacinto deformed belts (Alfaro and Rodríguez-Rodríguez, 2020).

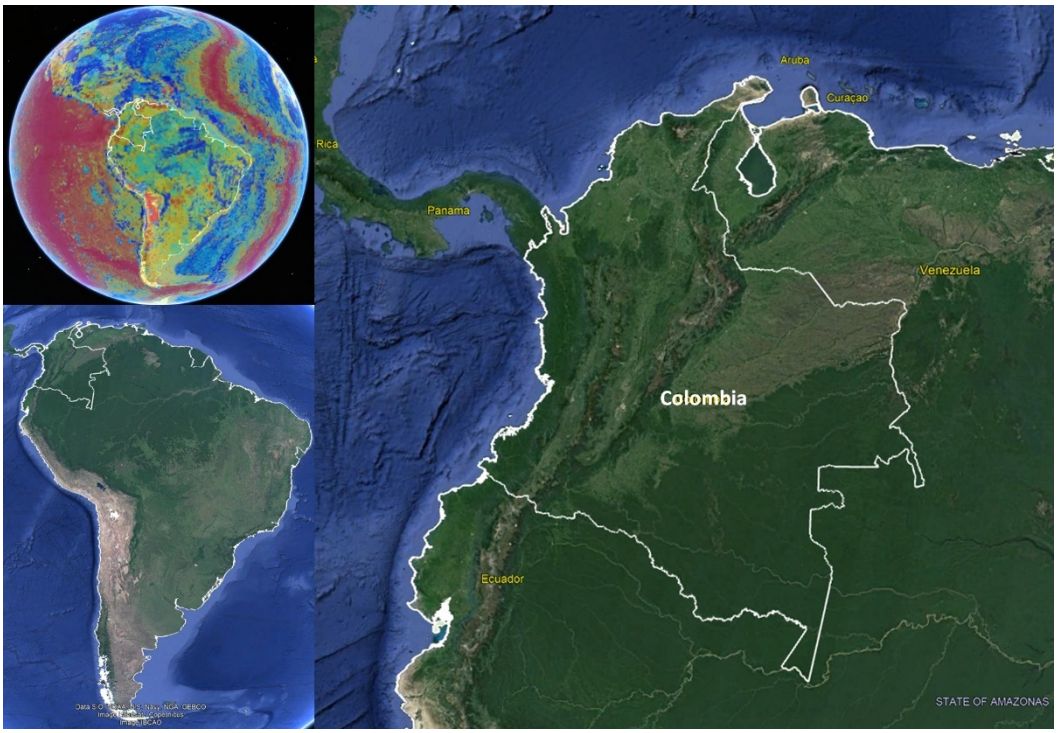

FIGURE 1: Location of Colombia in South America.
The global heat flow map as calculated by Lucazeau (2019) is in the top left corner

This deformation involves the sequences accumulated on top of the Caribbean Province's igneous basement, both before and after its accretion. The North Andean Block includes smaller triangular blocks bounded on all sides by wrench faults, such as the Maracaibo Block with the Santa Marta Block in the northeast corner. This complex pattern of plate and block interactions results in seismotectonic regimes on which transpressive and transtensive deformations play an important role (Gómez et al., 2007).

The basement of the Earth's crust in Colombia consists of two contrasting types of rocks that are separated by the Cauca-Almaguer fault. The fault is a structure that runs in a N-S direction along the western flank of the Andean Central Cordillera: To the east of this fault, it is made of Proterozoic metamorphic rocks, mainly sialic, whereas to the west of the fault it consists of Upper Cretaceous igneous, mainly volcanic rocks of simatic composition (Alfaro and Rodríguez-Rodríguez, 2020). A similar separation between sialic and simatic basements can be observed in the Guachaca and the Simarua faults at the Sierra Nevada de Santa Marta and at the Guajira Peninsula. The separation between these different types of basements is interpreted as a result of the accretion of oceanic lithosphere fragments to the continental active margin of South America during the Eocene (Gómez et al., 2007).

The Colombian territory has been divided into five crustal geological provinces, as presented in Figure 3. The description of these geological provinces is as follows: Within the eastern Proterozoic metamorphic basement, three geological provinces have been recognized, believed to be related to the Paleozoic history of collision between the Gondwana and Laurentia continents. From east to west these provinces are (INGEOMINAS, 2007):

1. Rio Negro-Juruena Province (RNJP): The Rio Negro-Juruena Province is a part of the Guayana Shield, the Gondwanic autochthonous craton around which the Laurentian continent's fragments were nucleated by amalgamation (formation of a crystal from a solution) to its west during the latter's northern drift relative to Gondwana.

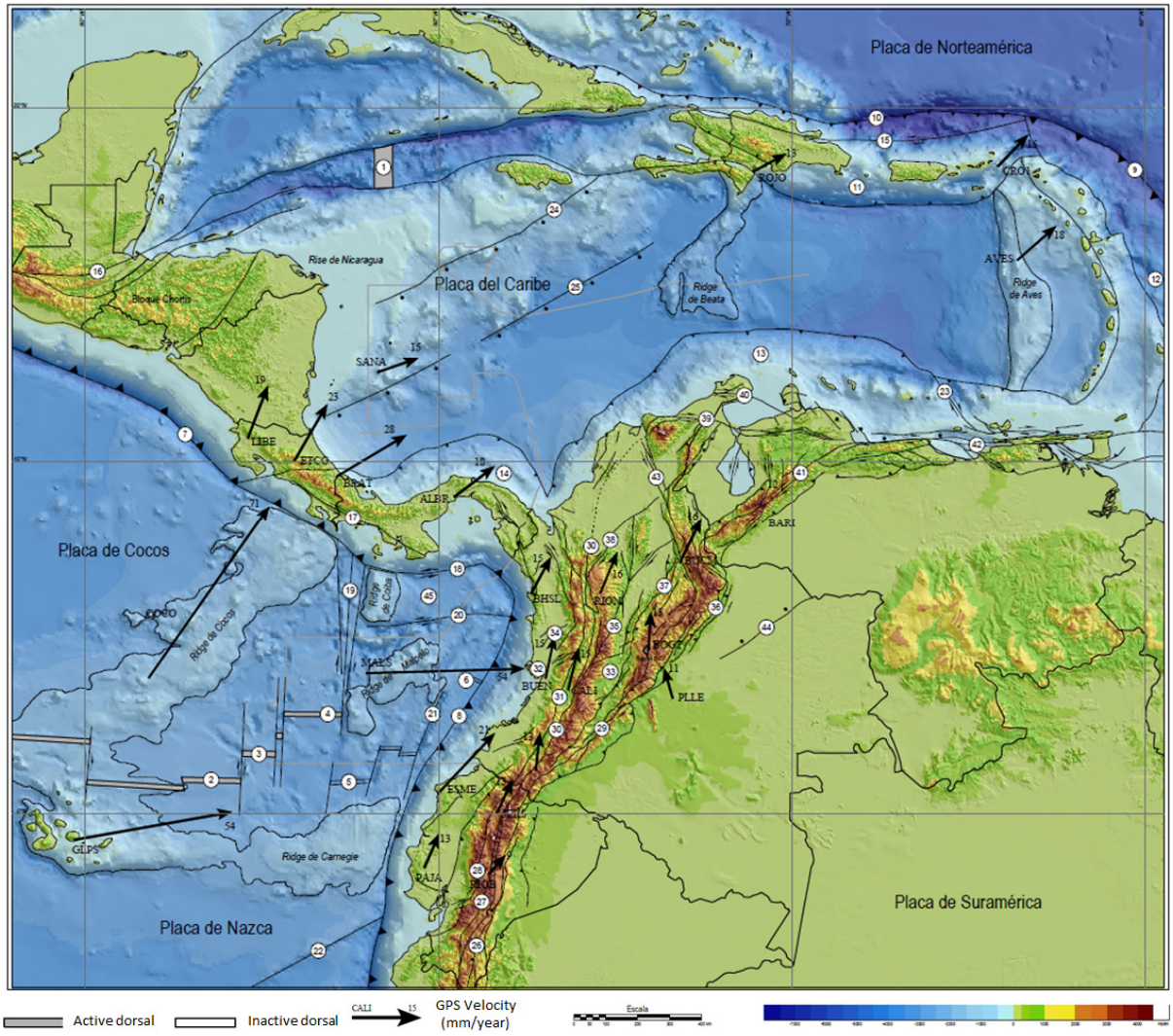

FIGURE 2: Tectonic scheme of the northern part of South America and the Caribbean (Gómez et al., 2007). Active Midocean Ridges: (1) Cayman, (2) Galapagos, (3) Ecuador) y (4) Costa Rica. Inactive Mid-ocean Ridges: (5) Malpelo, (6) Buenaventura. Oceanic trenches, active subduction zones: (7) Middle Aerican Zone, (8) Colombo -Ecuadorian Zone, (9) Caribbean. Oceanic trenches, inactive subduction zones: (10) Puerto Rico. Accretionary prisms– deformed belts: (11) Los Muertos, (12) Lesser Antilles, (13) Caribbean (14) Panama. Transformational fault zones:(15) North - East (16) Motagua – Swan, (17) Celmira – Ballena, (18) Jordan, (19) Panama, (20) Hey, (21) Yaquina, (22) Grijalva y (23) Los Roques. Oceanic Normal Faults: (24) Pedro Bank (25) Hess. Main continental plate faults: (26) Consaga, (27) Peltetec, (28) Pallatanga – Pujili, (29) Algeciras, (30) Cauca – Almaguer, (31) Cali – Patía, (32) Garrapatas, (33) Ibagué, (34) Itsmina Fault zone, (35) Palestina, (36) Guaicaramo, (37) La Salina, (38) Espiritu Santo (39) Oca, (40) Cuisa, (41) Boconó, (42) El Pilar, (43) Bucaramanga (44) Algarrobo, (45) Meta

2. Grenvillian Colombian Province (GCP): Amphibolites and granulites crop out as insulated faulted blocks at the Serranías de La Macarena y San Lucas, at the Garzón and Santander Massifs and at the Sierra Nevada de Santa Marta and Guajira Peninsula as well as roof pendants brought out by Mesozoic plutons.

3. Arquía Province (AP): The westernmost metamorphic basement corresponds to the Arquia Province of oceanic affinity that outcrops as a long and narrow belt bounded to the east by the Silvia-Pijao fault and to the west by the Cauca-Amaguer fault that is made of quartzsericitic, amphibolic schists, and amphibolites that are granatiferous in some locations.

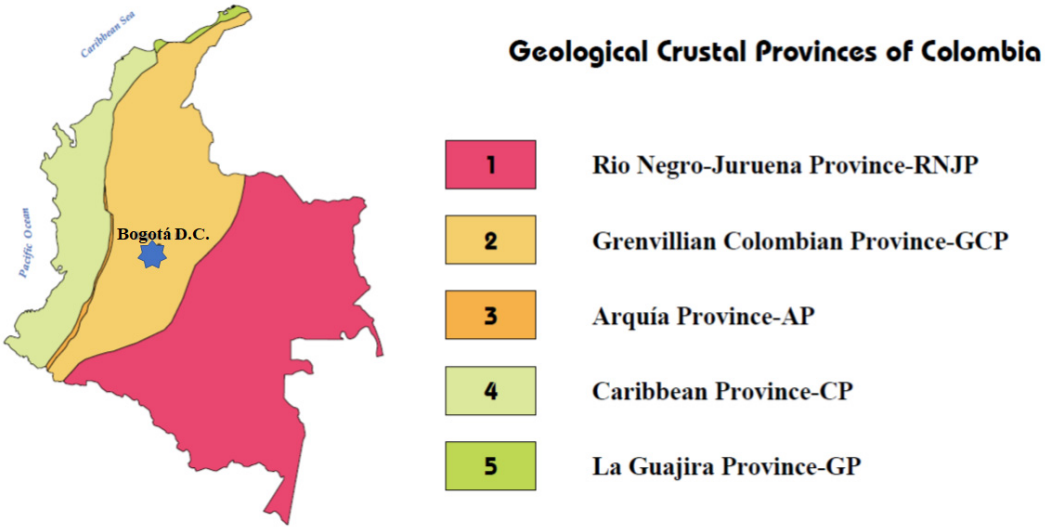

FIGURE 3: Geological Provinces of Colombia (INGEOMINAS, 2007)

4. Caribbean Province (CP): It is made of crustal fragments of oceanic affinity. The rocks of this province outcrops on the western foothills of the Central Cordillera, the western Cordillera and the Serranía del Baudó. These rocks consist of strongly faulted and deformed slivers of ultramafic rocks, associated gabbros and tonalites and also associations of basalts and felsites. There are also outcrops of picrites and komatites of which those of Gorgona Island are a unique example of Phanerozoic komatites. Lithological characteristics of this province, like the presence of picrites and komatites accompanied by the bimodal basaltic-rhyolitic composition of the volcanic rocks, as well as their trace element composition, suggests that these rocks were formed in an oceanic plateau.

5. La Guajira Province (GP): Oceanic cretaceous rocks of La Guajira and the Sierra Nevada de Santa Marta make up La Guajira Province. Given the low level of understanding, it has not been possible to establish the type of oceanic crust and the time of accretion.

2.1.1 Azufral volcano geothermal area

The Azufral volcano geothermal area (size around 20 km x 27 km) is located in the Western Cordillera in the southwestern part of Colombia (Figure 4). The area is classified as a high-temperature geothermal system by chemical geothermometers (Alfaro et al., 2015a). The crater of the volcano is 2 km x 3 km. Inside the volcano there are exogenous domes, which were built above volcanic edifices and produced important andesitic lava flows dated 0.58 ±0.03 M.a. (Bechon and Monsalve, 1991) and thick ignimbrite layers. No historical activity has been reported. It is considered to be an active volcano based on the age of its deposits, which ranges between 17,500 and 280 years B.P. (Bechon and Monsalve, 1991; Fontaine, 1994; Cortés and Calvache, 1997; Calvache, 1999; Torres et al., 2001). Both active surface hydrothermal manifestations (hot springs, fumaroles, and domes) and recent hydrothermal manifestations are found (Gómez and Ponce, 2009; Cortés et al., 2009).

Several intrusive bodies are found, like Azufral domes and Colimba domes, which are composed of pyroclastic deposits of andesitic composition and andesitic domes (Pinilla et al., 2007). The main basin comprises the Sapuyes river and Las Juntas river. It is composed of volcanoclastic deposits from the Cumbal, Azufral, and Pajablanca volcanoes, specifically some units like Cumbal Antiguo andesitic lavas, Quitasol and Olaya andesitic lavas, andesitic lavas of Pajablanca volcano and pyroclastic deposits from Azufral volcano (Pinilla et al., 2007). Basement rocks comprise the Dagua Group and Diabásico Group, the latter composed of basalts and meta-siltstones of late cretaceous age (González et al., 2002) with outcrops in the northeast part of the area.

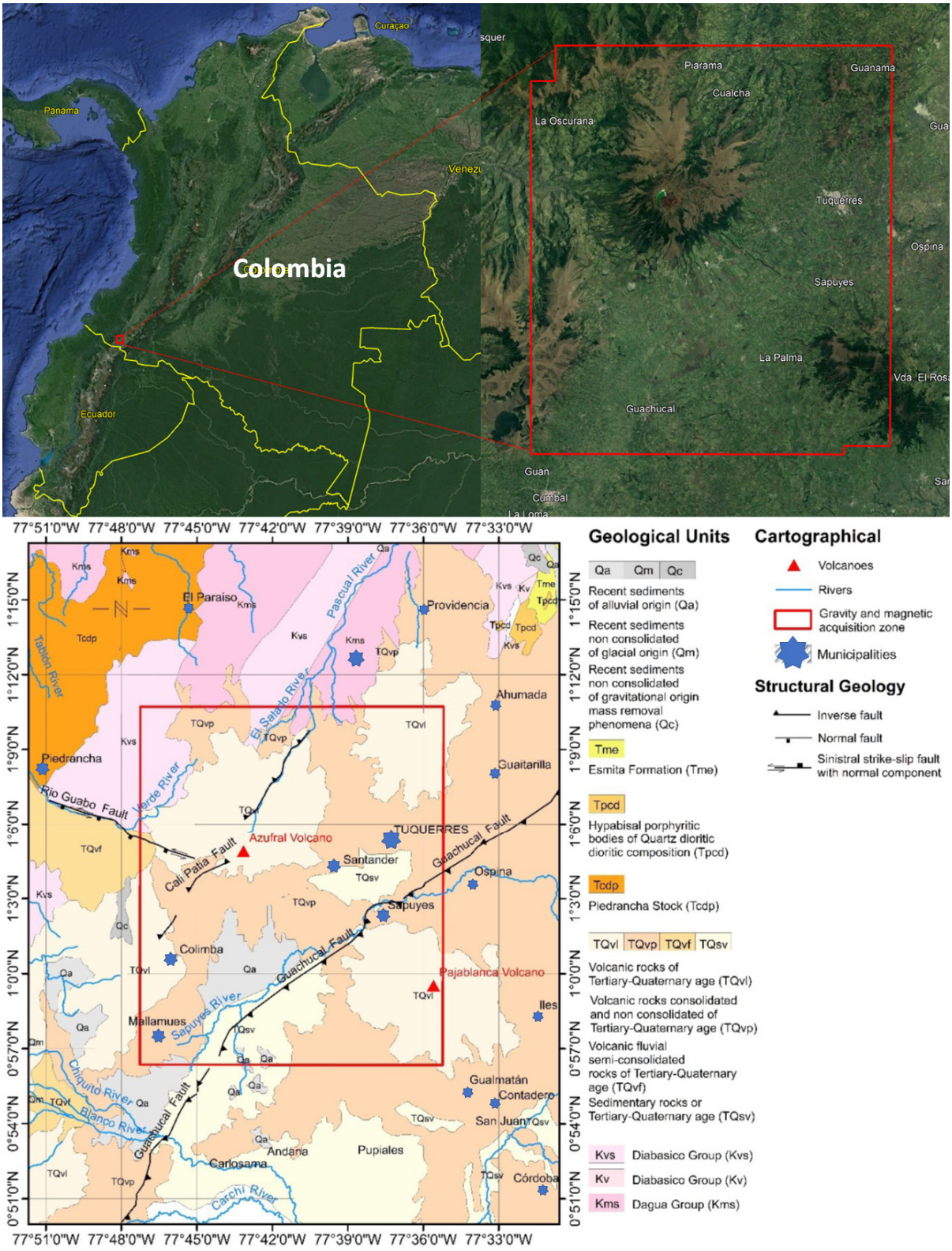

FIGURE 4: Location of Azufral volcano geothermal area (top) and a geological map of Azufral volcano (bottom) (Beltrán, 2020, based on Calvache et al. (2003); González et al., 2002); Pinilla et al., 2007; and Rodríguez and Rueda, 2017)

The Western Cordillera is characterized by active faults parallel to the Romeral system, between the Central and Western Cordilleras, along the Cauca-Patía valley. Although the faults of this system are dominantly reversed, towards the southwest (where Azufral volcano is located) they have a close to NE

direction and dextral movement (Taboada et al., 1998). Cauca-Patía fault (also named Cali-Patía fault) belongs to this system. This structure crosses the Azufral volcano crater and is partially covered by its pyroclastic deposits. Further to the south, it crosses the volcanic edifices of Cumbal, Chiles-Cerro Negro, and extends to Ecuador (González et al., 2002). Structural geology comprises the Cali-Patía fault. This fault separates an oceanic domain to the west and a continental domain to the east (Gónzalez et al., 2002).

The Cali-Patía fault has a left-lateral movement with a minor normal component in latitude over 5°. To the south, this fault has a reverse sense of movement with a minor right-lateral component (Rodríguez and Rueda, 2017). El Diviso-Túquerres fault consists of two parallel lineaments trending NW-SE as identified by remote sensors analysis (Cepeda, 1985). Its intersection with the Cauca-Patía fault would lead to the emergence of the Azufral volcano. The Guachucal fault, trending NE-SW, is a reverse fault with a dextral component and dipping to the east, with multiple variations in its strike along its trace and it controls the Sapuyes river. It is a pure strike slip in the south and an extensional strike slip fault in the north (Velandia et al., 2006).

To the north, this structure has a major reverse component (Rodríguez and Rueda, 2017). Some of the most relevant aspects of the geology regarding the geothermal system include the rhyodacitic composition of the aforementioned deposits which show a higher magma differentiation when compared to nearby volcanoes of andesitic composition. This suggests that the magma has remained and heated up the crust, resulting in a high temperature anomaly. The existence of young rhyodacitic domes (~4,000 years) highlights the new magma and shallow heat contribution, and the pyroclastic surge deposits indicate the abundance of water, likely coming from a geothermal reservoir. The thick lava and welded ignimbrites, related to ancient Azufral's volcanic edifices and other nearby volcanoes, could host a secondary permeability geothermal reservoir. Altered pyroclastic flows and surge deposits could act as a cap rock for the reservoir.

2.1.2 Eastern Llanos sedimentary basin

The Eastern Llanos sedimentary basin covers an area of approximately 200,000 km^2 and is limited to the west by the foothills of the Eastern Cordillera, to the north by Cordillera de Mérida (northeast section of the Eastern Cordillera), to the east by the Guyana Precambrian shield and to the south by the Amazonas and Caguan-Putumayo basins (Bachu et al., 1995). The geographical location and a geological cross-section of the sedimentary basin is shown in Figure 5.

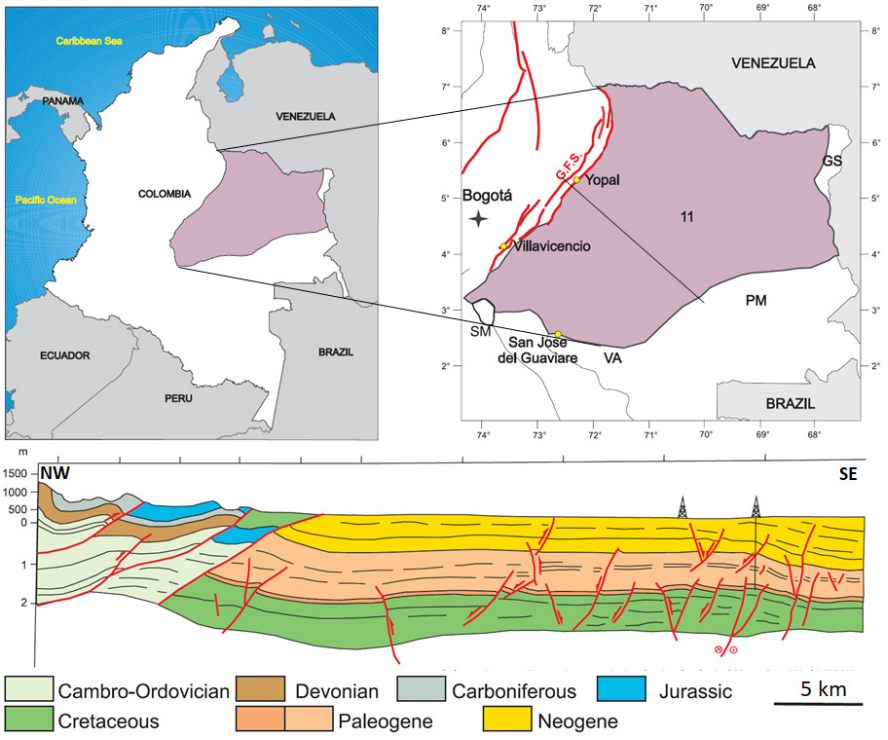

FIGURE 5: The Eastern Llanos sedimentary basin and a geological cross-section of the area (Barrero et al., 2007)

The Eastern Llanos sedimentary basin is located in the northern part of Los Andes Cordillera range, at the end of a series of sub-Andean basins that generally extend eastwards, from south Argentina to north Venezuela (Bayona et al., 2008). Geographically, the Eastern Llanos sedimentary basin is limited to

the north by the boundary with Venezuela and includes the departments of Arauca (to the north), Casanare (north-central), Meta (southwest) and Vichada (east) (Barrero et al., 2007). Within the geotectonic framework, the Eastern Llanos basin is a sub-Andean foreland basin (Dueñas Jiménez and van der Hammen, 2007), located between the Guyana Precambrian shield to the east and the Eastern Cordillera to the west (Cooper et al., 1995). The northwest corner of South America has experienced different geological events that have controlled the distribution, genesis, filling and structural boundaries of the sedimentary basins in this region (Bachu et al., 1995). The Eastern Llanos basin is a part of the eastern tectonic domain of Colombia, which is limited to the west by the Eastern Cordillera foothills (Barrero et al., 2007). This domain consists of a Precambrian and Paleozoic basement (Guyana Shield), with a Paleozoic-Cenozoic sedimentary cover that has undergone mild deformation (Barrero et al., 2007).

The sedimentary basin consists primarily of a siliciclastic sedimentary sequence (Figure 6), with ages ranging from Paleozoic age to the most recent deposited in an unusual tectonic framework on the crystalline basement of the Guyana shield (Cooper et al., 1995). The Guayana shield consists of a very old remnant (more than 2,700 million years) of high-grade metamorphic terranes that were remobilized in a large-scale thermal event approximately 2000 million years ago (Bachu et al., 1995). Subsequently, they were covered by sedimentary rocks, which underwent intense metamorphism 1000 to 1600 million years ago (Grenvillian Orogeny). In turn, acidic igneous rocks intruded these sedimentary rocks during the Orinoco episode approximately 1,300 Ma ago (Barrero et al., 2007). The basement in most of the sedimentary basin of the Eastern Llanos corresponds to metamorphosed Neo-Proterozoic sedimentary rocks (Bachu et al., 1995). Figure 6 shows a chart of the stratigraphic column and a cross-section of the formations in a northwest-southeast direction for characterizing the outcropping crystalline basement on the surface (García González et al., 2009).

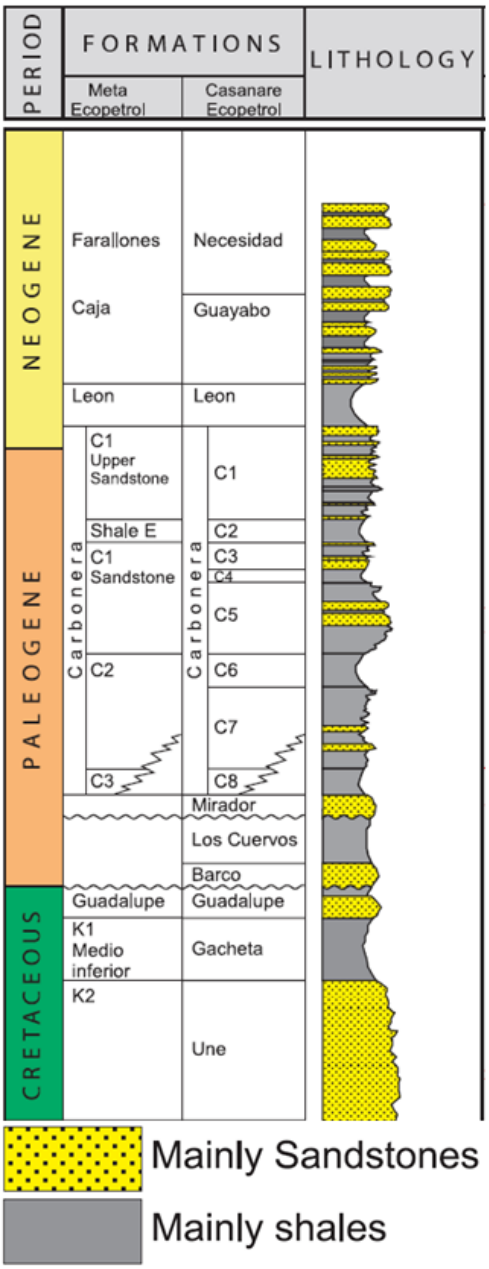

FIGURE 6: Chart of the generalized stratigraphic column with the formations of the Eastern Llanos basin (Barrero et al., 2007)

2.2 Geothermal exploration in Colombia

Colombia hosts convective geothermal systems resulting in hydrothermal surface manifestations related to volcanoes and faults. In the sedimentary basins, there are indications of possible conductive systems (Figure 7). Most of the geothermal systems related to volcanoes are in the Central Cordillera and in the south of the Western Cordillera. However, an isolated geothermal area is found in the Eastern Cordillera: Paipa, related to an ancient volcano and Iza, related to a cryptodome (Alfaro and Rodríguez-Rodríguez, 2020). Geothermal exploration has mainly been focused on the hydrothermal systems. The systems are presumably related to faults found in the Andes but also in the Caribbean, Pacific and Eastern Llanos sedimentary basin (Alfaro et al., 2021).

Assuming the preliminary geothermal gradient anomalies in sedimentary basins are indicative of

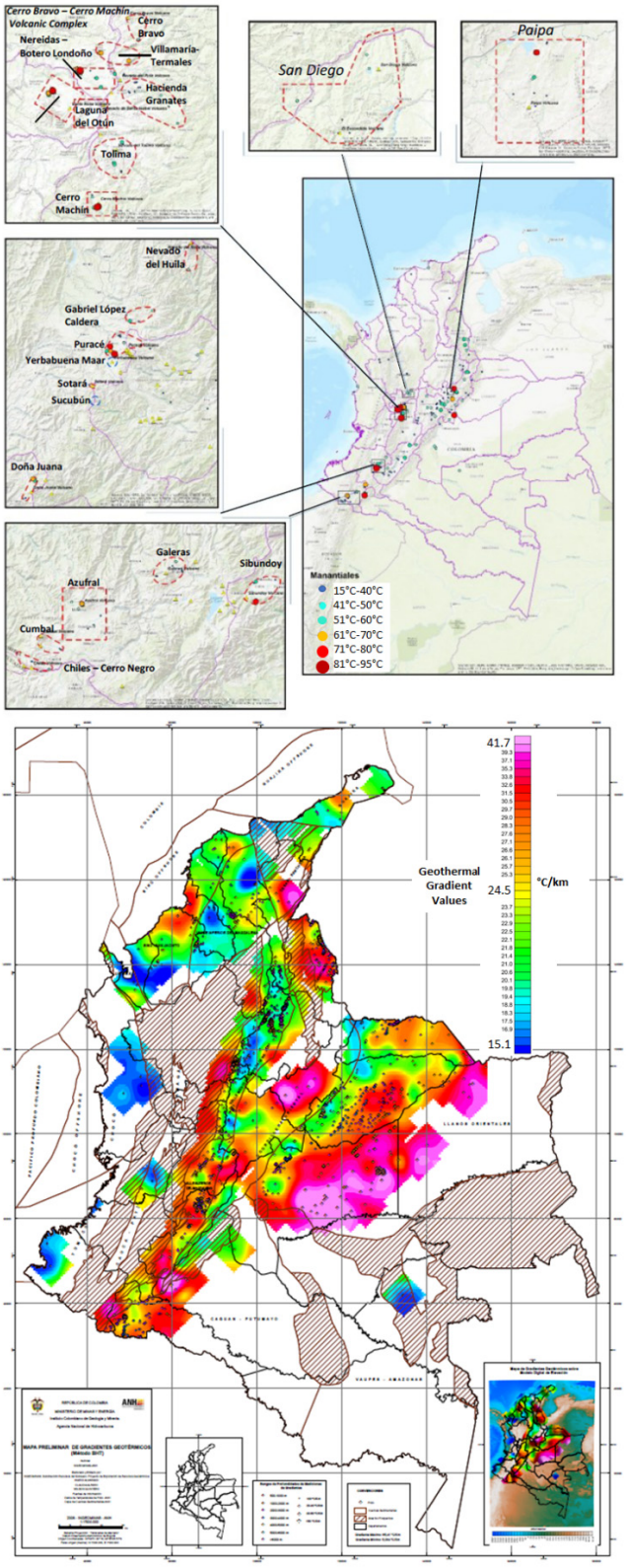

FIGURE 7: Geothermal resources in Colombia (Alfaro and Rodríguez-Rodríguez, 2020). Upper figure: Convective systems related to volcanoes and faults, on the hot springs map (Alfaro et al., 2021). Lower figure: Geothermal gradient map (INGEOMINAS-ANH, 2008)

possible low enthalpy conductive systems where temperatures of above 150°C at more than 5 km depth have been registered, the most promising areas for this type of resource are the following sedimentary basins: Eastern Llanos, Cagüan – Putumayo, Catatumbo, Central Cordillera and Middle Magdalena Valley (INGEOMINAS-ANH, 2008).

The Colombian Geological Survey (SGC) has been trying for more than 50 years to promote the development of Colombia's geothermal potential through the exploration of the country's geothermal resources (Alfaro et al., 2000). The main geothermal study areas are Paipa, Azufral volcano, San Diego, and new geothermal areas in the Cerro Bravo-Cerro Machín volcanic complex (Alfaro and Rodríguez-Rodríguez, 2020).

The area with the largest number of geoscientific studies is the Paipa geothermal area. The particularity of this geothermal area is that it is isolated from the main volcanic systems (located in the Eastern Cordillera in a sedimentary basin). Various surface studies have been concluded. The latest conceptual model constructed by Alfaro et al. (2020) (Figure 8) contains domes mapping and dating (Rueda, 2017), descriptions and chemical analysis of core samples from shallow wells (50 and 100 m) and of an altered igneous intrusion (El Durazno) (Rodríguez and Alfaro, 2015), a 3D geological model constrained by gravity and magnetic data (Llanos et al., 2015), 2D and 3D resistivity modelling (González-Idárraga and Rodríguez-Rodríguez, 2017; Siripunvaraporn, 2016 in Alfaro and Rodríguez-Rodríguez, 2020), updating of the geoelectrical study (VES) (Franco, 2016 in Alfaro and Rodríguez-Rodríguez., 2020), and a local isotopic meteoric line (Malo and Alfaro, 2020). Two geothermal gradient wells were drilled in the geothermal area, reaching depths between 432 m and 500 m, with core recovery, identification of the sedimentary sequence, aquifer levels with known thickness and geothermal gradient values (around 180°C/km) (Rodríguez, 2020).

The Azufral volcano geothermal area is one of the two study areas in this work. In the geothermal area, geoscientific surface studies have been completed, consisting of an MT

resistivity survey (Meqbel, 2017 in Alfaro et al., 2015b), 3D model (Rodríguez-Rodríguez, 2020), and preliminary radon surveys in soil and air (Malo and Alfaro, 2018), in addition to geological studies, potential field studies (magnetics and gravity), and geochemistry whose results are included in the preliminary conceptual model (Alfaro et al., 2015b). The updated conceptual model (Figure 8) will be presented by Alfaro-Valero et al. (in prep).

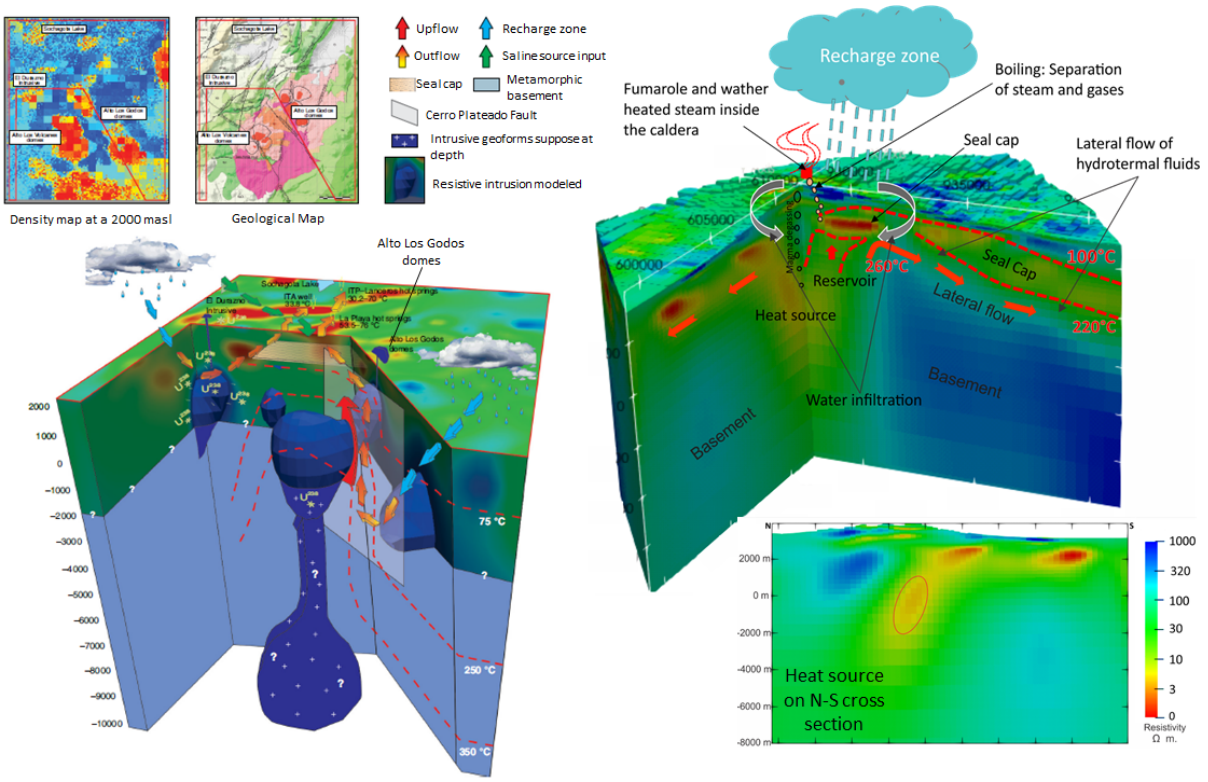

FIGURE 8: Conceptual models of the Paipa geothermal area (Alfaro and Rodríguez-Rodríguez, 2020) to the left and the Azufral volcano geothermal area (Alfaro-Valero et al., in prep) to the right

Colombia's northernmost geothermal area is called the San Diego geothermal area. It is delineated by hot springs reported in the national inventory of hydrothermal manifestations and includes the San Diego Maar and El Escondido de Florencia volcano (discovered during the reconnaissance mission before starting the geological work of the area (Monsalve, 2013). The activities carried out involve geological mapping in scale 1:25,000, analysis of the geochemistry of fluids, as well as gravity, magnetic and magnetotelluric resistivity studies (Alfaro and Rodríguez-Rodríguez, 2020).

The first geothermal well in Colombia, Nereidas-1, was drilled in 1998 (Monsalve et al., 1998) in the Nevado del Ruiz geothermal area, the most studied geothermal area in Colombia (CHEC, 1968). The well remains the only deep geothermal well drilled so far in Colombia and has been of interest for stakeholders of the power generation sector such as CHEC-EPM and ISAGEN. For new exploration activities, the Paramillo de Santa Rosa volcano geothermal system was selected due to the high discharge temperatures of hot springs (up to 91°C), the type of water (sodium chloride) and high resource temperature based on geothermometers calculations (>225°C) (Alfaro et al., 2002). The Cerro Machín volcano geothermal area was chosen because of the hot springs discharge temperatures (up to 96°C), which were measured recently by the SGC, and high estimated resource temperatures (200 to 220°C) (Inguaggiato et al., 2017). However, the hot springs are sodium bicarbonates water type with low chloride, boron and lithium concentrations. Currently, other exploration studies are being completed in the Cerro Machín area like acquisition of gravity, magnetic and magnetotelluric data, as well as geothermal fluid sampling for updating the water analysis and gas phase characterization (Alfaro and Rodríguez-Rodríguez, 2020).

2.3 The beginning of geothermal utilization

Based on the exploration carried out in the country, two cases of success in the use of geothermal energy have been recorded. The first heat pump for cooling purposes was installed in an industrial park located in Tocancipá, 40 km north of Bogotá, D.C. (Alfaro and Rodríguez-Rodríguez, 2020). The heat pump works full time and cools a room of 90 m^3 down to -10°C. The change in the temperature of the water that is circulated in the subsoil at 2500 l/h is 6°C, from 15°C to 21°C. Two vertical wells of 70 m and one of 80 m depth support the system (Ortiz, 2017).

For geothermal power generation, Parex Resources Inc. (a joint project with Universidad Nacional de Colombia-Medellin and the national government through the Ministry of Mines and Energy) inaugurated a small-scale 100 kW unit, capable of generating up to 72,000 kWh/month, at the Las Maracas field in Casanare (Eastern Llanos sedimentary basin). The generation equipment was designed, built, and commissioned by the Spain-based ORC manufacturer Rank. The utilization of the resource helps to reduce emissions from fossil fuel-sourced power generation by around 550 tons of CO_2 per year. The implementation takes advantage of high temperature gradients, permeable rocks, and fresh water which could be brought to the surface without additional cost as a co-product of oil extraction (Franco et al., 2021).

3. METHODS

3.1 Gravimetric processing

Of the geophysical techniques used in the exploration of geothermal resources, gravimetry focuses on identifying changes in rock density and density anomalies. In turn, it is possible to identify lateral density variations in the subsurface with a focus on structural interpretation to find depth variations in the basement, caldera rim, intrusions, rock alteration, porosity variations, faults, or dykes (Hersir et al., 2022). Even at the utilization stage, by measuring gravitational acceleration and elevation changes over time, it is possible to estimate the net total mass extraction from the reservoir (Gudnason et al., 2015).

3.1.1 Bouguer gravity anomaly

To measure the spatial variation of the gravitational acceleration, gravimetry requires corrections for instrumental drift, latitude, solar and lunar effects, elevation and terrain to obtain the final Bouguer gravity anomaly map (Hersir et al., 2022). The calculation of the Bouguer anomaly (mGal) is expressed as (Kearey et al., 2002):

$$BA = g_{obs} - g_{\phi} + FAC \pm BC + TC(\pm EC) \quad (1)$$

where g_{obs} = Observed gravity (mGal);
g_{ϕ} = Latitude correction (correction for the gravity variation due to latitude, i.e., the non-spherical shape of the Earth);
FAC = Free-air correction (in mGal – correction for the different elevation of gravity stations with reference to the geoid (sea-level); its calculation is determined as $FAC = 0.3086\ (mGal) * h$, where h is the height of the station in meter about see level);
BC = Bouguer correction (correction is applied for the excess mass material between the station and sea-level);
TC = Terrain correction (correction made to account for topographic relief in the vicinity of the gravity station); and
EC = Eötvös correction (correction applied to gravity measurements taken on a moving vehicle, such as a ship or an aircraft).

3.1.2 Regional-Residual separation

The Bouguer anomaly is characterized as a regional anomaly with local variations, which makes it necessary to remove or separate the regional field from the residual anomalies (Kearey et al., 2002). Generally, the residual anomalies of interest are derived from relatively shallow sources (Hinze et al., 2013). In the literature, the removal of the residual signal is generally referred to as "regional-residual separation" and various analytical techniques can be used to perform the process (Lowrie, 2007).

One method to isolate the shallow signal is using continuation filters. The upward continuation filter projects (recalculates) the observed gravity anomaly to a higher elevation. A smooth, long-wavenumber pass filter emphasizes the anomalies from broader, deeper sources at the expense of the shallow-source anomalies and noise in the anomaly field (Hinze et al., 2013). To calculate the height at which the regional-residual separation of the signal should be made, the power spectrum method is used. This method consists of a graphic that allows seeing the behaviour of the energy logarithm of a signal as a function of its frequency or wavenumber (Beltrán, 2020). Taking as a reference the equation used by the Oasis Montaj software for the estimation of the regional grid obtained by the upward continuation, the filter $L(k)$is given as (Seequent, 2020):

$$L(k) = e^{-2\pi hk} \tag{2}$$

where h = Distance in ground units relative to the plane of observation (m); and
k = Wavenumber.

The residual contribution is calculated from the difference between the original grid and the regional grid (obtaining the maximum depth of the new residual signal computed by the power spectrum).

3.1.3 First vertical derivative

Assuming that the first vertical derivative (FVD) is the rate of vertical change of gravity (Figure 9), the calculation of the derivative accentuates the representation of the shallow features and the density gradient to infer and locate anomalies (Hersir and Björnsson, 1991; Dentith and Mudge, 2014; Hinze et al., 2013; Kearey et al., 2002; Lowrie, 2007; Jacoby and Smilde, 2009). The first vertical derivative was calculated to enhance small details related to geological structures. It can help to recognize the signatures of small-scale features that are not easily recognized in the Bouguer anomaly map (Mohamed et al., 2014). Derivatives can be calculated with Fast Fourier Transform (FFT) since the transformation of the data to remove cultural noise can be complex and symmetric. The equation below shows the application of the first vertical derivative (mGal/m):

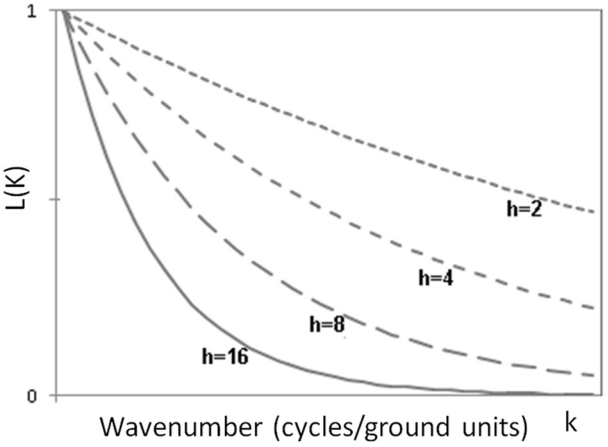

FIGURE 9: Graphics for the height (h) in meters to be applied (obtained from the signal power spectrum) in separating the residual signal from the noise (Seequent, 2020)

$$FVD = -\frac{dBA_{residual}}{dz} \tag{3}$$

where $\frac{dBA_{residual}}{dZ}$ corresponds to the residual signal of the total Bouguer anomaly (mGal); and means that the derivative is applied in the direction of the vertical component (depth, m).

3.2 Machine learning (classification algorithms)

Machine learning is an alternative to conventional researching approach for the design of an algorithmic solution. The conventional research design flow starts with the acquisition of domain knowledge (the problem of interest is studied in detail) and then a mathematical model is produced that captures the physics of the set-up under study. In comparison, the machine learning approach is based on the collection of a sufficiently large number of samples of "desired behaviour" to become a training set for the algorithm. In Figure 10, the machine learning methodology is explained. Acquisition of data for the training set is combined with domain knowledge which allows for the selection of an appropriate hypothesis class (Simeone, 2018).

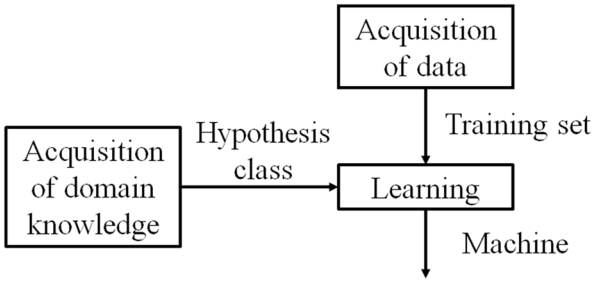

FIGURE 10: Machine learning methodology (Simeone, 2018)

Supervised training is a part of the machine learning approach. Classification is the process of dividing a set of data into categories. It can be done on both structured and unstructured data and predicting the class of provided data points is the first step in the procedure (Rezvanbehbahani et al., 2017; Lösing and Ebbing, 2021). The classification predictive modelling maps input variables to discrete output variables, where the main goal is to identify into which class/category the new data fall, specifically in binary classification algorithms (two outcomes) (Brown et al., 2020; Faulds et al., 2020; Smith et al, 2021). In this study, two types of classifiers are applied to construct a classification model based on the training data before getting data for predictions. These are known as eager learners and correspond to support vector machines (SVM) and Random Forest.

3.2.1 Support vector machine

The support vector machine (SVM) consists of heuristical algorithms based on statistical learning theory (Vapnik, 1995). The SVM uses a set of data vectors with known class labels acquired through a priori knowledge to design a linear hyperplane for separating various classes (Figure 11) where the data vectors serve as a training set and each data vector within the set is distinguished by unique features on which the classification is based (Abedi et al., 2012). SVM is also called a nonparametric model, which means that the learning (training) is critical. The parameters are not predefined, and their number depends on the training data used (parameters that define the capacity of the model are data-driven in such a way as to match the model capacity to data complexity) (Sretenović et al., 2018). To apply SVM to problems in the real geoscience world that needs practical solutions, the data points of different class memberships (clusters) overlap one another, and this makes linear separability difficult as the basic linear decision boundaries are often not sufficient to classify patterns with high accuracy (Mountrakis et al., 2011).

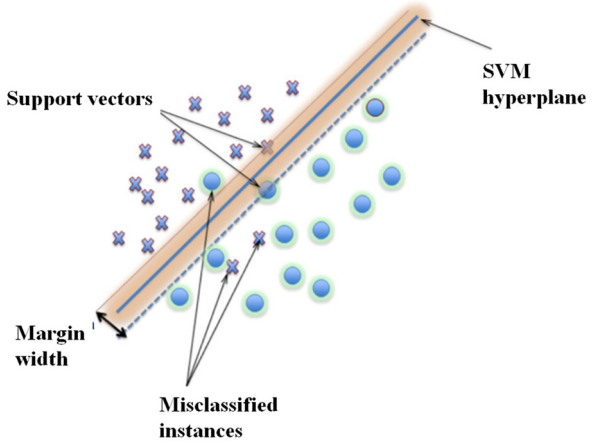

FIGURE 11: Linear Support Vector Machine example (Mountrakis et al., 2011)

3.3 Geothermal gradient

The geothermal gradient is the temperature variation with depth. The gradient is more irregular at the surface than in the mantle or core, with large lateral changes (Clauser, 2009). In subduction zones, low gradients are observed due to the cooling produced by sediments saturated with water carried under the

crust (Alfaro et al., 2009). In tectonically stable areas, such as shields and sedimentary basins, the average gradient ranges from 15 to 30°C/km (Pollack et al., 1993). In basement rocks, such as granites, which contain higher concentrations of radiogenic minerals than other types of crystalline rocks (Alfaro et al., 2015a), the heat generated by radioactive decay could become a target for extraction (hot dry rock geothermal systems) if the overlying rocks have low thermal conductivity and act as a thermal insulating coating, trapping heat within the granite (Clauser, 2009).

Geothermal gradients calculated from bottom-hole temperatures correspond to the gradients estimated from the bottom-hole temperature measurements at the maximum depth with respect to the mean surface temperature (Carvalho and Vacquier, 1977). This is an apparent gradient because it disregards the vertical variation associated with lithological changes and the heat transfer properties of the rock (Alfaro et al., 2015a). For the calculation, it is also of great importance to know if the surrounding rock was in thermal equilibrium after drilling. From the heat flow equation (4), the calculation of the geothermal gradient is derived in the following way:

$$q = \left\{\left(\frac{T_{BHT} - T_0}{Z_{BHT} - Z_0}\right)\right\} * \left\{\left(\sum_{i=1}^{n} R_i h_i\right)\right\} \tag{4}$$

$$heat\ flow = geothermal\ gradient\ (°C/m) * thermal\ conductivity(W/m°C)$$

where q = Heat flow (mW/m^2);
T_{BHT} = Bottom hole temperature (°C);
T_0 = Surface temperature (°C);
Z_{BHT} = Bottom hole depth (m);
Z_0 = Surface elevation (m); and
$\sum_{i=1}^{n} R_i h_i$ = Cumulative thermal conductivity (mW/m°C).

3.3.1 Bottom-hole temperature (BHT) corrections

Geothermal gradients can be estimated using the apparent geothermal gradients if wellbore temperature profiles are not available (Alfaro et al., 2009). The standard method in the oil industry is to estimate geothermal gradients by vitrinite reflectance analysis (a measurement of the maturity of organic matter), whose variation is linear and thus requires no comparative validations with the rock composition (Kellogg et al., 2005).

The BHT method including corrections is the method that least disturbs the temperature measurements after drilling, but the measured values are not the precise formation temperatures due to the thermal disturbance of drilling (Alfaro et al., 2009) since any empirical or theoretical correction to BHT data incorporates an unknown amount of error (Alfaro et al., 2015a).

The American Association of Petroleum Geologists (AAPG) correction is applied to approximate the true formation temperature values. This empirical correction was first applied to well measurements for geothermal purposes in the 1970s and compares the bottom-hole temperatures with the equilibrium temperatures measured in 602 wells located in the United States (Deming, 1989). The differences between the two data sets (bottom-hole temperatures and equilibrium temperatures) are adjusted by applying a third order least squares polynomial to calculate the mean geothermal gradient correction as a function of depth (Deming, 1989). The resulting gradient correction can be multiplied by the depth to correct the formation temperature as a function of depth (Deming, 1989). The third order least squares polynomial corresponds to:

$$\Delta T = az + bz^2 + cz^3 + dz^4 \tag{5}$$

where ΔT = Temperature correction (°C);
z = Depth of the well (km); and
a, b, c, d = Coefficients of the third-order polynomial equation.

The topographic correction of geothermal gradients is estimated primarily by correcting the temperature measurements taken in exploration wells (Balling et al., 1981). This correction, albeit approximate, considers the complexity of the topography, is easily applicable to deep exploration wells, and provides sufficiently accurate corrected data (Bullard, 1938).

3.4 Proposed methodology

The research methodology called Numerical Baseline and Classification Algorithms (NUBA) is divided into two stages (Figure 12). The first phase (numerical baseline) consists of processing the total Bouguer anomaly, realizing the regional-residual separation by using the upward continuation filter, a process that relies on the power spectrum to find the depth of the shallow signal with the regional noise. Once the residual signal is obtained, the first vertical derivative is applied to obtain the variation of gravity as a function of the depth of the signal.

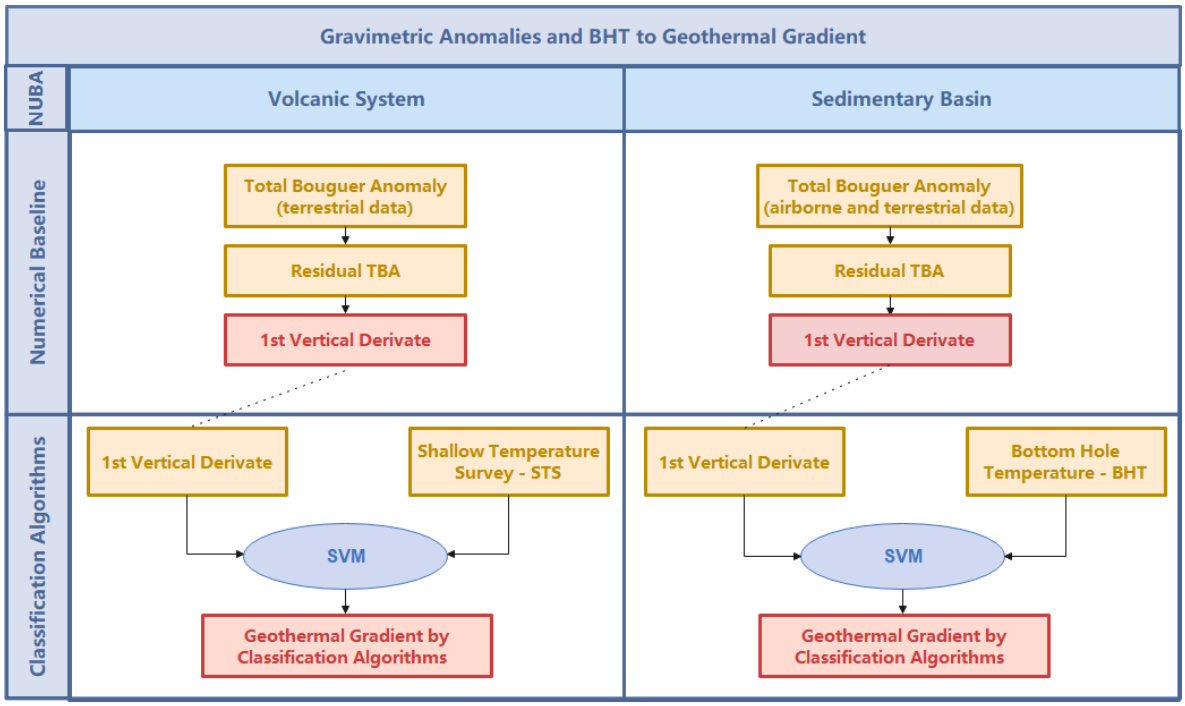

FIGURE 12: Proposed research methodology

The second phase (classification algorithms) applies the machine learning approach with a focus on the classification algorithms to estimate the geothermal gradient from BHT data in the sedimentary basin and from STSs in the volcano system, respectively.

4. INPUT DATA

4.1 Gravity anomalies

The training data correspond to the first vertical derivative of the residual component of the total Bouguer anomaly. For each study area, the availability of information was checked, and the type of survey stations was verified. The parameters used to correct the observed gravity and obtain the total Bouguer anomaly were also described. Due to the different stations in each study area, an upward continuation filter was applied to perform the regional-residual separation. The software used was Geosoft Oasis Montaj, currently owned by Seequent (Seequent, 2020). Power spectrum estimation

calculations were done for Azufral volcano geothermal area and the Eastern Llanos sedimentary basin (Figure 13 and 14).

4.1.1 Azufral volcano geothermal area

The total Bouguer anomaly was generated from 303 gravimetric stations (Beltrán, 2020). According to the power spectrum estimation, a depth of investigation of approximately 4 km was reached in the geothermal area (Figure 13). This depth was taken as a reference value to carry out the regional-residual separation by means of the upward continuation filter. Once the residual anomaly was calculated (subtracting the regional noise from the total Bouguer anomaly), the first vertical derivative was calculated. Figure 15 shows the total Bouguer anomaly, the regional anomaly, the residual anomaly and the first vertical derivative for the Azufral volcano geothermal area.

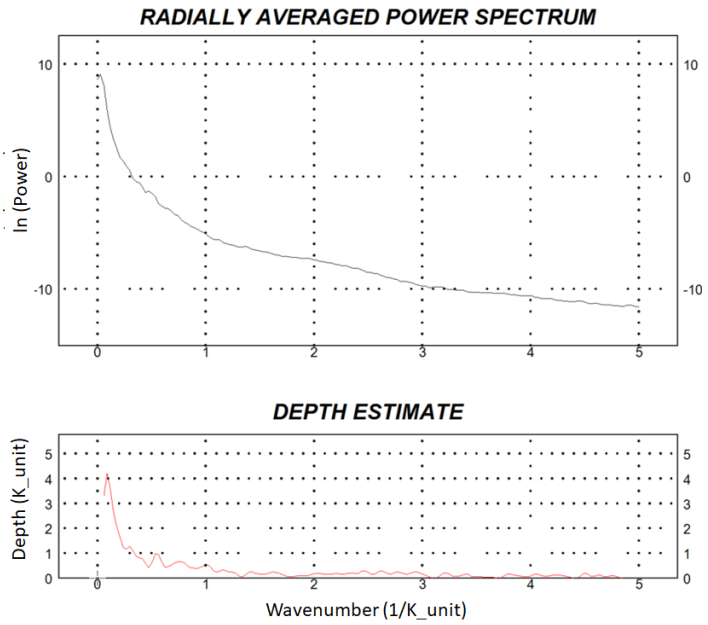

FIGURE 13: Estimated depth of investigation and the power spectrum

4.1.2 Eastern Llanos sedimentary basin

The information available on the sedimentary basin corresponds to the total Bouguer anomaly map of the Republic of Colombia (Graterol Graterol and Vargas Gomez, 2010). The anomaly map was generated from both air and land gravity observations in Colombia, with marine stations in the Caribbean and the Pacific, and with land and air stations in Venezuela and eastern Panama. For the corrections of the anomaly, the NASA digital terrain model, the 90 m SRTM, was used. For the calculation of the simple and total Bouguer anomaly on land, in the sea and air, the International Gravity Standard Network 1971 (ISGN71) and the World Geodetic System 1984 (WGS84) were used as reference. Densities of 2.67 g/cm^3 for land and 1.03 g/cm^3 for ocean water were used for Bouguer and terrain corrections (Graterol Graterol and Vargas Gomez, 2010). In the sedimentary basin, a depth of inquiry of nearly 60 km was reached, according to the power spectrum estimation (Figure 14). The upward continuation filter was applied to perform the regional-residual separation using this depth as a reference value. The total Bouguer anomaly, the regional anomaly, the residual anomaly and the first vertical derivative for the Eastern Llanos sedimentary basin are shown in Figure 16.

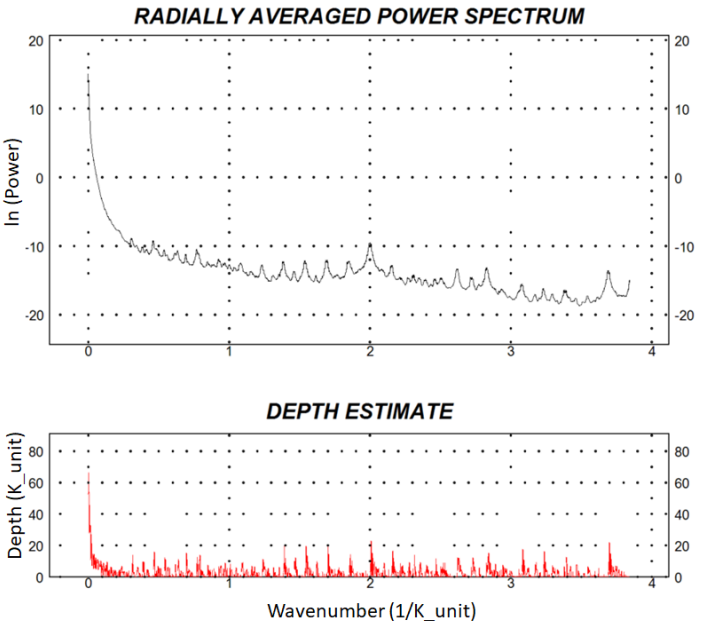

FIGURE 14: Estimated depth of investigation and the power spectrum

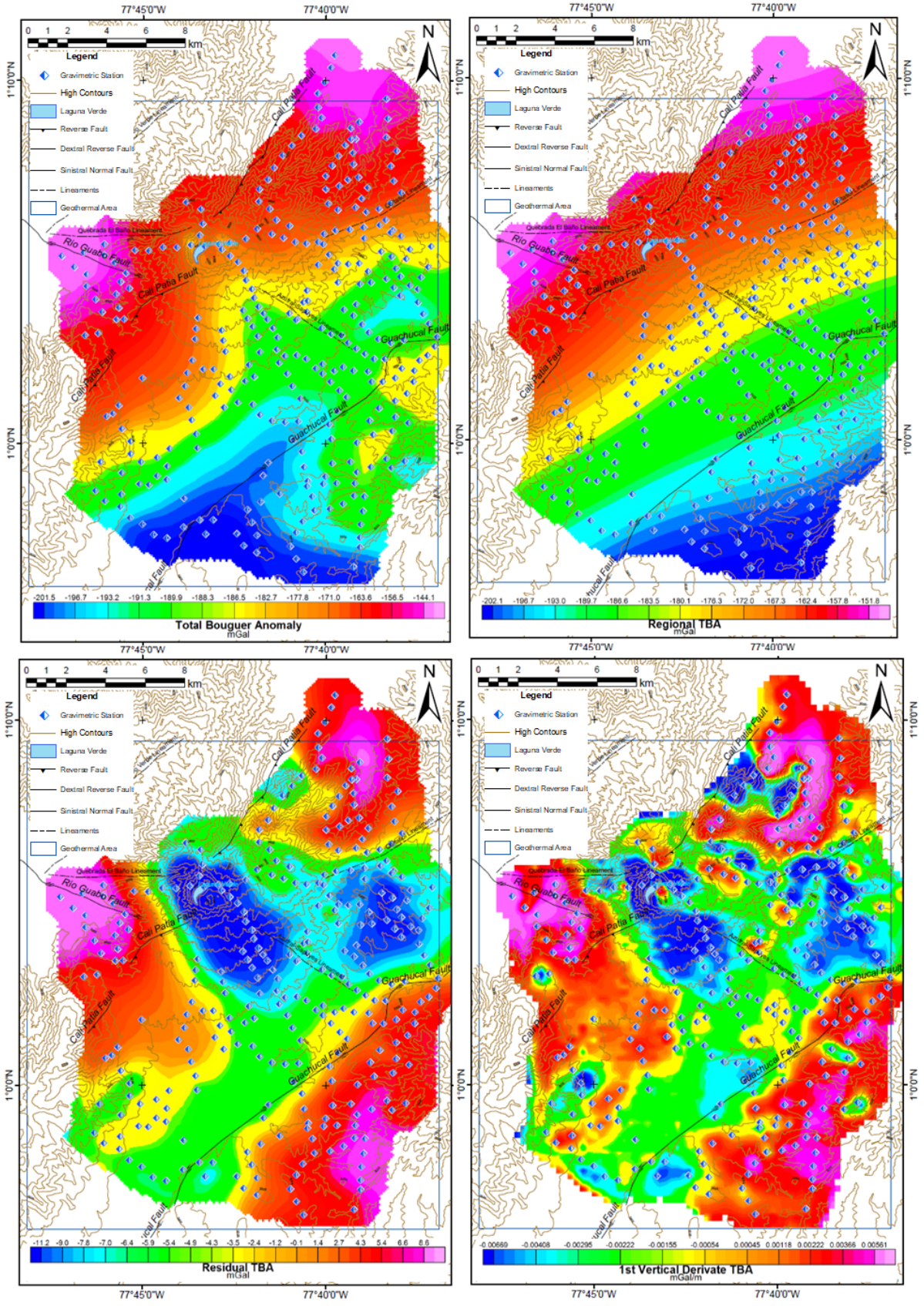

FIGURE 15: The total Bouguer anomaly (upper left), the regional anomaly (upper right), the residual anomaly (bottom left) and the first vertical derivative (bottom right) of the Azufral volcano geothermal area. Gravimetric stations are marked as blue diamonds, faults as continuous black lines, lineaments as dashed black lines, and the geothermal area is marked with a blue rectangle

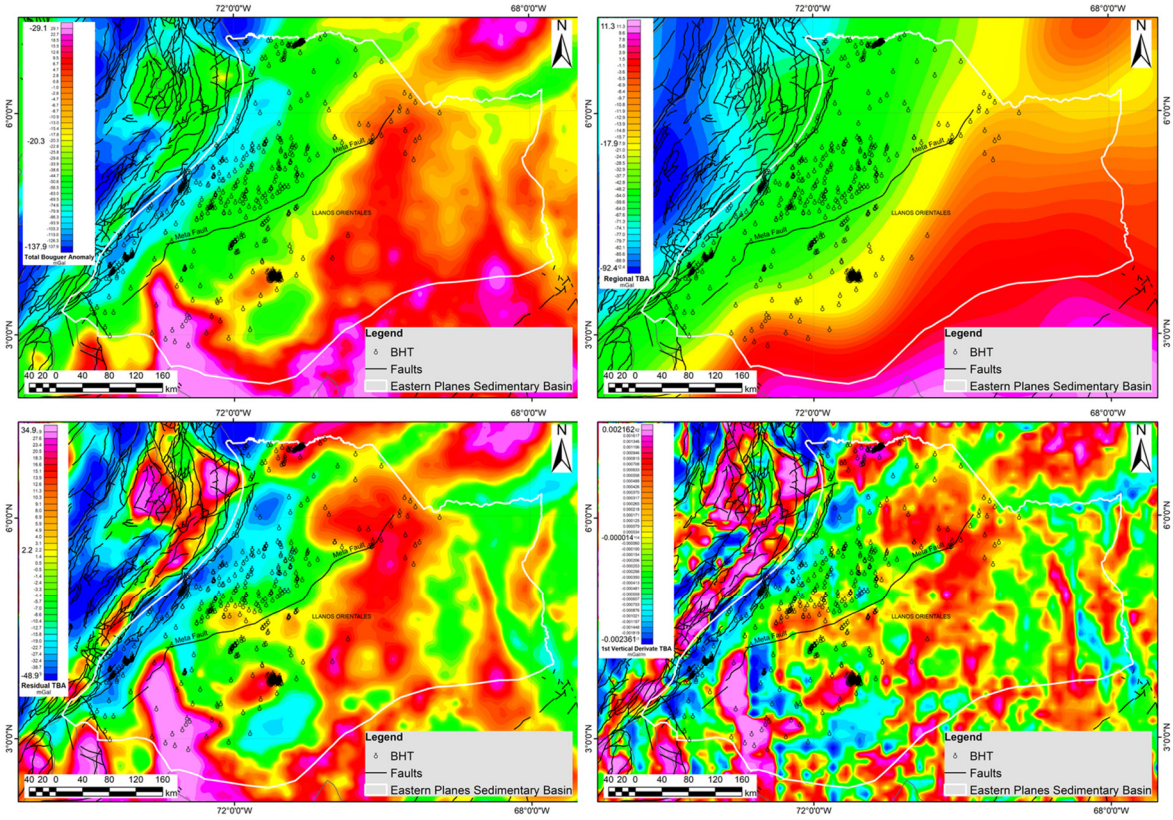

FIGURE 16: The total Bouguer anomaly (upper left), the regional anomaly (upper right), the residual anomaly (bottom left) and the first vertical derivate (bottom right) of the Eastern Llanos sedimentary basin. BHT measurements are shown as black points, faults as continuous and dashed lines, and the border of the sedimentary basin with a white line

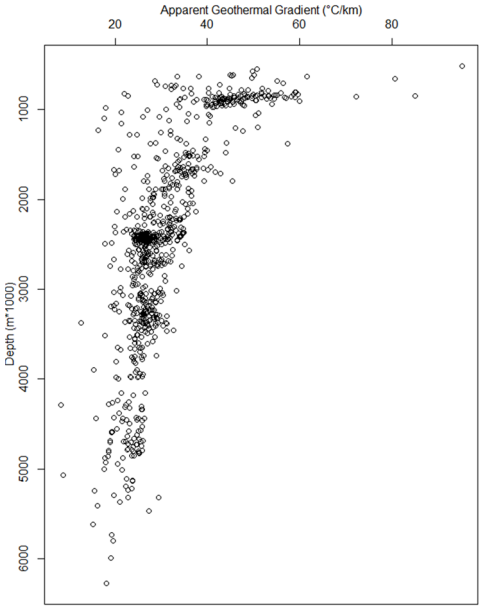

FIGURE 17: Geothermal gradients applying AAPG and terrain correction (Matiz-León, 2018)

4.2 Bottom hole temperature and geothermal gradient

The bottom hole temperature values are from oil wells. These types of wells are located in the sedimentary basins of Colombia. For this study, 861 oil wells with BHT values located within the Eastern Llanos sedimentary basin, found in the SGC Petroleum Information Bank, were used. According to Matiz-León (2018), the average BHT values were 84.5°C (between 41.1°C and 166.1 °C). Applying the AAPG correction to obtain the rock formation temperature, the average value was 92.4°C (between 42.7°C and 178.9°C). The observed geothermal gradient varies between 5.0°C/km and 57.6°C/km, at depths between 515 m and 6273 m. After applying the AAPG correction and the terrain correction, an average of 30.8°C/km (gradient range between 8.2°C/km and 95.1°C/km) was obtained (Matiz-León, 2018). Figure 17 shows the variation of the geothermal gradients after the corrections have been applied, and Figure 18 presents the graphical analysis, showing heterogeneity in the data with high gradient values but low frequency.

With the corrected geothermal gradient values, an interpolation was generated using the minimum curvature method (Figure 19).

4.3 Shallow temperature surveys

Shallow temperature surveys have been performed in all areas that are explored and under exploration by the SGC. For this work, we selected the geothermal area of the Azufral volcano with temperature values have been measured in 329 stations at a depth of 1.5 m with an average of 14.5°C (the range varies between 6.6°C and 49.1°C) (Rodríguez, 2016). The surface temperature values were generated using the relationship shown in Equation 6 and described by Eslava (1992). With the values at 1.5 m depth and the surface temperature values, the shallow geothermal gradient was calculated.

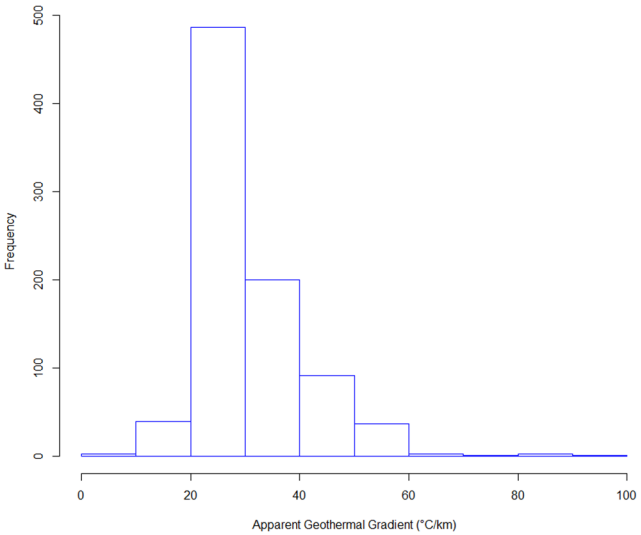

FIGURE 18: Histogram of geothermal gradients (Matiz-León, 2018)

$$T_S^O = 28.1 - 0.00553 * h \quad (6)$$

where T_S^O = Average annual air temperature (°C); and
h = Altitude in meters.

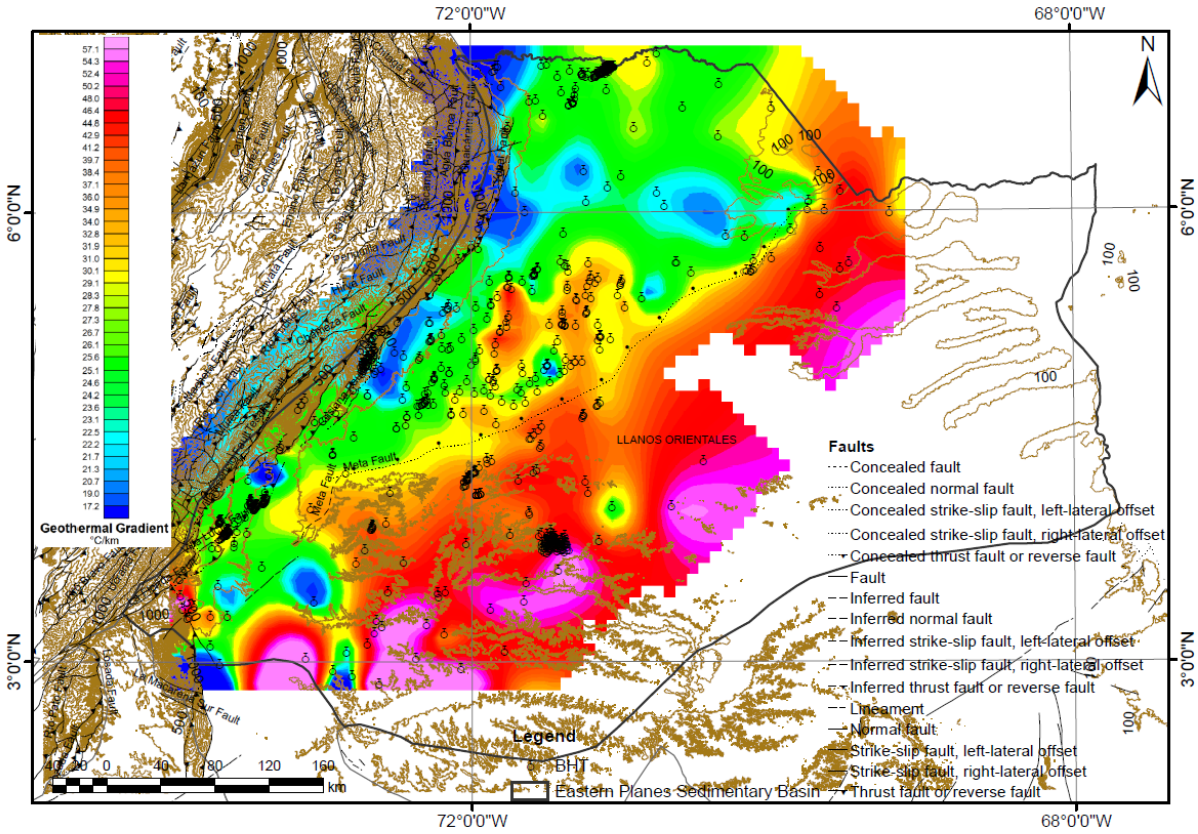

FIGURE 19: Geothermal gradient and BHT from oil wells within the Eastern Llanos sedimentary basin

The scatter plot (Figure 20) shows the presence of negative geothermal gradient values. This effect is due to the influence of the sun on shallow or near-surface temperatures. The histogram (Figure 21) shows that the most frequently occurring values correspond to gradients between -3°C/m and 0°C/m. The locations of the shallow temperature surveys and the interpolated grid for the geothermal gradient are shown in Figure 22.

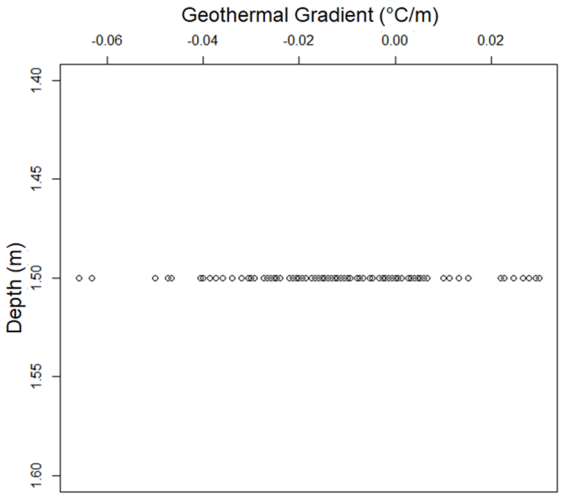

FIGURE 20: The geothermal gradient as a function of depth in the Azufral volcano geothermal area

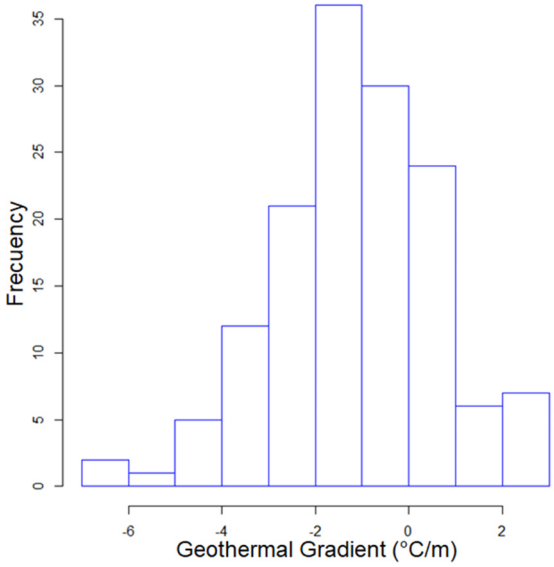

FIGURE 21: Histogram of geothermal gradients in the Azufral volcano geothermal area

5. RESULTS AND ANALYSIS

5.1 Azufral volcano geothermal area

The SVM method uses a C-svc function as a classification kernel. The classes used in the process correspond to 11 divisions with 1°C/m intervals. The classification yielded 141 support vectors for the differentiation of the executed classes. The value assigned to C corresponds to a constant of 50. This made the resulting model less flexible because the tolerance to misclassification is very limited. The sigma found was 72.5, which means that the level of non-linearity is low (a high sigma corresponds to a decision boundary with a linear trend). In turn, the training error was calculated as 0.4. Table 1 is a summary of the values of the objective function, the support vectors (alpha vectors) and the negative intercepts of the function (b).

As part of the validation of the prediction results, the confusion matrix is obtained. This allows the evaluation of thematic accuracy. The kappa coefficient is a

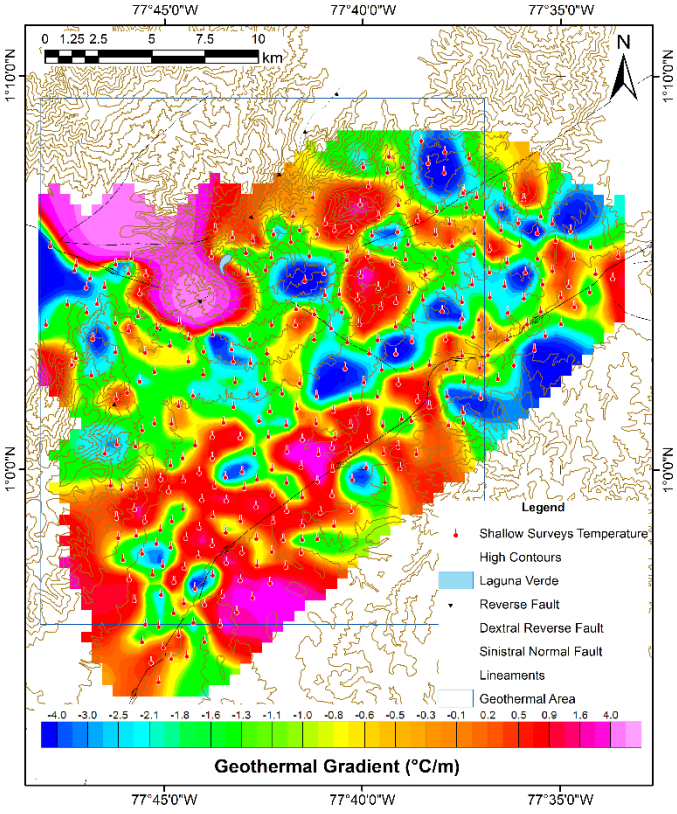

FIGURE 22: Geothermal gradient and the shallow temperature surveys (in red) at 1.5 m depth for the the Azufral volcano geothermal area

condition number of the confusion matrix that measures how much the output value of the function can change due to a small change in the input argument. It was found to be 6.2. The error of the confusion matrix could not be calculated. This may be due to the high error found in the confusion matrix. The mismatches found were calculated to be 1 which indicates only very limited correlation between the training data and the test data. Table 2 shows a summary of the confusion matrix, the kappa value, the error, and the mismatch for the classification.

TABLE 1: Summary and support vectors resulting from the SVM method for the Azufral volcano geothermal area

Summary	Support Vector Machine object of class "ksvm" SV type: C-svc (classification) parameter: cost C = 50 Gaussian Radial Basis kernel function. Hyperparameter: sigma = 72.4964705460041 Number of Support Vectors: 141 Objective Function Value: -3.1123 -5.0811 -3.8157 -5.8683 -102.6084 -81.7138 -8.9555 -5.05 -3.819 -1.3359 -1.796 -14.2877 -3.0187 -90.852 -99.9518 -33.1011 -1.762 -48.7513 -1 -105.2234 -248.4586 -201.711 -224.2671 -198.3521 -31.4615 -103.9511 -1.6662 -324.6271 -510.7573 -461.2894 -525.1617 -125.8832 -72.6408 -12.5418 -856.4306 -603.0478 -638.1021 -384.3382 -102.4616 -2.3496 -1574.239 -1107.622 -599.9886 -243.6212 -99.9232 -809.4419 -318.9438 -457.6199 -68.0812 -188.9169 -220.7084 -23.454 -7.638 -1.9574 -1.6443 Training error: 0.41958
Support (Alpha) vectors	[1] 2.5641685 3.1120912 0.5479227 [[2]] [1] 0.4632371 1.7371865 2.4458407 0.4758068 2.6351034 0.4562545 1.9484592 [[3]] [1] 0.5161361 0.3850105 0.1479770 0.3838162 2.2003102 0.3044398 1.3059469 0.1298210 0.3084642 0.3342993 1.6156009 ⋮ ⋮ [[54]] [1] 0.1208249 0.1486620 0.1500465 0.1130962 0.3137626 0.6773609 1.9572599 0.2163846 0.2171221 [[55]] [1] 0.3192991 0.3222014 0.3556091 0.2038201 0.0875884 0.3556904 1.6442085
b	[1] 0.45162168 -0.53693133 -0.61487806 -0.80935207 -1.24936111 [6] -2.18619530 -0.87019967 -0.61736174 -0.43292493 0.33396471 [11] -0.70764791 -0.86305995 -0.92455415 -1.24591349 [15] -1.04448555 -1.97528640 -0.74000540 -0.74705127 [19] 0.00000000 -0.23755140 -0.94044507 -0.96479618 [23] -0.88484260 -0.77470533 -0.14119819 -0.09117347 [27] 0.66622266 -0.14084823 -0.73613645 -0.68826143 [31] -0.80918066 0.26441427 -0.01224770 0.84498938 -0.47932475 [36] -0.08691200 0.17058389 0.73754566 0.44809566 0.90629841 [41] 0.19312760 0.33171193 1.23389182 0.76193667 1.22345930 [46] 0.09899716 0.86861533 1.26677921 4.34358215 0.58438536 [51] 0.54801577 3.02346045 0.17857924 0.78259388 0.64424985

TABLE 2: Confusion matrix and coefficients resulting from the SVM method for the Azufral volcano geothermal area

Confusion Matrix	true predicted 1 3 4 5 6 7 8 9 10 -5.51551914215088 0 0 0 0 2 0 0 0 0 -4.66540670394897 0 0 0 0 0 0 1 0 0 -4.52533531188965 0 0 0 0 1 0 0 0 0 -4.46836233139038 0 0 0 1 0 1 0 0 0 -4.16263771057129 0 0 0 0 0 0 2 0 0 -3.90017008781433 0 0 0 0 1 0 0 0 0 -3.83633875846863 0 0 0 0 0 0 1 0 0 -3.75892448425293 0 0 0 0 1 0 0 0 0 ⋮ 13.7643823623657 0 0 0 1 0 0 0 0 0 14.4611177444458 0 0 0 0 0 0 0 0 1 14.5567378997803 0 0 0 0 1 0 0 0 0 16.6410884857178 0 0 0 0 2 0 0 0 0 17.0004711151123 0 0 0 0 0 0 0 0 0 17.2435283660889 0 0 0 0 0 0 0 1 0
kappa	6.186671
error	NaN
mismatch	1

The plots for the separability analysis show that there is a concentration of classes towards the central part. It should be highlighted that the support vectors fulfil the function of discriminating for most of the proposed classes, reducing the mismatches between each one of them. Figure 23 shows the 2D separability analyses.

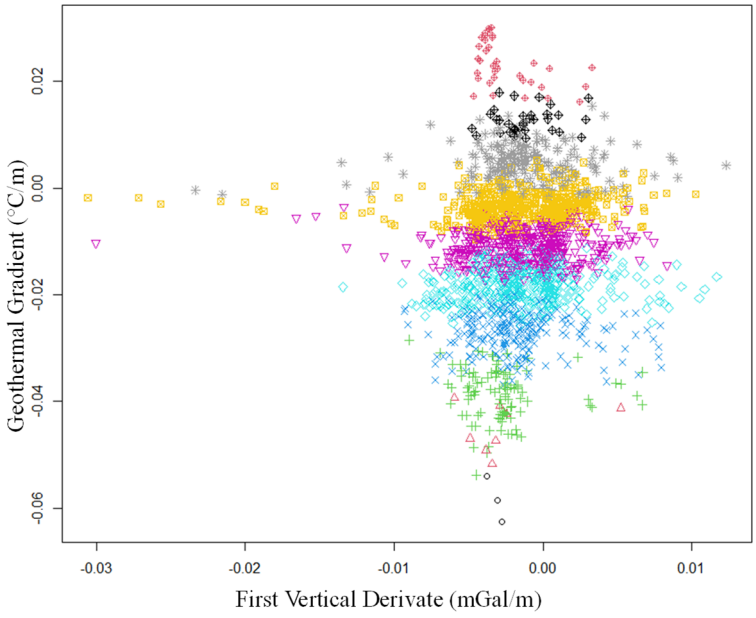

FIGURE 23: Separability analysis of the 11 proposed classes

After the classification with the training and test values, the final class map is obtained for the geothermal gradients in Azufral volcano geothermal area (Figure 24). This prediction allows the observation of a wide zone of negative geothermal gradients between -1°C/m and -0.4°C/m (pixels in yellow) between the Guachucal and Cali-Patia faults. Parallel to each fault there is a series of gradients between -2°C/m and -1.1°C/m (cyan pixels).

No pixel clusters of positive anomalies (red pixels) or negative anomalies (dark blue pixels) can be clearly identified. Also, where the hot springs are located in the geothermal area, there is no direct relation to high anomalies of the geothermal gradient. As well as the geothermal gradient map calculated by direct temperature measurements, the presence of negative gradients near the surface is most likely caused by the influence of the sun. The highest geothermal gradients reach values of approximately 2°C/m.

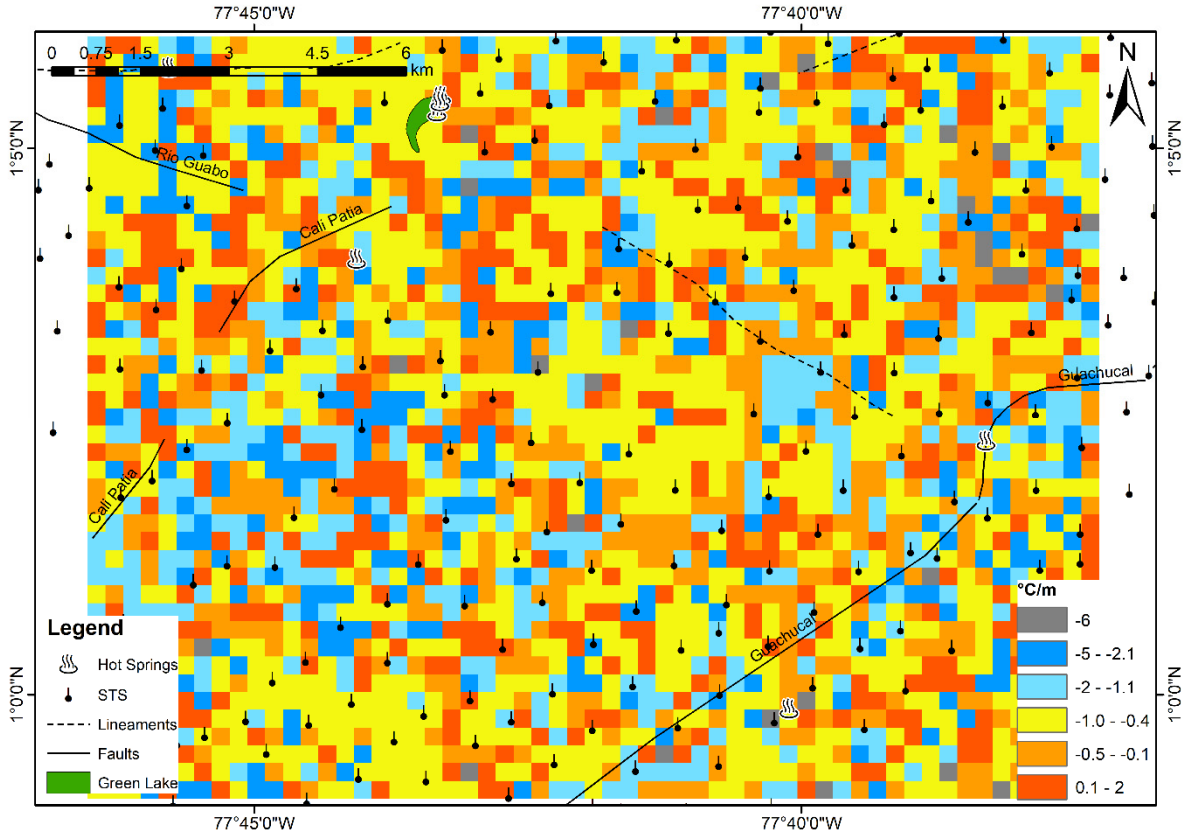

FIGURE 24: The geothermal gradient map for the Azufral volcano geothermal area using the SVM method

5.2 Eastern Llanos sedimentary basin

For the SVM method for the Eastern Llanos sedimentary basin we also used a C-svc function as a classification kernel. 53 classes segmented by 1°C/km intervals were used for the process. The value assigned to C is a constant of 50 (the resulting model is less flexible because the tolerance to misclassification is very limited). A total of 863 support vectors were calculated for the added classes. The sigma found was 48.2 which means that the level of non-linearity is low (a high sigma corresponds to a decision boundary with a linear trend). The training error was calculated as 0.7. Table 3 is a summary of the objective function values, the support vectors (alpha vectors) and the negative intercepts of the function (b).

Calculating the confusion matrix to evaluate the thematic accuracy of the classification in the Eastern Llanos sedimentary basin, a kappa value of 10.7 was obtained. This value expresses that the classification has a confidence of 10% better than the expected confidence. It was not possible to calculate the error of the confusion matrix. This could indicate that the error of the confusion matrix is high. The number of mismatches observed was calculated to be 1, indicating a significant lack of connection between the training and test data. Table 4 shows a summary of the confusion matrix, the kappa value, the error, and the mismatch for the classification.

The separability analysis plots reveal that there is a concentration of classes in the central area. It is also worth noting that the support vectors perform the task of differentiating many of the proposed classes, eliminating incompatibilities between them. For the classes present in the sedimentary basin, there are gaps between the classes with gradient values between 40°C/km and 45°C/km, and a rather wide dispersion for the geothermal gradients with values between 55°C/km and 65°C/km. The separability

evaluations in 2D are shown in Figure 25.

TABLE 3: Summary and support vectors resulting from the SVM method for the Eastern Llanos sedimentary basin

Summary	Support Vector Machine object of class "ksvm" SV type: C-svc (classification) parameter: cost C = 50 Gaussian Radial Basis kernel function. Hyperparameter: sigma = 48.1892335371568 Number of Support Vectors: 860 Objective Function Value: -1.3333 -36.5647 -2.0874 -13.2791 -3.1053 -101.7428 -101.6433 -73.5102 -61.8983 -101.7819 -190.2593 -184.5512 -157.0322 -97.3685 -186.1652 -101.7885 -42.7237 -21.2023 -34.8543 -127.4514 -7.7541 -82.336 -7.0568 -98.4749 -2 -32.3604 -5.4708 -3.1379 -42.7788 -3.0512 -32.7848 -91.3802 -40.3036 -9.0857 -52.561 -2.8221 -41.2335 -47.7182 -1.8893 -3.3455 ⋮ ⋮ -88.9596 -1.9987 -2.3976 -2.3887 -1.3327 -1.3327 -1.3327 -1.3327 -1.3365 -1.3327 -2.3995 -2.3906 -1.3333 -1.3333 -1.3333 -1.3333 -1.3333 -1.3333 -104.5834 -8.2806 -4.4603 -1.4998 -2.334 -1.4998 -49.8229 -2.2089 -4.6587 -1.4963 -2.6239 -1.4963 -47.4753 -1.0019 -1 -1.0091 -1 -1 -1 -9.2921 -1 -2.7453 -1 -1 -1 -1 -1.563 -1 Training error: 0.760743
Support (Alpha) vectors	[[1]] [1] 1.3330078 0.6660156 0.6669922 [[2]] [1] 1.2396426 35.0814831 35.3254645 0.7232259 0.7603981 [[3]] [1] 1.068524 1.068524 1.018862 1.018862 ⋮ ⋮ [[1376]] [1] 1 1 [[1377]] [1] 1.563049 1.563049 [[1378]] [1] 1 1
b	[1] 3.333333e-01 -2.396130e-01 0.000000e+00 -4.695381e-01 [5] -5.506760e-01 -7.642363e-01 -8.279955e-01 -1.720341e+00 [9] -1.528247e+00 -8.564408e-01 -1.966759e+00 -2.388141e+00 [13] -9.756241e-01 -1.004443e+00 -9.920417e-01 -8.635592e-01 [17] -1.664196e+00 -1.383007e+00 -1.366936e+00 -4.191709e+00 ⋮ ⋮ [1359] 7.385119e-01 4.963356e-01 7.226613e-01 4.963360e-01 [1363] 5.729336e-01 0.000e+00 0.000e+00 0.000e+00 0.0000e+00 [1368] 0.000e+00 0.0000e+00 0.000e+00 0.000e+00 0.000e+00 [1373] 0.000e+00 0.000e+00 0.00e+00 0.00e+00 0.000e+00 0.00e+00

TABLE 4: Confusion matrix and coefficients resulting from the SVM method for the Eastern Llanos sedimentary basin

Confusion Matrix	true predicted 5 7 8 9 10 11 12 13 14 15 16 17 19 21 22 23 25 27 29 32 34 38 41 45 50 2.56267738342285 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 12.9706001281738 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 13.7296276092529 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 14.0704288482666 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 14.6821346282959 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 15.1917514801025 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 15.5084972381592 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 17.0330390930176 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 17.076868057251 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 17.4558372497559 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 17.6955223083496 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 17.7040214538574 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 18.0231475830078 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 ⋮ ⋮ 55.336669921875 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 56.1368713378906 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 56.7421951293945 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 56.8430824279785 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 57.2803268432617 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 57.7577476501465 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 58.1677474975586 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 59.4567375183105 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 64.3648147583008 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 73.2776641845703 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
kappa	10.75905
error	NaN
mismatch	1

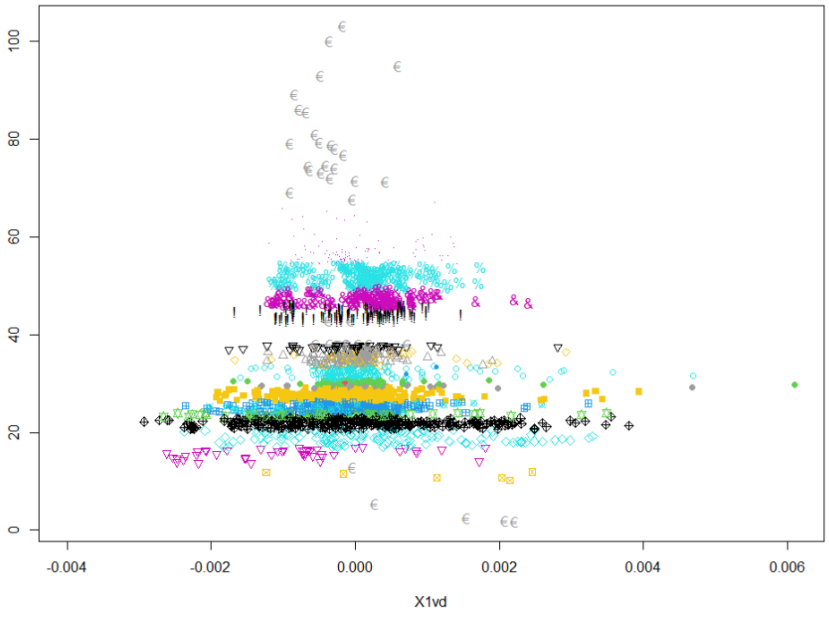

FIGURE 25: Separability analysis of the 53 proposed classes

The final SVM classification map for the Eastern Llanos sedimentary basin is shown in Figure 26. The apparent geothermal gradients were obtained with the training data and the proposed test values. The southernmost part of the sedimentary basin shows a consolidation of clusters with positive geothermal gradient anomalies (red pixels between 51°C/km and 72°C/km). It is also possible to identify a structure with N-S orientation near the eastern boundary of the sedimentary basin. For the western border (which retains a SW to NE orientation and is consistent with the location of faults present in high altitude zones outside the basin under study), trends with positive apparent geothermal gradient anomalies are observed. These positive anomalies are surrounded by high apparent geothermal gradient values (orange pixels between 38°C/km and 50°C/km). Most of the sedimentary basin has gradient values between 21°C/km and 28°C/km (dark blue and cyan pixels).

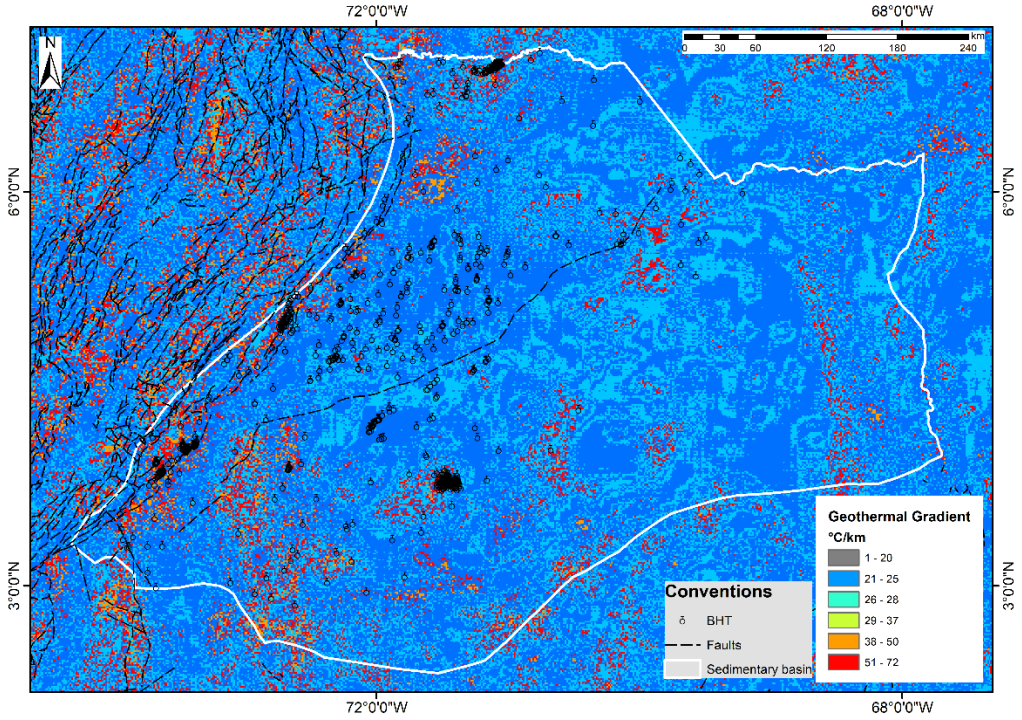

FIGURE 26: Apparent geothermal gradient map for the Eastern Llanos sedimentary basin using the SVM method

6. CONCLUSIONS

The main accomplishment of this work is the conclusion that it is possible to build a "training correlation" between geophysical data (in this case, total Bouguer anomaly and geothermal gradient) for decision makers in geothermal exploration investigations using the machine learning technique. Even if a relationship between physical properties does not exist, consideration of directly comparable relationships such as the slope of rock density in the vertical component, calculating the first vertical derivative of the residual signal of the total Bouguer anomaly, versus the vertical (depth) variation of temperature towards the Earth's surface, are highly practical. Such approaches provide decision makers at the geothermal exploration stage with better tools and results than the surface studies performed in the pre-feasibility phases.

The inclusion and consideration of multipurpose data applied to the exploration of geothermal resources greatly enriches the results in the pre-feasibility phase. Many times, entities and companies have information that is only used in a conventional way to obtain information about the areas and to develop conceptual models. However, it is important to embrace new approaches and new methodologies that

generate new knowledge for the geothermal community. The analytics found by the methods articulated in machine learning largely supplement the research, development and innovation of relevant information and can increase the precision and accuracy of surface studies.

Considering the different datasets that are acquired in each target area, it is possible to apply classification processes framed in machine learning methodologies for different geothermal environments (such as hydrothermal areas and sedimentary basins). The availability of the data and their subsequent structuring (consistency at the database level) has to be checked before included in the programming scripts.

For the geothermal area of the Azufral volcano, the indicators of the thematic accuracy of the calculated classification show poor results. When using shallow temperature surveys at 1.5 m depth, the temperatures are most likely influenced by the effect of the sun as these are shallow measurements close to the surface. The resulting negative shallow geothermal gradients greatly influence the result estimates in a negative way. A qualitative analysis of the shallow geothermal gradients calculated from direct observations (Figure 22) versus the classified gradients (Figure 24) shows that in the clusters of the computed classes, the values range between -1.0°C/m and -0.4°C/m, while in the direct interpolation the values vary between -1.8°C/m and -1.0°C/m. This denotes the discrepancy between the support vectors that generate mismatches between the classes, decreasing the geothermal gradient values.

The parameters describing the classification's thematic accuracy show unsatisfactory results for the Eastern Llanos sedimentary basin. In the southernmost part, some structures with gradient values ranging between 38°C/km and 60°C/km are observed. Similar average geothermal gradient values are found based on the conventional geothermal gradient map and the applied classification. In the qualitative comparison, a range of positive gradient anomalies (from 34°C/km to 58°C/km) in SW-NE direction are observed in the map with the conventional methodology for BHT (Figure 19). There is also a trend of mean gradients between 23.6°C/km and 30.1°C/km. For the same sectors of the classified map (Figure 26), the low gradient values (between 21°C/km and 25°C/km) are mixed with intermediate gradients and positive anomalies (between 38°C/km and 72°C/km). This denotes the lack of accuracy and positionality of the classes when making the support vectors to divide the trained classes.

The approach proposed in this research can be applied to methodologies such as Play Fairway Analysis, direct uses of low temperature, hidden systems, as well as contribute to studies that influence the generation of conceptual models of geothermal areas.

7. RECOMMENDATIONS

An access to geothermal gradient wells is necessary to avoid shallow surveys that are influenced by external factors or cultural noise.

BHT values must be replaced by temperature data in profiles, temperature logs.

When structuring the data, the technical part related to the cooling of the well and the time elapsed after drilling must be considered when estimating the temperature gradient with less uncertainty.

A quantitative validation between the classified results and the ground truth is imperative to know the true correlation between the estimated geothermal gradient values. It is recommended to use a pixel-to-pixel correlation as the Pearson coefficient between the resulting grids (Matiz-León et al, 2019a; Matiz-León et al., 2019b; Matiz-León et al., 2020).

The shallow temperature variation in the horizontal component versus the first horizontal derivative of the total Bouguer anomaly could be considered, relating the two variables in terms of the lateral flow variation that they may have.

ACKNOWLEDGEMENTS

I would like to thank GRÓ-GTP for the opportunity to attend this great experience. Especially to all the staff comprising Director Gudni Axelsson, Málfrídur Ómarsdóttir (muchas muchas gracias), Ingimar G. Haraldsson, Vigdís Hardardóttir and Markús A.G. Wilde. Thank you for believing in Colombia so that the country could become a part of the GTP family and for allowing this beautiful experience called Iceland to be possible. Thanks for the patience, effort, advice, and dedication with which you do your work so that we do not feel so far from home. Infinitas gracias!

I am deeply grateful to my supervisors Gylfi Páll Hersir and Egill Árni Gudnason for supporting and supervising my final project; for supporting new methodologies that can drive geothermal exploration from different perspectives and contribute to the achievement of reliable and accurate results, and not just stick with conventional methods. Thanks to all my lecturers, especially the lecturers in geophysics specialization: Ásdís Benediktsdóttir, Sigrídur Kristjánsdóttir, and Knútur Árnason, for their time and dedication.

Thanks to Servicio Geológico Colombiano – SGC, especially to the Technical Direction of Geosciences for providing the data for the execution of this project (NUBA). Many thanks to my boss Claudia Alfaro for motivating my passion for geothermal from afar.

A very loving thank you to my wife Andrea, for empowering me to fulfil this dream, for her love, full support, and patience. To my parents and brothers for supporting me and for the good cheers.

Thanks to all the GRÓ-GTP 2021 fellows for being part of this wonderful experience, for sharing your perspectives and points of view, which made my stay in Iceland a much more enriching experience.

Above all, to the Lord God almighty! By his work and grace, this adventure could be carried out. Thank you for guaranteeing my life, my health, and the strength to finish the training.

REFERENCES

Abedi, M., Norouzi, G.H., Bahroudi, A., 2012: Support vector machine for multi-classification of mineral prospectivity areas. *Computers & Geosciences, 46,* 272-283.

Alfaro, C., Aguirre, A., and Jaramillo, L.F., 2002: *Inventario de fuentes termales en el Parque Nacional Natural de Los Nevados*, INGEOMINAS, internal report, Bogotá D.C., 101 pp.

Alfaro, C., Alvarado, I. and Manrique, A., 2015a: Heat flow evaluation at Eastern Llanos sedimentary basin. *Proceedings of the World Geothermal Congress 2015, Melbourne, Australia*, 25 pp.

Alfaro, C., Alvarado, I., Quintero, W., Hamza, V., Vargas, C., and Briceño, L.A., 2009: Mapa preliminar de gradientes geotérmicos de Colombia. *Memorias del XII Congreso Colombiano de Geología, Colombia*, 24 pp.

Alfaro, C., Bernal, N., Ramírez, G., and Escovar, R., 2000: Colombia, Country Update. *Proceedings of the World Geothermal Congress 2000, Kyushu - Tohoku, Japan*, 10 pp.

Alfaro, C., Ponce, P., Monsalve, M.L., Ortiz, I., Franco, J.V., Ortega, A., Torres, R., and Gomez, D., 2015b: A preliminary conceptual model of Azufral geothermal system, Colombia. *Proceedings of the World Geothermal Congress 2015, Melbourne, Australia*, 10 pp.

Alfaro, C. and Rodríguez-Rodríguez, G., 2020: Status of the geothermal resources of Colombia: country update. *Proceedings of the World Geothermal Congress 2020+1, Reykjavik, Iceland*, 16 pp.

Alfaro, C., Rueda-Gutiérrez, J.B., Casallas, Y., Rodríguez, G., and Malo, J., 2021: Approach to the geothermal potential of Colombia. *Geothermics, 96*, 102-169.

Alfaro-Valero, C.M., Rueda-Gutiérrez J., Rodríguez-Rodríguez, G.F., Rodríguez-Ospina G.Z., Malo-Lazaro J., and Matiz-León J.C., in prep.: *A descriptive conceptual model of the Azufral volcano. A geothermal system in a subduction zone in Nariño, Colombia*. Servicio Geológico Colombiano, Bogotá D.C.

Bachu, S., Ramon, J.C., Villegas, M.E., and Underschultz, J.R., 1995: Geothermal regime and thermal history of the Llanos basin, Colombia. *Am. Assoc. Pet. Geol. Bull., 79,* 116–129.

Balling, N., Haenel, R., Ungemach, P., Vasseur, G., and Wheildon, J., 1981: *Preliminary guidelines for heat flow density determination*. Commission of the European Communities, technical report EUR 7360, Luxembourg, 40 pp.

Barrero, D., Pardo, A., Vargas, C.A., and Martínez, J.F., 2007: *Colombian sedimentary basins: Nomenclature, boundaries and petroleum geology, a new proposal*. Agencia Nacional de Hidrocarburos – ANH, 92 pp.

Bayona, G., Valencia, A., Mora, A., Rueda, M., Ortiz, J., and Montenegro, O., 2008: Estratigrafía y procedencia de las rocas del Mioceno en la parte distal de la cuenca antepais de los Llanos de Colombia. *Geol. Colomb., 33*, 23–46.

Bechon, F. and Monsalve, M.L., 1991: Activité récente préhistorique du volcán Azufral (S-W de la Colombie). *C.R. Acad. Sci. Paris, 313(II)*, 99-104.

Beltrán, M., 2020: Gravity and magnetic interpretation of Azufral volcano geothermal zone. *Proceedings World Geothermal Congress 2020+1, Reykjavik, Iceland*, 10 pp.

Brown, S., Coolbaugh, M., DeAngelo, J., Faulds, J., Fehler, M., Gu, C., Queen, J., Treitel, S., Smith, C., and Mlawsky, E., 2020: Machine learning for natural resource assessment: An application to the blind geothermal systems of Nevada. *Geothermal Resources Council Transactions, 44*, 14 pp.

Bullard, E.C., 1938: The disturbance of the temperature gradient in the earth's crust by inequalities of the height. *Geophys. J. Int, 4*, 360–362.

Calvache, M.L., 1999: *Informe volcán Azufral. Proyectos B98G04 Y C98R06*. Unidad Operativa Pasto, INGEOMINAS, Pasto, internal report, 44 pp.

Calvache, M.L, Cortés, G.P, Torres, M.P, Monsalve, M.L., and Cepeda, H., 2003: *Mapa de cartografía geológica del volcán Azufral*. Memoria explicativa, INGEOMINAS, Bogotá D.C.

Carvalho, H.D.S. and Vacquier, V., 1977: Method for determining terrestrial heat flow in oil fields. *Geophysics, 42*, 584–593.

Cepeda, H., 1985: Anotaciones acerca de la geología del volcán Galeras (Colombia). *Memorias del VI Congreso Latinoamericano de* geología, Tomo I, Bogotá D.C.

CHEC, 1968: *Proyecto de investigación geotérmica en la región del Macizo Volcánico del Ruiz*. Central Hidroeléctrica de Caldas (CHEC) and Ente Nazionale per L'Energia Elettrica (ENEL), technical report, 41 pp.

Clauser, C., 2009: Heat transport processes in the earth's crust. *Surv. Geophys, 30,* 163–191.

Cooper, M.A., Addison, F.T., and Alvarez, R., 1995: Basin development and tectonic history of the Llanos basin, Colombia, *Pet. basins South Am*., 659–665.

Cortés, G.P. and Calvache, M.L., 1997: *Geología del volcán Azufral primera fase*. INGEOMINAS, internal report, 31 pp.

Cortés, G.P., Calvache, M.L., and Murcia H.F., 2009: *Mapa de amenaza del volcán Azufral, Colombia*, INGEOMINAS, Manizales, internal report, 27 pp.

Deming, D., 1989: Application of bottom-hole temperature corrections in geothermal studies. *Geothermics, 18*, 775–786.

Dentith, M. and Mudge, S., 2014: *Geophysics for the mineral exploration geoscientist*. Cambridge University Press, New York, 516 pp.

Dueñas Jiménez, H., and van der Hammen, T., 2007: Significado geológico y asociaciones palinológicas de las Formaciones Diablo Inferior (Mioceno Tardío) y San Fernando Superior (Mioceno Medio), Piedemonte Cuenca de los Llanos Orientales, Colombia, *Revista de la Academia Colombiana de Ciencias, 31*, 481–4493.

Eslava, J., 1992: Altitudinal profile of the average air temperature in Colombia. *Geophysics Colombian, 1*, 37-52.

Faulds, J., Brown, S., Coolbaugh, M., DeAngelo, J., Queen, J., Treitel, S., Fehler, M., Mlawsky, E., Glen, J., Lindsey, C., Burns, E., Smith, C., Gu, C., and Ayling B., 2020: Preliminary report on applications of machine learning techniques to the Nevada geothermal play fairway analysis. *Proceedings 45th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California,* 6 pp.

Franco, J., 2016: *Actualización geoeléctrica en el área geotérmica de Paipa – Boyacá. Servicio Geológico Colombiano.* Internal report, Medellín, 27 pp.

Franco, C., Martinez, D., Gutierrez, M., Pataquiva, J., Rojas, J., Jaramillo, D., Foo1, G., Cespedes, S., and Cortes, F., 2021: Oilfield application of co-produced fluid geothermal power in Colombia's Llanos Orientales basin. *First EAGE Workshop on Geothermal Energy in Latin America*, 5 pp.

Fontaine, E., 1994: *Évolution volcanologique et géochimique du volcan Azufral, Colombie, Amérique du Sud*. M.Sc. thesis, Université de Montreal, 215 pp.

García González, M., Mier Umaña, R., Cruz Guevara, L.E., and Vásquez, M., 2009: *Informe ejecutivo: Evaluación del potencial hidrocarburífero de las cuencas colombianas*. Agencia Nacional de Hidrocarburos – ANH, internal report, 14 pp.

Gómez, J., Nivia, Á., Montes, N.E., Jiménez, D.M., Tejada, M.L., Sepúlveda, M.J., Osorio, J.A., Gaona, T., Diederix, H., Uribe, H., and Mora, M., 2007: *Mapa geológico de Colombia 2007*. Escala 1:1'000.000, compilación, INGEOMINAS, Bogotá D.C., 2 pp.

Gómez. D.M. and Ponce, A.P., 2009: *Informe de visita técnica al volcán Azufral – Laguna verde, departamento de Nariño*. INGEOMINAS, OVP, internal report, 37 pp.

González, H., Zapata, G., and Montoya, D.M., 2002: *Geología y geomorfología de la Plancha 428 Túquerres. Departamento de Nariño*. Memoria explicativa, INGEOMINAS, Medellín, 88 pp.

González–Idárraga, C.E. and Rodríguez–Rodríguez, G.F., 2017: *Modelo resistivo del área geotérmica de Paipa a partir de datos magnetotelúricos*. Servicio Geológico Colombiano, internal report, Bogotá D.C., 96 pp.

Graterol Graterol, V.R. and Vargas Gomez, A., 2010: *Mapa de anomalía de Bouguer total de la República de Colombia.* Agencia Nacional de Hidrocarburos – ANH, technical report, Bogotá D.C., Colombia, 9 pp.

Gudnason, E.Á., Arnaldsson, A., Axelsson, G., Berthet, J-C.C., Halldórsdóttir, S., and Magnússon, I.Th., 2015: Analysis and modelling of gravity changes in the Reykjanes geothermal system in Iceland, during 2004-2010. *Proceedings of the World Geothermal Congress 2015*, *Melbourne, Australia*, 8 pp.

Hersir, G.P. and Björnsson, A., 1991: Geophysical exploration for geothermal resources. Principles and applications. Report 15 in: *Geothermal Training in Iceland 1991*. United Nations University Geothermal Training Programme, Reykjavík, Iceland, 94 pp.

Hersir, G.P., Gudnason, E.Á., and Flóvenz, Ó.G., 2022: *Geophysical exploration techniques*. In: Letcher, T. (ed.), Comprehensive Renewable Energy – 2nd edition, Vol 7, Elsevier, Oxford, 54 pp.

Hinze, W., von Frese, R.B., and Saad, A.H, 2013: *Gravity and magnetic exploration. Principles, practices, and applications*. Cambridge University Press, New York, 525 pp.

INGEOMINAS, 2007: *Geological map of Colombia.2:8'000.000. Second edition*. Bogotá. Website: *http://www.sgc.gov.co/Geologia/Mapa-geologico-de-Colombia.aspx*

INGEOMINAS-ANH, 2008: *Mapa preliminar de gradientes geotérmicos (Método BHT) 1:1´500.000*. Instituto de Geología y Minas y Agencia Nacional de Hidrocarburos, executive report, Bogotá D.C.

Inguaggiato, S., Londoño, J.M., Chacón, Z., Liotta, M., Gil, E., and Alzate, D., 2017: The hydrothermal system of Cerro Machín volcano (Colombia): New magmatic signals observed during 2011–2013. *Chemical Geology, 460*, 60-68.

Jacoby, W. and Smilde, P., 2009: *Gravity interpretation. Fundamentals and application of gravity inversion and geological interpretation*. Springer-Verlag Berlin Heidelberg, Berlin, 413 pp.

Kearey, P., Brooks, M., and Hill, I., 2002: *An introduction to geophysical exploration (3rd edition)*. Blackwell Science Ltd, Oxford, 281 pp.

Kellogg, J., Ojeda, G., Duque, H., and Cerón, J., 2005: Crustal structure of the Eastern Cordillera, Colombia. *6th Int. Symp. Andean Geodyn. – ISAG*, 424–427.

Llanos, E., Bonet, C., and Zengerer, M., 2015: *3D geological – geophysical model building and forward and inverse modeling of magnetism and gravimetry data from Paipa Geothermal Area, Colombia – Final report*. Contrato No. 363 de 2015 between Servicio Geológico Colombiano e Intrepid Geophysics, Melbourne, Australia, 106 pp.

Lösing, M. and Ebbing, J., 2021: Predicting geothermal heat flow in Antarctica with a machine learning approach. *Journal of geophysical research: Solid Earth*, *126(6),* 16 pp.

Lowrie, W., 2007: *Fundamentals of geophysics (second edition)*. Cambridge University Press, New York, 393 pp.

Lucazeau, F., 2019: Analysis and mapping of an updated terrestrial heat flow data set. *Geochemistry, Geophysics, Geosystems, 20(8)*, 4001-4024.

Malo, J. and Alfaro, C., 2018: *Emisiones de gas radón en aire del suelo. Área Geotérmica del Volcán Azufral.* Servicio Geológico Colombiano, internal report, Bogotá D.C., 36 pp.

Malo, J. and Alfaro, C., 2020: Local meteoric water line of the central zone of Boyacá, Colombia. *Proceedings of the World Geothermal Congress 2020+1, Reykjavik, Iceland*, 9 pp.

Matiz-León, J.C., 2018: *Metodología para determinar el modelo espacial del gradiente geotérmico en las cuencas sedimentarias del Valle Medio del Magdalena, Cordillera Oriental y Llanos Orientales en Colombia*. Universidad Distrital Francisco José de Caldas, Master degree thesis, Bogotá D.C., Colombia, 120 pp.

Matiz-León, J.C., Rodríguez-Rodríguez, G., and Alfaro-Valero, C., 2019a: Modelos de temperatura del suelo a partir de sondeos superficiales y sensores remotos para el área geotérmica de Paipa, Boyacá-Colombia. *Boletín de Geología, 41(2)*, 71-88.

Matiz-León, J.C., Rodríguez-Rodríguez, G., and Alfaro-Valero, C., 2019b: Modelos de temperatura del suelo a partir de sondeos superficiales de temperatura y sensores remotos para el área geotérmica del volcán Azufral. *Ciencia e Ingeniería Neogranadina, 29(1)*, 19-36.

Matiz-León, J.C., Rodríguez-Rodríguez, G., and Alfaro-Valero, C., 2020: Application of digital image processing techniques of earth observation satellites in the estimation of surface temperature models in support of geothermal exploration in Colombia. *Proceedings of the World Geothermal Congress 2020+1, Reykjavik, Iceland,* 10 pp.

Meqbel, N., 2017: *3D inversion results of Azufral MT data set, Colombia.* Internal report. Berlin, Germany, 12 pp.

Mohamed, H., Abdel Zaher, M., and Senosy, M., D., 2014: Geothermal gradients in the North Western desert, Egypt as deduced from bottom-hole temperature and aerogravity data. *Pure and Applied Geophysics, 172,* 1585–1597.

Mohamed, H., Zaher, M., Senosy, M., Saibi, H., Nouby, M., and Fairhead, D., 2015: Correlation of aerogravity and BHT data to develop a geothermal gradient map of the Northern Western desert of Egypt using an Artificial Neural Network. *Proceedings of the 2015 World Geothermal Congress, Melbourne, Australia*, 4 pp.

Monsalve, M.L., 2013: *Reconocimiento Vulcanológico Área Geotérmica de San Diego*. Internal report, Servicio Geológico Colombiano, Bogotá D.C., 35 pp.

Monsalve, M.L., Rodríguez, G., Méndez, R., and Bernal, N., 1998: Geology of the Well Nereidas 1, Nevado Del Ruiz Volcano, Colombia. *Transactions Geothermal Resources Council, 22,* 263-267.

Mountrakis, G., Im, J., and Ogole, C., 2011: Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 66,* 247–259.

Ortiz, A., 2017: Caso exitoso en Colombia del aprovechamiento de la geotermia de baja entalpía para climatización. *Memorias de la Segunda Reunión Nacional de Geotermia (RENAG), Manizales, Colombia*, 22 pp.

Pollack, H.N., Hurter, S.J., and Johnson, R., 1993: Heat flow from the earth´s interior: Analysis of the global data set. *Rev. Geophys, 31*, 267–280.

Pinilla, A., Ríos, P., Rodríguez, B., Roa, H., and Ladino, F., 2007: *Memoria explicativa de la cartografía geológica y geomorfológica del Altiplano Nariñense*. INGEOMINAS, Bogotá D.C., technical report, 121 pp.

Rezvanbehbahani, S., Stearns, L.A., Kadivar, A., Walker, J.D., and van der Veen, C.J., 2017: Predicting the geothermal heat flux in Greenland: A machine learning approach. *Geophysical Research Letters*, *44*, 12271-12279.

Rodríguez, G.F., 2016: *Análisis de la distribución de calor en el área geotérmica del volcán Azufral a partir de sondeos superficiales de temperatura*. Internal report, Bogotá D.C., 80 pp.

Rodríguez, G., 2020: *Avances en las perforaciones de gradiente térmico y conocimiento geológico en el Área Geotérmica de Paipa.* Geoflash Servicio Geológico Colombiano – SGC, Grupo Investigación Geotérmica, Bogotá D.C. Website: *https://bit.ly/3mxaSH1*

Rodríguez, G. and Alfaro, C., 2015: *Caracterización de núcleos de perforación en las zonas de El Durazno, Paipa y Criptodomo de Iza.* Servicio Geológico Colombiano, internal report, Bogotá D.C., 57 pp.

Rodríguez, G. and Rueda, J., 2017: *Geología estructural del área geotérmica del Volcán Azufral*. Technical report, Bogotá D.C., 93 pp.

Rodríguez-Rodríguez, G., 2020: Magnetotelluric models for the characterization of the geothermal system of Azufral volcano, Colombia. *Proceedings of the World Geothermal Congress 2020+1, Reykjavik, Iceland*, 15 pp.

Rueda, J., 2017: *Cartografía de los cuerpos dómicos del área geotérmica de Paipa.* Servicio Geológico Colombiano, internal report, Bogotá D.C., 25 pp.

Seequent, 2020: *Upward continuation (CNUP)*.

Simeone, O., 2018: A very brief introduction to machine learning with applications to communication systems. *IEEE Transactions on Cognitive Communications and Networking*, *4,* 648-664.

Siripunvaraporn, W., 2016: *3D MT inversión results from Paipa volcanic area.* Contrato asesoría SGC – Mahidol University, Mahidol, Bangkok, 182 p.

Smith, C., Faulds, J., Brown, S., Coolbaugh, M., Lindsey, C., Treitel, S., Ayling, B., Fehler, M., Gu, C., and Mlawsky, E., 2021: Characterizing signatures of geothermal exploration data with machine learning techniques: An application to the Nevada play fairway analysis. *Proceedings 46th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California*, 13 pp.

Sretenović, A., Jovanović, R., Novaković, V., Nord, N., and Živković, B.D., 2018: Support vector machine for the prediction of heating energy use. *Thermal Science, 22,* 4 pp.

Taboada, A., Dimaté, C., and Fuenzalida, A., 1998: Sismotectónica de Colombia: Deformación continental activa y subducción. *Física de la Tierra, 10*, 111-147.

Torres, M.P., Cortes, G.P., Calvache, M.L., and Monsalve, M.L., 2001: *Geología y estratigrafía del Volcán Azufral*. INGEOMINAS, internal report, 11 pp.

Vapnik, V.N.,1995: *The nature of statistical learning theory.* Springer-Verlag, New York.

Velandia, F., Romero, D., Rodríguez, B., Hincapié, G., and Salazar, E., 2006: Esquema estructural para exploración hidrogeológica en el Altiplano Nariñense. *Memoria del Congreso Colombiano de Hidrogeología. Bucaramanga, Santander, Colombia,* 11 pp.