

MINISTERIO DE MEDIO AMBIENTE Y
RECURSOS NATURALES

DIRECCIÓN GENERAL DEL OBSERVATORIO AMBIENTAL

GERENCIA DE GEOLOGÍA
ÁREA DE SISMOLOGÍA

Generación de mapa de velocidades
promedio a 30 metros de profundidad (Vs30)
de El Salvador”

Luis Ernesto Mixco Durán

Especialista I / Sismólogo

agosto de 2020

TABLA DE CONTENIDO

1.	INTRODUCCIÓN.....	4
2.	AMPLIFICACIÓN DEL SITIO.....	6
3.	CATEGORÍAS DE RESPUESTA SÍSMICA SIMILAR EN LA AMPLIFICACIÓN DE SITIO.....	7
4.	INFORMACIÓN PARA BASE DE DATOS.....	9
4.1.	ENSAYOS DOWNHOLE DE 15 PERFORACIONES DE PROYECTOS MARN	10
4.2.	ENSAYOS DOWNHOLE OBTENIDOS DE OTROS PROYECTOS.....	12
4.3.	PERFILES GEOLÓGICOS DE POZOS PLATAFORMA SIHI-MARN	14
4.4.	PERFILES GEOLÓGICOS DE AFLORAMIENTOS.	15
4.5.	PERFILES DE VELOCIDAD A PARTIR DE ENSAYOS DE MICROTREPIDACIÓN.....	16
4.6.	ENSAYOS DE PENETRACIÓN ESTÁNDAR (SPT)	19
4.7.	OTRAS FUENTES DE POZOS DE PERFORACIÓN Y ENSAYOS GEOFÍSICOS	19
5.	ESTIMACIÓN DE VS PARA PERFILES GEOLÓGICOS	22
5.1.	CLASIFICACIÓN DE IGNACIO GARCÍA, 2016	22
5.2.	CLASIFICACIÓN DE CELINA KATTAN, MARN 2011	25
5.3.	RECLASIFICACIÓN DE ESPESORES Y ASIGNACIÓN DE VS.....	26
5.3.1.	AFLORAMIENTOS DOLORS FERRÉS	27
5.3.2.	PLATAFORMA DE SISTEMA DE INFORMACIÓN HÍDRICA (SIHI)	32
5.3.2.1.	ESTIMACIÓN DE LA TS EN EL VALLE DE ZAPOTITÁN.....	34
5.3.2.2.	REALIZACIÓN DE CORTES GEOLÓGICOS EN EL AMSS	35
5.3.2.3.	ESTUDIOS PARA LA ESTIMACIÓN DE LA PROF. DE LA TS EN LA ZONA DE NEJAPA 47	
5.3.2.4.	ESTUDIOS PARA LA ESTIMACIÓN DE LA PROF. DE LA TS EN LA ZONA DE SS.....	55
5.3.2.5.	MAPA DE ISOLÍNEAS DE PROFUNDIDADES DE LA TS EN LA ZONA DEL AMSS.....	64
6.	CONSOLIDACIÓN DE BASE DE DATOS.....	66
6.1.	DETERMINACIÓN DE VS30 Y PERÍODO DOMINANTE	68
6.2.	CORRELACIÓN DE GEOLOGÍA CON VELOCIDADES DE CORTE	71
6.3.	INTEGRACIÓN DE DATOS DE USGS A BASE DE DATOS	74

7.	GENERACIÓN DE MAPA INTEGRADO DE VS30	75
7.1.	MODELOS DE APRENDIZAJE SUPERVISADO	76
7.1.1.	RANDOM FOREST CON PREDICCIÓN ESPACIAL	77
7.1.1.1.	TEORÍA SOBRE RANDOM FOREST	77
7.1.1.2.	ALGORITMO DE RF (TOMADO DE GARCÍA RUIZ, M. 2018)	79
7.1.1.3.	EVALUACIÓN DE DESEMPEÑO DE UN MODELO EN RF	80
7.1.1.3.1.	OOB OUT OF BAG ERROR (OOB)	80
7.1.1.3.2.	VALIDACIÓN CRUZADA	81
7.1.1.3.3.	IMPORTANCIA DE LAS VARIABLES	83
7.1.1.4.	CONCEPTOS ADICIONALES EN RF	83
7.1.1.5.	CONSIDERACIÓN DE AUTO-CORRELACIÓN ESPACIAL	85
7.1.1.6.	APLICACIÓN DE RFSP PARA GENERAR MAPA DE VS30	85
7.1.1.6.1.	ASIGNACIÓN DE PESOS	86
7.1.1.6.2.	SELECCIÓN DE COVARIABLES	87
7.1.1.6.3.	ALGORITMO DE RFSP PARA GENERAR PREDICCIONES	89
7.1.1.6.3.1.	PREDICCIONES CON EL MODELO FINAL	103
7.1.2.	TÉCNICA DE REDES DE NEURONAS ARTIFICIALES (RNA)	110
7.1.2.1.	RNA PROBLEMA COMO VARIABLE CATEGORÍCA	110
7.1.2.1.1.	MODELO 1: TIPO DE SUELO CD	111
7.1.2.1.2.	MODELO 2: TIPO DE SUELO BC	112
7.1.2.1.3.	MODELO 3: TIPO DE SUELO D	113
7.1.2.1.4.	MODELO 4: TIPO DE SUELO C	114
7.1.2.1.5.	ALGORITMO PARA SELECCIÓN DE UMBRAL	114
7.1.2.1.6.	ENSAMBLE DE SUB-MODELOS DE RNA	115
7.1.2.1.7.	ENSAMBLAJE DE RESULTADOS DE MODELOS (RANDOMFOREST SP & RNA)	119
8.	CONCLUSIONES, RECOMENDACIONES Y COMENTARIOS FINALES	122
9.	REFERENCIAS	125

1. Introducción

Este documento abarca la metodología para la generación de un mapa de velocidades de corte promedio a 30 metros de profundidad de El Salvador, el cual se asocia a una estimación robusta de los efectos de amplificación del suelo, convirtiéndose en una herramienta importante para propósitos de clasificación de tipos de suelo a nivel nacional. Se pretende que este mapa funcione como un diferenciador geográfico de tipos de suelo y que de manera general se logre agrupar geoméricamente aquellos estratos de suelo que podrían tener un comportamiento homogéneo frente al paso y propagación de ondas sísmicas a lo largo de sus subunidades.

Este trabajo nace como parte de las acciones para la actualización de la normativa nacional para diseño y construcción sismorresistente y plan nacional para la reducción de la vulnerabilidad sísmica, que está llevando a cabo el Ministerio de Medio Ambiente y Recursos Naturales de El Salvador (MARN) junto al Ministerio de Obras Públicas (MOP), la Universidad Centroamericana José Simeón Cañas (UCA) y como agencia ejecutora: el Banco Interamericano de Desarrollo (BID), dentro del marco del Proyecto de Cooperación: “Plan de Acción para la implementación del índice de gobernabilidad y política pública para la gestión de riesgo de desastre”.

En primer lugar, se inició un proceso de búsqueda y compendio de información relacionada principalmente en valores de velocidad de corte medidos directamente en pozos de perforación, acompañada de información geológica en profundidad, estudios de microtrepidación y otros estudios de investigación afines a este mismo tema. Dentro de la información recopilada, fue evidente la calidad de la misma, la cual se tomó en consideración al momento de ser utilizada. Dado que hasta la fecha ha habido muy poca investigación, relacionada a la caracterización dinámica de los suelos del país, se obtuvo un reducido número de ensayos geofísicos de medición directa de Vs (llamados Downholes).

Para subsanar esta escasez de información en el espacio, se recopilaron una serie de perfiles de pozos geológicos con descripciones generales de sus unidades, los cuales se encuentran contenidos dentro de la plataforma SIHI desarrollada por el MARN. Cabe mencionar, que la información de dicha plataforma se comprende mayoritariamente con información recabada por la institución ANDA; la cual es principalmente obtenida para contar con pozos de monitoreo para medir niveles de agua y toma de muestras para determinación de los componentes químicos para verificar la calidad del agua.

De acuerdo con la información obtenida, se planteó una metodología para poder estimar fronteras y profundidades límites de determinadas unidades geológicas, las cuales, mediante el apoyo de geólogos y abundante literatura, se pudieron trazar y lograr estimar la velocidad de corte de dichas unidades, logrando construir un perfil de velocidad que ayudaría a calcular el valor de VS30. Aunque, los valores obtenidos no

corresponden a una medida directa de la velocidad de corte para cada unidad litológica, y en ocasiones una misma unidad puede tener diferentes velocidades, dependiendo de variaciones en su grado de compactación, densidad y dureza, se hizo uso del detalle de las descripciones litológicas para realizar una reclasificación cualitativa de las unidades a las cuales se les asocia un valor de velocidad de corte. Adicionalmente, se integró información de perforaciones realizadas por el Ministerio de Obras Públicas (MOP), la Oficina de Planificación del Área Metropolitana de San Salvador (OPAMSS), empresas privadas y la Comisión Ejecutiva Hidroeléctrica del Río Lempa (CEL). Asimismo, se cuenta con información litológica en superficie (afloramientos) en algunos sectores a nivel nacional suministrada por investigadores geólogos que han trabajado para la caracterización vulcanológica del país.

Dentro de la metodología, se incluye la consolidación de los datos recopilados; esto se realizó utilizando una hoja de cálculo de Microsoft Excel, donde se iban ploteando los perfiles de velocidad y demás propiedades, tales como el Coeficiente de Poisson, densidad, Modulo de Young, entre otros. A partir de esta información se realiza el cálculo del valor de Vs30, se extrae la estimación donde se considera Roca (donde será aplicado un posible terremoto para propagación en los posteriores modelos analíticos unidimensionales), Periodos predominantes empíricos y asignación de tipo de suelo según NEHRP (*National Earthquake Hazards Reduction Program*). Finalmente, se elaboró una hoja de Excel resumen con los valores de VS30 calculados, su fuente bibliográfica y demás propiedades.

En total, se recopilaron 519 puntos a nivel nacional, los cuales se concentran principalmente en el Área Metropolitana de San Salvador (AMSS). Debido a la aún limitada cantidad de información en otros lugares del país, se optó por agregar algunos valores de VS30 calculados a partir de las pendientes del terreno, correspondiente al producto del USGS, específicamente para valores superiores e iguales a 640 m/s; siendo cuidadoso en los lugares, como los grábenes, donde se sabe que no existen afloramientos rocosos. Esto se hizo seleccionando espacialmente una muestra aleatoria de 100 puntos, para enriquecer la información a ser tratada con las técnicas de *machine learning* y garantizar una distribución de los datos.

Finalmente, se procede a la generación de un modelo estadístico espacial que represente de manera óptima un mapa, que registre los valores de Vs30 a nivel nacional, el cual será traducido de acuerdo a los rangos de velocidades establecidos en la clasificación de tipos de suelos según la NEHRP 2020. Estos modelos se basarán en la utilización de herramientas de aprendizaje supervisado (*machine learning*) manejando como variable objetivo el valor numérico de Vs30 (problema de regresión), o bien considerando variables categóricas como las clases de tipo de suelo, según la NEHRP 2020 (problema de clasificación). Se aplicarán las técnicas de los algoritmos de Random Forest con Predicción espacial (*Rfsp*) y la Técnica de redes de neuronas artificiales.

Con los resultados de ambas técnicas, se obtendrán dos mapas de Vs30, de los cuales, posteriormente, se superpondrán entre sí, logrando construir un Mapa Integrado de Vs30, que capture las bondades de ambos algoritmos; mejorando así la calidad de las zonificaciones representadas, asociándose a lugares con respuesta sísmica homogénea.

2. Amplificación del sitio

Los terremotos son considerados como uno de los fenómenos naturales que producen mayor destrucción, con respecto a pérdidas humanas y materiales. Su relación directa se asocia comúnmente con daños en componentes estructurales de viviendas y edificios que albergan a las personas; asimismo, la distribución geográfica de estos daños parece correlacionarse directamente con el grado de amplificación causada por las capas sedimentarias superiores del subsuelo, y que afectan en la respuesta estructural de diferentes tipologías o configuraciones constructivas a lo largo del país.

Por lo tanto, es menester caracterizar los efectos de amplificación local del suelo donde subyacen las obras ingenieriles; así como también, en lugares donde se emplazan estaciones de movimiento fuerte, las cuales permiten obtener series de tiempo-historia reales de la aceleración que sufre el terreno frente a un terremoto, permitiendo utilizarlos para futuros análisis dinámicos de estructuras.

La amplificación del sitio posee una estrecha relación con la topografía y la geología de la estratigrafía superficial (principalmente en unidades blandas y poco consolidadas), haciendo que la señal sísmica se modifique tanto en su duración, contenido frecuencial y, por supuesto, su amplitud. Respecto a la variabilidad de la duración, las discontinuidades laterales que presentan las columnas del subsuelo y la difracción de las ondas superficiales propiamente en los estratos de depósitos poco consolidados, representan un componente principal. Por otra parte, un factor clave para explicar la fluctuación del contenido frecuencial de las señales, es el contraste de impedancias que existe entre capas adyacentes (producto de la velocidad de ondas de corte por la densidad de la capa entre los materiales superficiales y el estrato rocoso); ya que al recibir dichas capas intensidades de movimiento diferentes producen diversas funciones de transferencia en la que se amplifican diferentes valores frecuenciales.

De modo general, los terrenos blandos amplifican las señales cuyas frecuencias predominantes son bajas (períodos altos), coincidentes con los periodos de vibración fundamentales típicos de estructuras ingenieriles altas (esbeltas), mientras que los suelos rígidos lo hacen con señales cuyas frecuencias predominantes son altas (períodos bajos). Por tanto, el grado de amplificación está en función de las propiedades geotécnicas de los terrenos tales como, el índice de poros, la consistencia y la densidad relativa, y la plasticidad (contenido de arcillas), factores que a su vez son el resultado de la historia geológica, origen y génesis de las litologías. [García, Ignacio, 2016]

3. Categorías de respuesta sísmica similar en la amplificación de sitio

En la actualidad existen dos métodos para la estimación de la amplificación del terreno debido a las condiciones geológicas. Los más utilizados son los métodos empíricos, los cuales se basan principalmente en el análisis de la respuesta dinámica del sitio utilizando las señales de registros acelerográficos, situados tanto en lugares con suelos poco consolidados como en suelos competentes; que permitan su calibración.

Por otro lado, están los modelos analíticos o de modelación numérica, los cuales son más complejos ya que se necesita información más precisa del comportamiento dinámico del suelo, tales como su amortiguamiento histerético y degradación de su módulo de corte (al acumularse deformación), el espesor de las diferentes capas, la estimación de la profundidad del estrato rocoso y señales sísmicas (en base a un estudio de amenaza sísmica de la zona) al ser aplicadas al sustrato rocoso para que se propague a través de las diferentes capas de la columna de suelo.

En el marco de este proyecto de actualización de la normativa, se desarrollarán ambos métodos; por lo que se obtendrán resultados interesantes sujetos a la información recabada y necesaria para su determinación. Previamente a la aplicación de dichos métodos, es necesario la realización de una **clasificación o agrupación de los materiales geológicos superficiales en categorías de respuesta sísmica similar**. Precisamente esto, es el objetivo de este trabajo, lograr representar a nivel nacional categorías de suelo que podrían tener comportamiento sísmico homogéneo, a pesar de las grandes incertidumbres que pueden existir siempre dentro de una misma categoría.

Lo idea anterior nace principalmente por la complejidad en la obtención de los parámetros dinámicos in situ, ya sea mediante ensayos de campo o de laboratorio, debido a: a) falta de equipo de laboratorio para cuantificar las propiedades dinámicas, b) áreas de análisis demasiado extensas, c) variaciones bruscas de los materiales tanto en la vertical como en la horizontal, e) dificultades de interpretación mediante las correlaciones empíricas.

Por ello, casi siempre, los análisis se realizan asumiendo parámetros de respuesta sísmica genéricos, como es la velocidad de corte (V_s) en función de la situación, edad, composición y origen de las formaciones geológicas. (García, Ignacio, 2016). Este valor de V_{s30} es uno de los parámetros más importantes considerados a nivel mundial para la clasificación de los suelos, que corresponde al promedio de las velocidades de ondas cortantes de las capas de suelo ubicadas desde la superficie hasta 30 metros debajo de ella (Schmidt, 2014).

El hecho de definir el promedio de velocidades a los 30 metros de profundidad como un criterio para la clasificación de suelos, obedece a que el suelo dentro de esta profundidad es el que afecta en mayor medida el comportamiento de las estructuras que se edifiquen

sobre él (Schmidt, 2014). A este respecto, diversas normativas sísmicas de varios países establecen varios tipos de suelos con respuestas sísmicas similares (Eurocódigo 8, UBC 1997 (*Uniform Building Code*), NEHRP 2003).

El Programa de Reducción de Amenazas por Terremotos de Estados Unidos (NEHRP) en su versión de 2003, define cinco clases de suelo, todos dependientes de Vs30 y de otros parámetros que ayudan a medir de resistencia del suelo, como lo es el valor de N del ensayo de SPT. Esta clasificación también es utilizada en el código sísmico de Estados Unidos, conocido como el ASCE 7-16; la cual se apega también a las recomendaciones propuestas por Borchardt en el año 1994, el cual plantea una gran variedad litológica y otras características que influyen en la respuesta sísmica del terreno, tales como el grado de saturación y espesor del sustrato. En la tabla No.1 se muestra dicha clasificación.

Tabla 1. Clasificación de suelos según NEHRP 2003

No	Clases de Suelo		Criterio Geotécnico (promedio de 30 m superiores)					
	Tipo	Descripción	Vs (m/s)		Vs (m/s) Media Geométrica	N or Nch	Su (psf)	
1	A	Roca dura	>	1500			NA	NA
2	B	Roca media		760.0	1500.0	1070	NA	NA
3	C	Roca blanda y suelo muy denso		360.0	760.0	523	> 50	> 2000
4	D	Suelo rígido		180.0	360.0	255	15 to 50	1000-2000
5	E	Suelo blando arcilloso	<	180.0			< 15	< 1000
6	F	Suelos que requieren análisis de respuesta del sitio de acuerdo con la Sección 21.1		Ver Sección 20.3.1				

La expresión del cálculo de Vs30 se define a continuación:

$$V_{s30} = \frac{\sum_{i=1}^n d_i}{\sum_{i=1}^n \frac{d_i}{V_{si}}}$$

Donde, di, es el espesor de cualquier capa entre 0 y 30 metros de profundidad, Vsi, es la velocidad de corte de la capa en particular. Esta clasificación ha servido mucho para diferentes autores para poder calibrar sus términos de efecto de sitio en las ecuaciones de predicción de movimiento fuerte. No obstante, el rango de clasificación del Vs30 se considera muy amplio, y podrían ajustarse los valores mínimos y máximos. Por ello, el NEHRP para su versión a ser publicada en este año 2020, ha sugerido incorporar más rangos de suelo. Lo anterior, se le suma a que en pasadas reuniones de los comités se acordó en incorporar espectros de diseño multiperiodo y con eso la determinación de factores de amplificación para una mayor variedad de suelos. A continuación, en la tabla 2 se muestra esta nueva clasificación.

Tabla 2. Clasificación de suelos según NEHRP 2020

No	Clases de Suelo				Criterio Geotécnico (promedio de 30 m superiores)					
	Tipo Clase 2016 NEHRP	Tipo Clase 2020 NEHRP	Descripción general	Descripción detallada	Vs (m/s)		Vs (m/s) Media Geométrica	N or Nch	Su (psf)	
1	A	A	Roca dura	Roca dura	>	1500			NA	NA
2	B	B	Roca media	Roca media		910	1500	1160	NA	NA
3		BC	Roca blanda	Roca blanda		640	910	760	NA	NA
4	C	C	Suelo duro	Suelo muy denso o arcilla dura.		440	640	530	> 50	> 2000
5		CD	Suelo medio duro	Suelo de arena densa o arcilla muy rígida.		300	440	360		
6	D	D	Suelo poco duro	Suelo de arena medianamente densa o suelo arcilloso rígido		210	300	250	15 to 50	1000-2000
7		DE	Suelo blando	Suelo de arena suelta o suelo arcilloso medio rígido		150	210	180		
8	E	E	Suelo muy blando	Suelos de arena muy suelta o suelo arcilloso blando	<	150			< 15	< 1000
9	F	F	Suelos que requieren análisis de respuesta del sitio de acuerdo con la Sección 21.1			Ver Sección 20.3.1				

Esta clasificación es la que se utiliza para la agrupación de los diferentes tipos de suelo de El Salvador, desarrollados en este trabajo. En los apartados posteriores, se detallará como se utilizaron los valores de esta tabla para asignar a las unidades geológicas encontradas dentro de la base de datos consolidada.

4. Información para base de datos

Para lograr una aceptable caracterización dinámica de los suelos, se debe contar con la mayor información posible que relacione el comportamiento elástico o inelástico de los mismos, durante son atravesados por las ondas sísmicas. Esta información se concentra en reunir la mayoría de mediciones de velocidades de las ondas de corte polarizadas (SH)

recabadas a partir de ensayos sísmicos-geofísicos como lo son Downholes. Estos ensayos ayudan a recrear la estructura del subsuelo por medio de la propagación de las velocidades de las ondas de compresión P y las ondas de corte S; con lo que se logra inferir cambios en las unidades geológicas que se están atravesando y por ende la determinación de otras propiedades geo-mecánicas.

Es sumamente importante que, acompañado a este trabajo de determinación de velocidades de corte, se ejecute un reconocimiento geológico tanto de las muestras obtenidas en la perforación como en los alrededores de la zona que se está estudiando.

Este reconocimiento o muestreo debe realizarlo un profesional en el área de geología, el cual debería proporcionar descripciones de las unidades geológicas relacionando su génesis eruptivo, edades y estimaciones de propiedades geotécnicas. Lo anterior, principalmente para poder centrar la atención en caracterizar de una manera empírica la variabilidad de deformación del suelo respecto al cambio en su módulo de corte y su amortiguamiento histerético; logrando capturar el comportamiento inelástico del suelo.

La información recabada para este trabajo se muestra a continuación:

- Ensayos Downhole de 15 perforaciones de proyectos MARN.
- Ensayos Downhole obtenidos de otros proyectos.
- Perfiles geológicos de pozos plataforma SIHI-MARN.
- Perfiles geológicos de afloramientos.
- Perfiles de velocidad a partir de ensayos de microtrepidación.
- Ensayos SPT que alcanzan más de 30 metros de profundidad
- Otras fuentes de pozos de perforación.

4.1. Ensayos Downhole de 15 perforaciones de proyectos MARN

Entre los años 2018 y 2020, el MARN participó en dos proyectos donde se realizaron 15 perforaciones de pozos; dentro de los cuales se ejecutaron ensayos Downhole hasta los 50 metros de profundidad.

El primer proyecto desarrollado entre el 2018 y 2019, tuvo como objetivo general contribuir a la caracterización de las diferentes unidades acuíferas portadoras del agua a nivel nacional. Por lo que el MARN ejecutó a través de este proyecto llamado: “Perforación de 8 pozos de monitoreo de aguas subterráneas-Fase II”, un proceso de monitoreo del comportamiento de los acuíferos en zonas definidas como prioritarias, ya sea por su potencial de recarga acuífera, por su alto valor medioambiental o por el mantenimiento de una gran cantidad de actividades socioeconómicas.

Lo anterior, se efectuó en base a la perforación de 8 pozos de monitoreo para medir niveles de agua y toma de muestras para determinación de los componentes químicos

para verificar la calidad del agua. Estos pozos contaron con recuperación de testigos o núcleos de perforación de diámetro PQ (3.345 plg.) con corona de insertos de diamante industrial. Estas perforaciones están distribuidas en diversos sitios de interés del territorio nacional, con fines de monitoreo hidrogeológico, y sus profundidades varían entre 60 m en Santiago Nonualco hasta 180 m en el parque Bicentenario. Los detalles de los lugares donde se ubicaron los pozos son los siguientes:

Tabla 1. *Ubicación de Downholes de proyecto hidrogeológico del MARN*

No.	Sitio	Municipio	Zona Prioritaria	MASub	Profundidad (ML)	Latitud Norte (Grados)	Longitud Oeste (Grados)
1	C.E. Cantón el Zope	Acajutla	ZP1 (GdeS - Banderas)	ESA-02	70	13.580053	-89.804306
2	Turicentro Atecozol	Izalco	ZP1 (GdeS - Banderas)	ESA-02	100	13.746314	-89.665947
3	Centro Escolar Ranulfo Castro	Chalchuapa	Río Paz	ESA-03	100	13.984726	-89.672913
4	Hospital Nacional de La Unión	La Unión	ZP4 (G. de San Miguel)	ESA-20	90	13.347242	-87.876904
5	Centro Escolar de Ereguayquín	Ereguayquín	ZP3 (Bahía de Jiquilisco)	ESA-12	100	13.34485	-88.389519
6	Centro Escolar Cantón Tecualuya	San Luis Talpa	ZP8 (Sucio-Acelhuate)	ESA-07	80	13.467648	-89.097382
7	Centro Escolar Caserío San Cristóbal	Santiago Nonualco	ZP2 (Estero Jaltepeque)	ESA-07	60	13.357693	-88.976945
8	Parque Bicentenario	SS / Antigua Cuscatlán	ZP8 (Sucio-Acelhuate)	ESA-06	150	13.685486	-89.253296

El segundo proyecto iniciado desde el año de 2019, es el que se encuentra desarrollándose en estos momentos, el cual se ejecuta como parte de las acciones para la actualización de la normativa nacional para diseño y construcción sismorresistente y plan nacional para la reducción de la vulnerabilidad sísmica. Este proyecto lleva como nombre: “Plan de Acción para la Implementación del Índice de Gobernabilidad y Política Pública para la Gestión de Riesgo de Desastre”.

En este caso, uno de los componentes del proyecto se basa primordialmente en la caracterización dinámica de los suelos de El Salvador; por lo que se planteó tanto la realización de 7 perforaciones de pozos en puntos estratégicos como la adquisición de equipos especializados para conocer a través de muestras las propiedades dinámicas de los mismos. La selección de los lugares fue analizada por el equipo de investigación del MARN, con el fin de cubrir aquellos lugares donde si bien se conocía la geología superficial, sin embargo, se tenía el conocimiento (a través de sismos importantes y estudios antecesores) que presentan importantes efectos de la amplificación del sitio, así

como también donde se concentran mayores núcleos de población e infraestructura importante como viviendas, edificios, etc.

Las 7 perforaciones se realizaron con recuperación de testigos o núcleos de perforación con diámetro HQ, con corona de insertos de diamante industrial. Estos pozos presentaron profundidades que van desde 40 m en La Laguna, Chalatenango hasta 80 m en Aviación Civil, Ilopango. Los detalles de los lugares donde se ubicaron los pozos son los siguientes:

Tabla 2. *Ubicación de Down-holes de proyecto BID-MARN-UCA-MOP*

No.	Ubicación	Latitud Norte (Grados)	Longitud Oeste (Grados)
1	Santa Ana, UNICAES	13.980201	-89.546733
2	ITCA, Frente a edificio Rectoría, Santa Tecla	13.674181	-89.279814
3	Aeropuerto de Ilopango	13.696616	-89.116861
4	DIGESTYC, Ciudad de Delgado	13.713719	-89.169868
5	Centro Escolar de La Laguna, Chalatenango	14.164557	-88.943201
6	Universidad de El Salvador San Miguel	13.438834	-88.159131
7	Unidad de Salud de Perquín Morazán	13.958808	-88.157734

Como se dijo anteriormente, en estas 15 perforaciones se realizaron ensayos Down-hole hasta una profundidad de 50 m, debido a la longitud y seguridad para introducir la sonda del equipo. Este ensayo se basó en la generación artificial de ondas sísmicas a través de golpes con almádana en dos placas de acero que se colocan en pequeños agujeros. Se miden las velocidades de las ondas de cortes de cada sismo que atraviesa los diferentes estratos en el sondeo.

4.2. Ensayos Downhole obtenidos de otros proyectos.

Durante la búsqueda de información se encontró que han existido proyectos de gran envergadura en los cuales, los proyectistas vieron la necesidad de realizar los ensayos Down-hole. A continuación, se presentan el resumen de estos ensayos.

Tabla 3. *Ubicación de Down-hole de otros proyectos*

No.	Ubicación	Latitud Norte (Grados)	Longitud Oeste (Grados)	Fuente
1	Falla de San Vicente	13.66260	-88.90910	García, Ignacio, 2016
2	Hermano Lejano	13.68470	-89.21800	MOP
3	Apopa 15-B	13.80401	-89.14841	ITALTEKNA
4	Apopa 5-B	13.80824	-89.14276	ITALTEKNA
5	Centro de Investigaciones Geotécnicas (CIG)	13.69870	-89.17330	ITALTEKNA
6	Instituto Geográfico Nacional (IGN)	13.71280	-89.16990	ITALTEKNA

Tabla 3. Ubicación de Down-hole de otros proyectos

No.	Ubicación	Latitud Norte (Grados)	Longitud Oeste (Grados)	Fuente
7	Hogar del Niño – San Jacinto	13.68769	-89.18814	ITALTEKNA
8	Universidad Centroamericana José Simeón Cañas (UCA)	13.68020	-89.23570	ITALTEKNA
9	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-1)	13.66190	-89.28670	LOTTI
10	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-2)	13.66220	-89.28870	LOTTI
11	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-3)	13.66180	-89.29090	LOTTI
12	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-4)	13.66400	-89.29340	LOTTI
13	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-5)	13.66300	-89.28490	LOTTI
14	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-6)	13.66450	-89.28740	LOTTI
15	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-7)	13.66460	-89.28950	LOTTI
16	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-8)	13.66560	-89.29540	LOTTI
17	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-9)	13.66540	-89.29830	LOTTI
18	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-10)	13.66500	-89.30080	LOTTI
19	Deslizamiento Las Colinas (2001), Santa Tecla (BAL-11)	13.66790	-89.30020	LOTTI

A continuación, se presenta una figura donde se distribuyen los ensayos Down-hole.

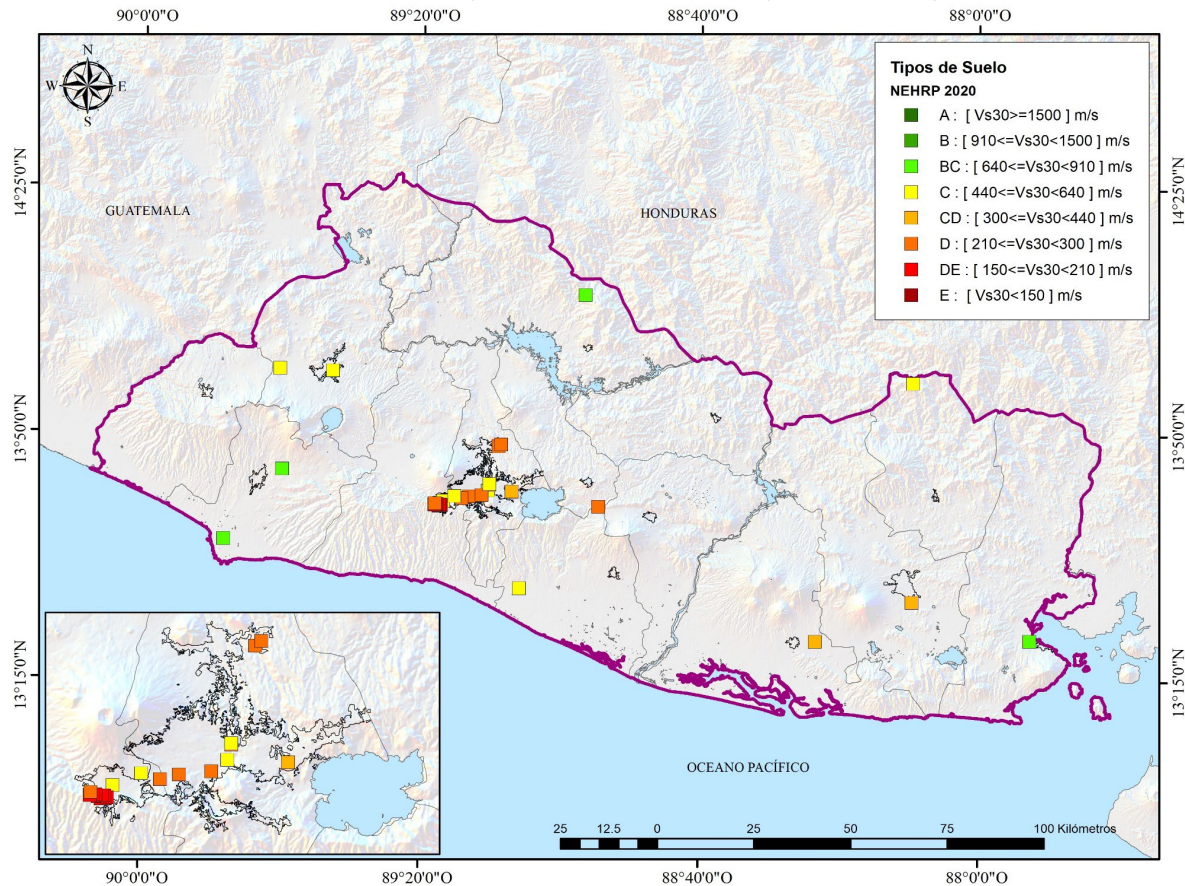


Figura 1. Ubicación geográfica de 34 ensayos Down-holes. Fuente: MARN. 2020

4.3. Perfiles geológicos de pozos plataforma SIHI-MARN

La plataforma SIHI es un sistema informático para la gestión, análisis y consulta de información hídrica, en apoyo a la toma de decisiones de los procesos de gestión y planificación del recurso hídrico en El Salvador.

La información de dicha plataforma se comprende mayoritariamente con información recabada por la institución ANDA; la cual es principalmente obtenida para contar con pozos de monitoreo para medir niveles de agua y toma de muestras para determinación de los componentes químicos para verificarla calidad del agua. En ella se pueden encontrar un total de 353 puntos de pozos con información litológica, con descripciones de cada unidad geológica enfocada a la rama de la hidrogeología.

Es de resaltar que estas columnas litológicas fueron realizadas por la percepción de un geólogo que en su momento pudo generalizar capas y que pudo tener incertezas en sus apreciaciones. Asimismo, no siempre los cambios de velocidad se asocian a cambios de estratos geológicos completos, sino que cuando haya un cambio marcado en las propiedades del suelo. Por ello, es bien importante determinar dichas propiedades físico-mecánicas de los suelos, que permiten agrupar de mejor manera esos estratos sísmicos.

Lo anterior, se pudo observar en los resultados obtenidos en los ensayos de Down-hole y la caracterización litológica por el geólogo. Sin embargo, estas descripciones son útiles para poder asignar una clasificación más general respecto a la geología nacional de El Salvador, y así poder inferir luego un valor de velocidad de onda de corte. Esto se explicará con mayor detalle en el apartado 5.3, en el que se utilizará una jerarquía de criterios que siempre serán respaldadas por la opinión experta de un geólogo.

Para ello, se formuló junto a estudiantes de Geofísica de la UES y opiniones de geólogos que conocen muy bien la estratigrafía y distribución de subunidades geológicas en un buen segmento del país, una metodología para poder inferir profundidades límites de determinadas unidades, de acuerdo a su génesis eruptiva y longevidad en sus edades de formación.

Con ello se pudo trazar y lograr estimar la velocidad de corte de dichas unidades, logrando construir un perfil de velocidad que ayudaría a calcular el valor de VS30. Si bien esta práctica no es una medida directa de la velocidad de corte, sin embargo, se trató de aprovechar en lo mayor posible la información geológica, la cual era abundante y abarcaba una buena distribución espacial. En el siguiente apartado se hablará con mayor detalle sobre esta metodología.

A continuación, en la figura 2 se muestra la distribución de los pozos que contiene SIHI, junto a la clasificación de tipo de suelo según NEHRP 2020, y valores de Vs30 calculado.

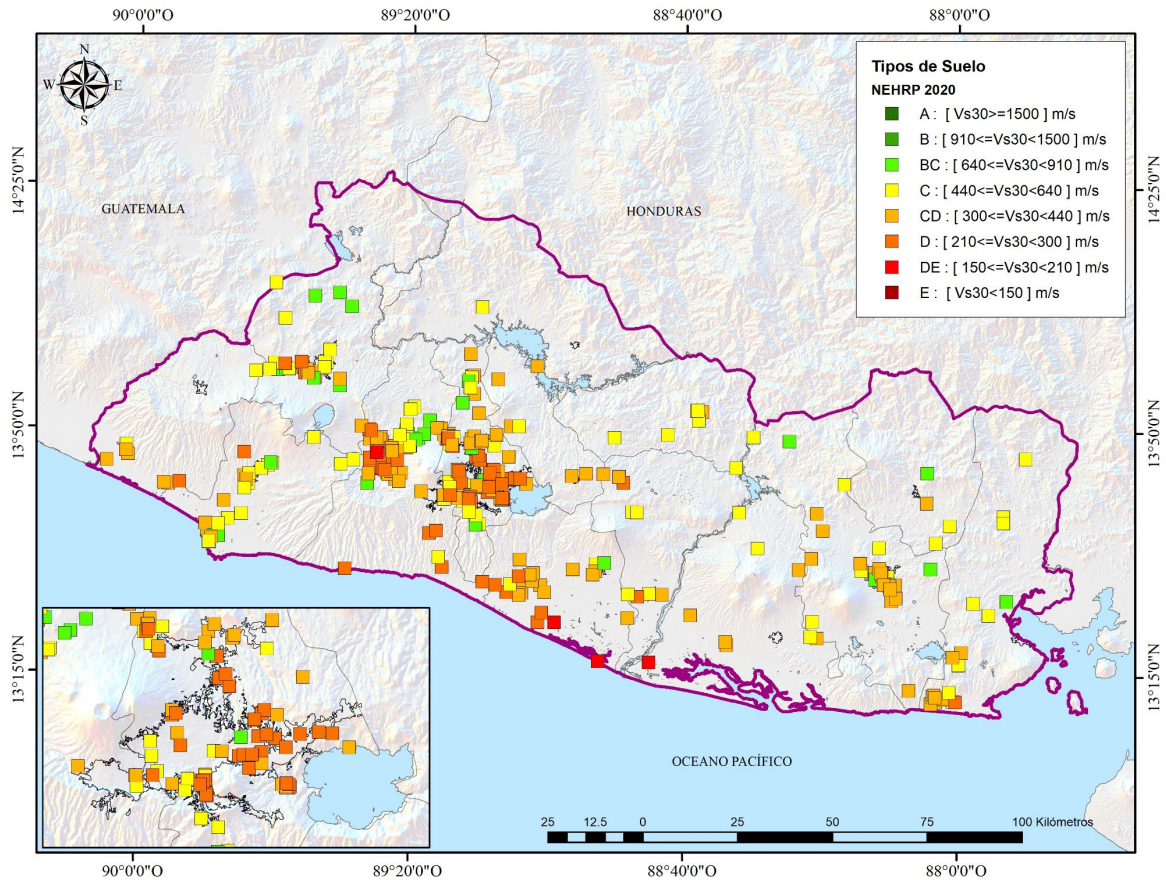


Figura 2. Ubicación geográfica de 353 pozos de plataforma SIHI. Fuente: MAR.N. 2020

4.4. Perfiles geológicos de afloramientos.

Esta información se basó en recabar puntos de afloramientos que diversos geólogos han construido para estudiar unidades geológicas específicas. La gran mayoría de estos afloramientos son muy superficiales que sirven para poder trazar isopacas de espesores de las unidades; y no podrían ser utilizados para el cálculo de la velocidad promedio de ondas de corte en los 30 metros; sin embargo, son puntos clave que sirven de control para visualizar la distribución de unidades geológicas y así comparar perfiles que sean semejantes a los perfiles de mayor profundidad, con el objetivo de extrapolar comportamientos homogéneos del subsuelo.

La información se concentra en los estudios de columnas litológicas de la Dra. Dolors Ferrés, el geólogo Walter Hernández y el estudio de investigación de Pedrazzi, D. et al. 2019. La primera fuente se centra en levantamientos estratigráficos de productos eruptivos del volcán de San Salvador (Boquerón), en los cuales, en ciertos puntos, figuran espesores de tobas de la caldera de Ilopango, especialmente la Tierra Blanca Joven (538 DC). Por otra parte, la segunda y tercera fuente, compilan puntos de espesores de productos o tobas de la caldera de Ilopango.

Con estas fuentes se pudo entender más sobre la cronología de los productos de dichos volcanes y definir con más justificación (junto a otra literatura de estudios sobre dichas unidades) los límites o fronteras de las tobas. Un total de 115 columnas litológicas fueron extraídas de la tesis doctoral de la Dr. Ferrés, de las cuales, únicamente 100 presentaban las descripciones y valores de espesores en su perfil. De estos, solamente 2 superan los 30 metros de profundidad. En la figura 3 se presenta la distribución de puntos de afloramiento de Walter Hernández y el investigador Darío Pedrazzi; donde mayoritariamente se registraban los espesores de TBJ, TB3 y TB4.

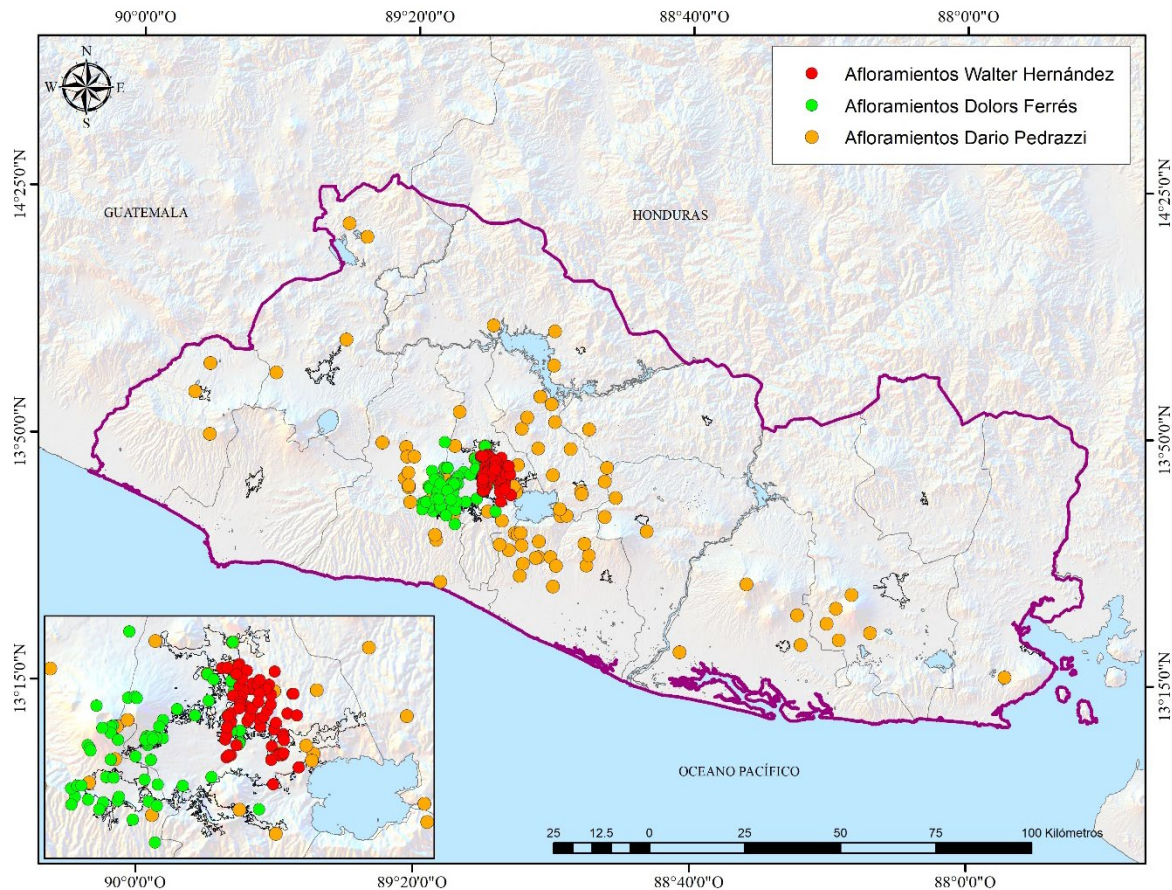


Figura 3. Ubicación geográfica de afloramientos. Fuente: MARN. 2020

4.5. Perfiles de velocidad a partir de ensayos de microtrepidación

Los microtemblores o microtemblores, son vibraciones generadas por eventos artificiales producto de la actividad humana tales como tráfico, maquinaria industrial y explosiones usando dinamita.

Estos serán discutidos con un geólogo para definir con mejor criterio el tipo de suelo que muestran las descripciones generales de los perfiles de pozos (Moreno & Alfaro, 2000). Los microtemblores, están compuestos principalmente por ondas Rayleigh y S y, al contrario de los microsismos, son de periodo corto (Aki, 1957; Nogoshi, 1961).

Los microtremores se localizan en el intervalo de frecuencias entre 0.01 a 30 Hz, sin embargo, en la exploración geofísica solo enfoca su interés en el intervalo de 0.1 a 10 Hz, debido a que, en esta banda, parte de la energía se transmite como ondas Rayleigh, cuyos modos y velocidades de propagación se pueden estimar con arreglos instrumentales para obtener información sobre los parámetros elásticos de la corteza (Asten, 1976).

El estudio de microtrepidaciones ha tenido excelentes resultados en muchos casos para la determinación de los períodos de vibración dominante de los suelos, pero en algunos otros los resultados son contrarios a los esperados, por esta razón, estudios de este tipo deben ser interpretados con cautela y en la mayoría de las veces, corroborados por modelos teóricos que evalúen con más certeza la respuesta dinámica de los suelos.

Los ensayos de microtrepidación fueron compilados a través de dos trabajos de graduación desarrollados con la Universidad Centroamericana José Simeón Cañas (UCA) con apoyo del MARN, utilizando una serie de sensores digitales (sismómetros) de 1 y 3 componentes. Estos equipos son propiedad de la UCA, siendo donados por la organización JICA-Japón. De igual manera, estos sensores fueron utilizados en un trabajo anterior, donde se buscaba caracterizar las propiedades dinámicas de ciudades importantes del país (JICA, et al 2007).

El primer trabajo de graduación consistió en determinar los períodos de vibración y perfiles de velocidad de corte de puntos geográficos de El Salvador mediante el análisis digital de registros de microtrepidaciones. Estos puntos coinciden con ubicaciones cercanas donde se realizaron los ensayos de Down-hole, mencionados anteriormente; los cuales tratan de cubrir lo más posible las formaciones geológicas planteadas en el mapa geológico de El Salvador. Estos resultados también servirán como puntos de calibración, al momento de la generación de modelos analíticos unidimensionales de la respuesta sísmica del suelo.

El segundo trabajo de graduación se basó específicamente en determinar los períodos de vibración y perfiles de velocidad de corte con el fin de caracterizar la respuesta sísmica del suelo de emplazamiento de algunas de las estaciones acelerográficas de El Salvador, a partir del análisis de registros de microtremores. El método de exploración de microtremores que fue practicado en estos trabajos de graduación, ha sido una formulación más versátil de la técnica (ampliamente utilizada por la comunidad de investigadores), conocida como Método de Autocorrelación Espacial (SPAC) desarrollado por Aki (1957). Esta técnica innovadora llamada *Centerless Circular Array* (CCA), es un algoritmo de exploración de microtremores que se puede utilizar para estimar las velocidades de fase de las ondas de Rayleigh mediante el análisis y procesamiento de registros de componentes verticales de microtremores que se obtienen con una matriz de tres o más sensores sísmicos colocados en la superficie del suelo alrededor de una circunferencia y no requiere registros en el centro del círculo.

Se ha confirmado, a través de pruebas de campo, la aplicabilidad del método CCA a matrices del orden de varios a varios cientos de metros en radios y se ha revelado su rendimiento notablemente alto en rangos de longitud de onda larga. El método cuenta con su propio modelo matemático que permite evaluar las relaciones de señal-ruido y el escrutinio de los datos de campo ha confirmado la hipótesis de que el ruido es el factor principal que sesga los resultados en rangos de longitud de onda larga (Cho, et al., 2006).

La principal razón del éxito que ha tenido la técnica de arreglos dentro de la comunidad científica, se justifica en que su aplicación práctica es relativamente sencilla, sobre todo porque utiliza fuentes pasivas que son las vibraciones ambientales, por lo que no es invasiva ni destructiva. Por lo que aplicarlas es posible identificar algunas características de las capas sedimentarias (espesor y velocidad cortante) a un costo bastante inferior.

A continuación, en la figura 4 se presenta un mapa con la localización de estos ensayos de microtrepidación. Cabe mencionar, que los valores calculados de velocidad de corte para el primer trabajo de graduación fueron sustituidos por los obtenidos con los ensayos Downhole, al tratarse de mediciones directas; sin embargo, sirvieron como puntos de control principalmente en aquellos casos que el nivel freático era muy somero y donde el ensayo Downhole podría tener incertidumbres en sus resultados.

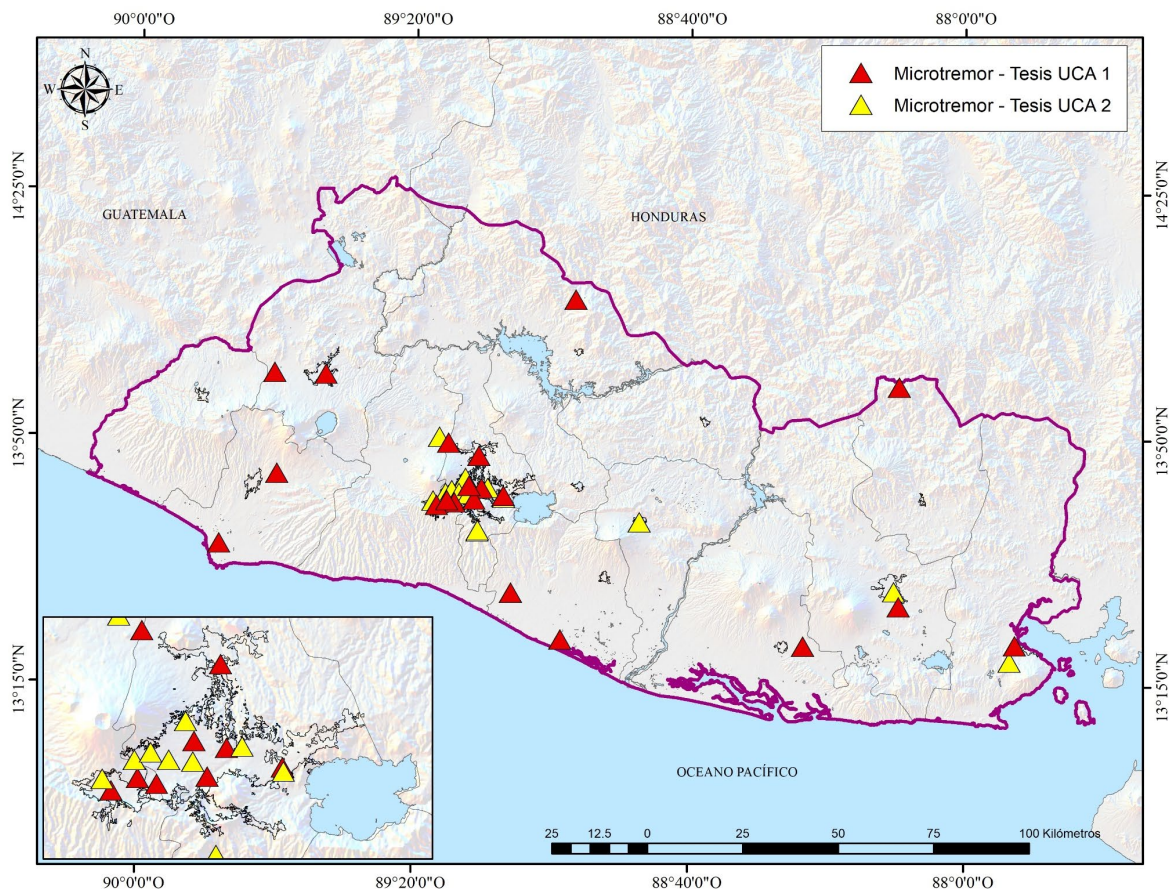


Figura 4. Ubicación geográfica de ensayos de microtrepidación con CCA. Fuente: MARN. 2020

4.6. Ensayos de Penetración Estándar (SPT)

Esta información fue recabada tanto del Ministerio de Obras Públicas (MOP), la Oficina de Planificación del Área Metropolitana de San Salvador (OPAMSS) y una serie de sondeos extraídos de los reportes publicados por la empresa Consorzio Salvador E. Italtekná-Italconult, para proyecto de caracterización sísmica y dinámica de suelos desarrollado en 1988, a raíz del terremoto del 10 de octubre de 1986.

Por parte del MOP se proporcionó una serie de registros de perforación basados en estudios de penetración estándar (SPT) las cuales presentaban profundidades menores a los 10 metros. Estos estudios contenían el valor de número de golpes (N) y una clasificación visual manual de los materiales encontrados en el subsuelo. Estos registros datan desde el año de 2007 al 2018, haciendo un total de 61 registros.

De igual manera, OPAMSS (con el cual se tiene un convenio de cooperación para intercambio de información) proporcionó una serie de puntos dentro del AMSS, los cuales se basaban en perforaciones de pozos con ensayos de SPT, donde en su mayoría poseen profundidades menores a los 10 metros. Un total de 3149 sondeos fueron recopilados, de los cuales 426 presentaban profundidades mayores a los 10 metros.

Si bien estos sondeos no abonarían para el objetivo de este trabajo (que es el de obtener el valor de Vs30), siempre es una información vital para poder correlacionar parámetros como la resistencia del suelo (número de golpes N) y el valor de velocidad de corte.

Existen una gran cantidad de correlaciones exponenciales ya documentadas en la literatura; de igual manera, el MARN está realizando un trabajo en paralelo con las nuevas perforaciones realizadas, ya que también se registraron los valores de N de SPT. Lo anterior, ayudaría a poder extrapolar valores de N y determinar velocidades de corte y ayudar en la calibración de los tipos de suelo del país.

A continuación, se muestran los mapas donde se visualiza la ubicación de los sondeos de la OPAMSS. Los sondeos del MOP y de Italtekná-Italconult no fueron graficados, ya que se encontraban en documentos en físico; por lo que se necesitaría un trabajo adicional para poder plotearlos.

4.7. Otras fuentes de pozos de perforación y ensayos geofísicos

En esta información se consolidan las diferentes fuentes en las que se pudieron obtener estudios de pozos de perforación. A continuación, se detallan dichas fuentes:

- 26 pozos profundos de proyecto FORGAES (que no se encontraban en plataforma SIHI), ubicados en AMSS; mostrando únicamente la litología de las unidades.

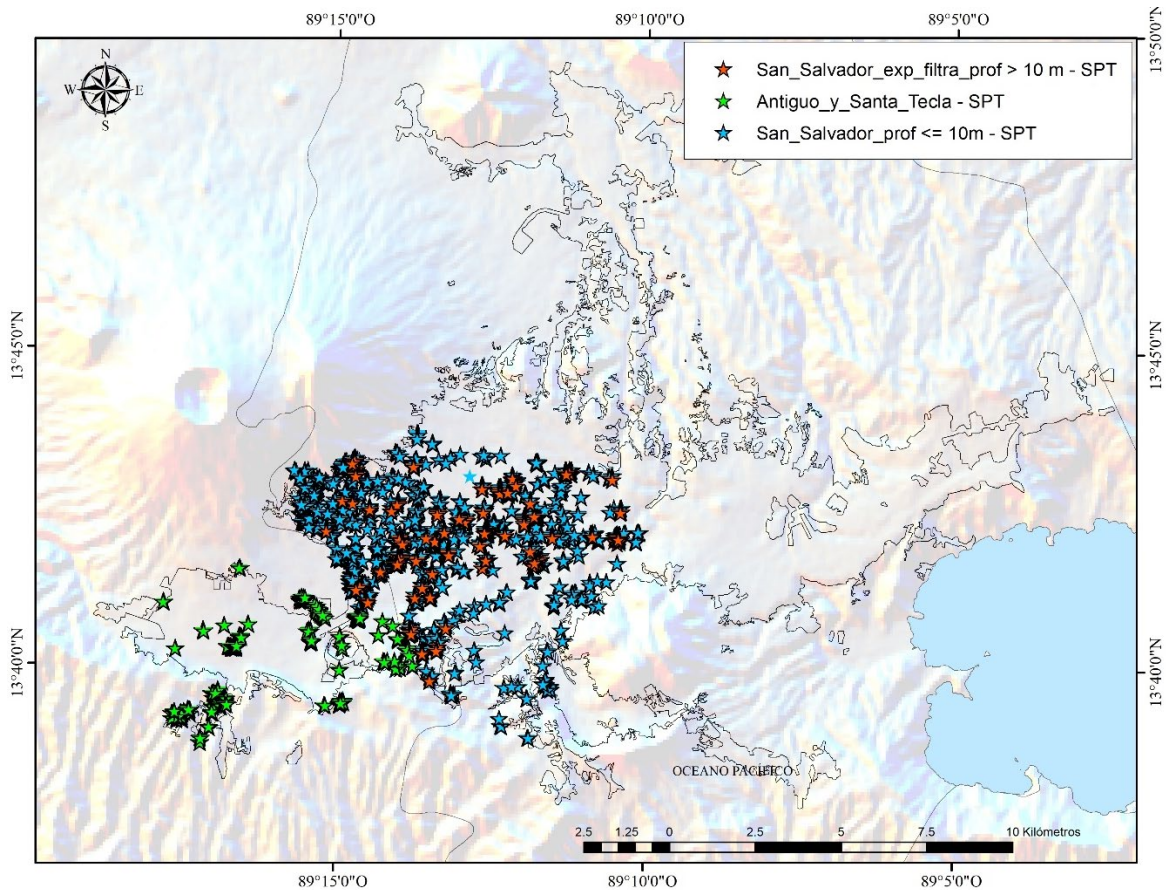


Figura 5. Ubicación geográfica de ensayos SPT proporcionados por OPAMSS. Fuente: MARN. 2020

- 46 pozos realizados por empresa Consorzio Salvador E. Italtেকna-Italconsult, para proyecto de caracterización sísmica y dinámica de suelos, desarrollado en 1988, a raíz del terremoto del 10 de octubre de 1986; ubicados en AMSS. Estos pozos mostraban una clasificación en unidades A y B de los estratos de suelo. La unidad A consiste en los depósitos de TBJ, Tobas superiores e Inferiores de la caldera de Ilopango y la unidad B a los depósitos lávicos (Boquerón). De igual manera, dentro de la unidad A, se clasificaron dos grandes grupos a los cuales se les determinó una curva dinámica característica. Estos fueron el grupo A, el cual consistía en suelos limos arcillosos, arenas no cementadas o poco cementadas de capas de TBJ, a menos de 5 metros de profundidad, con un 5 y 25 por ciento de contenido de arcilla. Por otro lado, el grupo B que consistía en suelos areno limosos (con menos del 5 por ciento de contenido de arcilla) provenientes de diferentes niveles del subsuelo dentro de las Tobas superiores.
- 7 pozos profundos de Estrada y Montenegro, pertenecientes a la información publicada por Consorzio Salvador E. Italtেকna-Italconsult, 1988, ubicados en AMSS. (Puntos que no se encontraban en la base de datos de Italtেকna)

- 17 pozos profundos de ANDA (que no se encontraban en plataforma SIHI), ubicados en AMSS, mostrando únicamente la litología de las unidades.
- 2 pozos geotérmicos en Complejo de Berlín CEL (TR5, TR5-A)
- 15 mediciones de ensayo geofísico MASW en zona de proyecto El Chaparral, Las Cruces, Las Maromas, Las Mesas, Las Marías, La Honda y Carolina. El ensayo MASW (Análisis multicanal de ondas superficiales), es un método para la caracterización geomecánica del terreno, basado en el análisis de dispersión de las ondas sísmicas superficiales para calcular un modelo de velocidad de ondas S del subsuelo; utilizando una línea de geófonos que se sitúa sobre el terreno. La señal sísmica puede proceder de una fuente controlada, o del propio ruido ambiente.
- 9 mediciones de ensayo geofísico MASW en zona de falla de San Vicente, llevado a cabo por el trabajo de investigación de Ignacio García en el año 2016.
- A continuación, se muestran los mapas donde se visualiza la ubicación de los anteriores pozos de perforación.

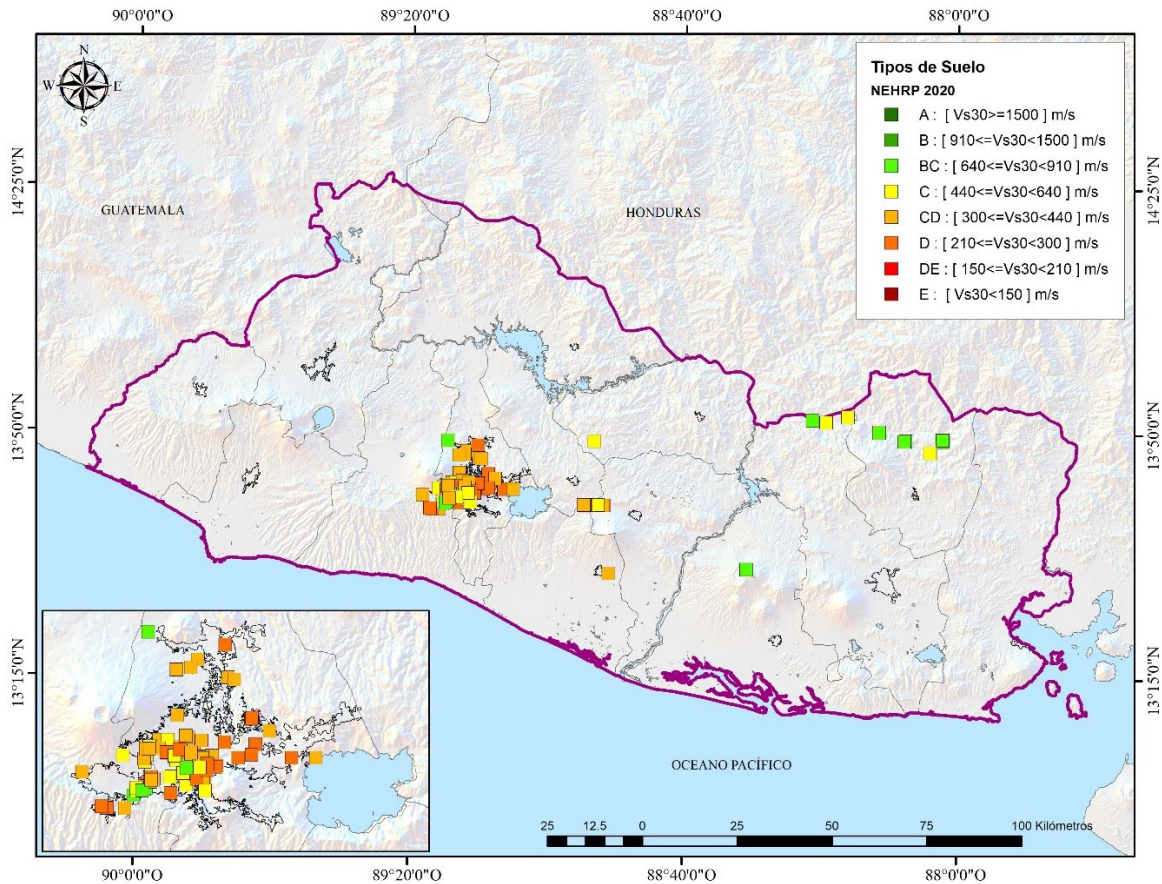


Figura 6. Ubicación geográfica de otros pozos de perforación y ensayos geofísicos. Fuente: MARN. 2020

5. Estimación de Vs para perfiles geológicos

Con el objetivo de aprovechar la información litológica contenida en los perfiles y columnas geológicas, se ve necesario realizar una agrupación estimada de las diversas unidades, según sus posibles comportamientos frente a sismos. Lo anterior, se realizará mediante una reclasificación controlada de dichas capas. Esta reclasificación será de mucha utilidad para contrarrestar la limitada cantidad de datos a lo largo de todo el territorio.

Para ello, es necesario una metodología basada principalmente en la opinión de geólogos y geotecnistas concedores de la geología de El Salvador, así como con la literatura existente sobre la misma; con el objetivo de delimitar diferentes espesores contenidos en los perfiles que presenten características similares asociados a influir en sus respuestas sísmicas.

Como parte de la metodología, se revisarán inicialmente dos estudios anteriores donde se han hecho agrupaciones de unidades geológicas del país; los cuales marcaran la línea para realizar las reclasificaciones de los perfiles recabados. El primer trabajo es la tesis doctoral de Ignacio García, en el año 2016, de la Universidad Politécnica de Madrid, enfocado en determinar los factores geológicos-geotécnicos que controlan los deslizamientos inducidos por terremoto. El otro trabajo corresponde al documento del MARN desarrollado en el 2011, por la Inga. Celina Kattan, orientado a conocer la amplificación del sitio en el AMSS.

5.1. Clasificación de Ignacio García, 2016

En el trabajo del Dr. García se realizó una clasificación desde un punto de vista geotécnico tomando como base el mapa geológico de El Salvador, a escala 1: 100,000, desarrollado por la misión Geológica Alemana (Bosse et al. 1978) y digitalizado por la Universidad Politécnica de Madrid. Este trabajo se complementó con datos geotécnicos de muestras obtenidas y otros datos bibliográficos. A este respecto, se utilizaron un conjunto de criterios geológicos (litología, edad, formación, génesis, ambiente deposicional), espesor de los depósitos blandos, criterios geotécnicos, grado de fracturación, grado de compactación, etc.

De igual manera, el autor incluyó cierto grado de conocimiento que fue adquirido en las diferentes campañas de campo realizadas con otros geólogos, ya que argumentaba que solo tener la información de las capas aflorantes que ofrece el mapa geológico, no sirve para determinar el comportamiento geotécnico existente. Finalmente, el autor logró unificar esta clasificación con otra categorización sísmica que corresponde a la elaborada por Borchardt en el año de 1994, así como también su formulación empírica para determinar factores de amplificación. A continuación, en la tabla 3 se muestra la clasificación hecha por Ignacio García. A simple vista, se observa una cuasi semejanza

en los tipos de suelo propuestos por el NEHRP 2020. Sin embargo, es importante recalcar que el autor no utilizó el concepto de las velocidades de ondas de corte como promedio de los 30 metros superiores; por lo que hay que tener cuidado al momento de correlacionar los tipos de suelo con los contemplados en la clasificación de NEHRP 2020.

Tabla 3. Clasificación de suelos según Ignacio García, 2016

No	Grupo litológico-geotécnico	Vs (m/s)				Factor de amplificación		
		Min	Med	Max	Ensayo Down-hole	Min	Med	Max
1	Roca dura	760	1130	1500	-	1	1	1
2	Roca media	760	1130	1500	-	1	1	1
3	Roca blanda	360	560	760	390-620	1	1	1
4	Suelo duro	250	300	350	-	1.3	1.7	1.7
5	Suelo medio	200	275	350	120-370	1.6	1.6	2
6	Suelo blando	100	150	200	-	1.6	1.8	2
7	Suelo muy blando	< 100	< 100	< 100	110	2	2.6	3
8	Suelo no consolidado	< 100	< 100	< 100	-	1	1	1

Lo interesante de esta tabla es que el autor compiló los diferentes miembros o unidades geológicas del mapa geológico de El Salvador 1:100 mil (Bosse et al. 1978), en categorías litológicas y geotécnicas. Por lo que es necesario recabar esta reclasificación de las unidades geológicas. A continuación, en la tabla 4 se muestran estas reclasificaciones.

Tabla 4. Clasificación de unidades geológicas en grupos geotécnicos, según Ignacio García, 2016

Grupo litológico-geotécnico	Descripción de los materiales	Formación (Unidad)
Roca dura	Copas rojas (conglomerados de cuarzo y caliza, areniscas, siltitas, lutitas); localmente vulcanitas intermedias intercaladas	Metapán (va)
	Efusivas básicas intermedias	Cuscatlán (c3)
	Efusivas básicas intermedias	Bálsamo (b3)
	Metasedimentos, metavulcanitas; principalmente ts; eventualmente más antiguas o más jóvenes	Metapán (ts'm)
	Rocas intrusivas ácidas hasta intermedias	l
Roca media	Calizas y calizas margosas con una intercalación de capas rojas (capas rojas mapeadas como ts)	Metapán (yo)
	Conglomerados de cuarzo principalmente rojos, siltitas, lutitas; localmente vulcanitas intermedias	Metapán (ts)
	Rocas efusivas ácidas	San Salvador (s3'b)
	Efusivas ácidas e intermedias ácidas (ocurrencias aisladas; eventualmente = ch2)	Cuscatlán (c2)
	Efusivas ácidas e ignimbritas; localmente piroclastitas	Morazán (m1'b)

Tabla 4. Clasificación de unidades geológicas en grupos geotécnicos, según Ignacio García, 2016

Grupo litológico-geotécnico	Descripción de los materiales	Formación (Unidad)
	Efusivas intermedias hasta intermedias-ácidas, piroclastitas subordinadas, alteración regional por influencia hidrotermal	Morazán (m2'a)
Roca blanda	Efusivas básicas intermedias; piroclastitas subordinadas	San Salvador (s2)
	Efusivas básicas-intermedias, piroclastitas volcánicas subordinadas (estratos no diferenciados y edificios volcánicos)	Bálsamo (b2)
	Piroclastitas ácidas, ignimbritas, epiclastitas; localmente efusivas ácidas intercaladas	Chalatenango (ch1)
Suelo duro	Rocas efusivas ácidas, piroclastitas ácidas subordinadas	Chalatenango (ch2)
	Piroclastitas ácidas hasta intermedias; en la parte basal localmente efusivas intermedias hasta intermedias - ácidas	Morazán (m1'a)
	Piroclastitas ácidas, epiclastitas volcánicas	Cuscatlán (c1)
	Piroclastitas intermedias hasta intermedias-ácidas; epiclastitas volcánicas; rocas efusivas subordinadas	Morazán (m2'b)
Suelo medio	Cenizas volcánicas y tobas de lapilli	San Salvador (s5'c)
	Conos de acumulación (escorias, tobas de lapilli, cinder)	San Salvador (s5'b)
	Epiclastitas volcánicas y piroclastitas, localmente efusivas básicas-intermedias intercaladas	Bálsamo (b1)
	Piroclastitas ácidas, epiclastitas volcánicas (Tobas Color Café)	San Salvador (s3'a)
	Piroclastitas ácidas, epiclastitas volcánicas; localmente efusivas básicas intermedias	San Salvador (s1)
Suelo blando	Cono de deyección	Cuaternario (Qd)
	Depósitos sedimentarios del Cuaternario	Cuaternario (Qf)
	Depósitos coluviales con y sin representación de depósitos subyacentes (por ejemplo ch2)	Cuaternario (Qc)
	Depósitos de playa: barra costera con nivel antiguo de costa	Cuaternario (p)
	Principalmente depósitos fluviales	Cuaternario (cf)
	Principalmente depósitos lacustres	Cuaternario (cl)
Suelo muy blando	Tierra blanca: piroclastitas ácidas e epiclastitas volcánicas subordinadas; localmente efusivas ácidas (s3'b)	San Salvador (s4)
Suelo no consolidado	Suelo anmoor	Cuaternario (Qs)
	Depósitos de estuario con y sin manglares	Cuaternario (Qe)

En la figura 7 se visualiza la distribución de las reclasificaciones de las unidades geológicas propuesto por Ignacio García Flórez. En resumen, cada grupo está formado por litologías caracterizadas por parámetros resistentes similares.

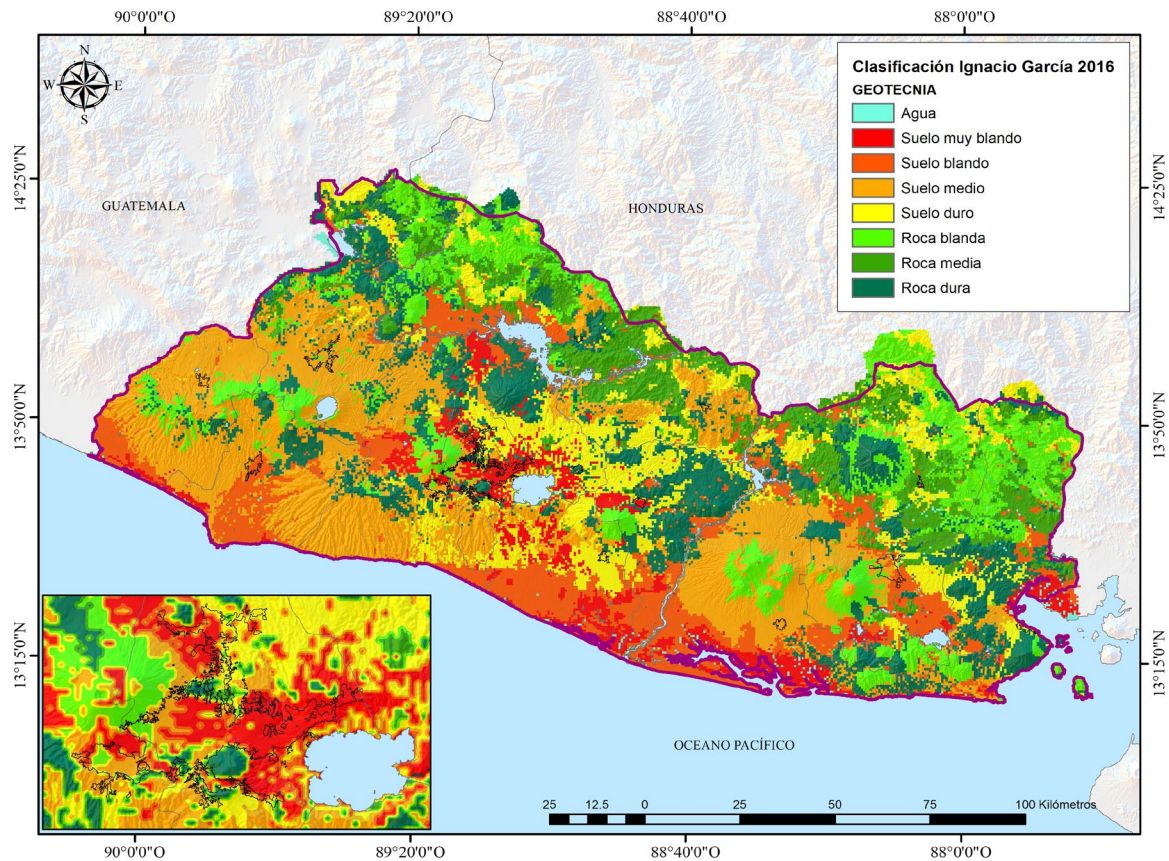


Figura 7. Mapa de distribución de unidades geológicas reclasificadas por Ignacio García Flórez, 2016.
 Fuente: MARN. 2020

Como se puede observar en la figura 7, la variabilidad litológica de El Salvador se traduce en una amplia gama de comportamientos geotécnicos que van desde rocas muy duras y competentes que no amplificarán la señal sísmica (o muy poco), hasta terrenos muy blandos que pueden amplificarla de forma relevante. (Ignacio García, 2016)

5.2. Clasificación de Celina Kattan, MARN 2011

El proceso de reclasificación de unidades geológicas (planteada en el documento de MARN, 2011 creado por la Inga. Celina Kattan); se basa en el acomodo de espesores de tipos de suelo en consonancia con orígenes de procesos eruptivos conocidos que han influenciado la estratigrafía del AMSS, y acoplándose a la categorización general propuesta por el Consorzio Salvador E. Italtেকna-Italconsult, explicada en el apartado 4.7.

Como se menciona en el documento, la ubicación geográfica del AMSS coincide con el graben que existe entre el volcán San Salvador y la caldera de Ilopango, por lo que los

productos de ambos centros eruptivos se encuentran intercalados en la mayor parte de dicha área. Las tefras de TBJ (Tierra Blanca Joven, unidad compuesta por depósitos de caída y flujos piroclásticos de la última erupción explosiva de la caldera de Ilopango y que tienen amplia distribución areal), se adelgazan ladera arriba del volcán San Salvador, pero hacia el oriente de la capital (Soyapango e Ilopango), alcanzan espesores de casi 60 m donde cubren otras tefras de la caldera (Tobas Superiores) y en conjunto superan los 100 m de espesor [MARN, 2012]. Esta reclasificación se muestra en la figura 8.

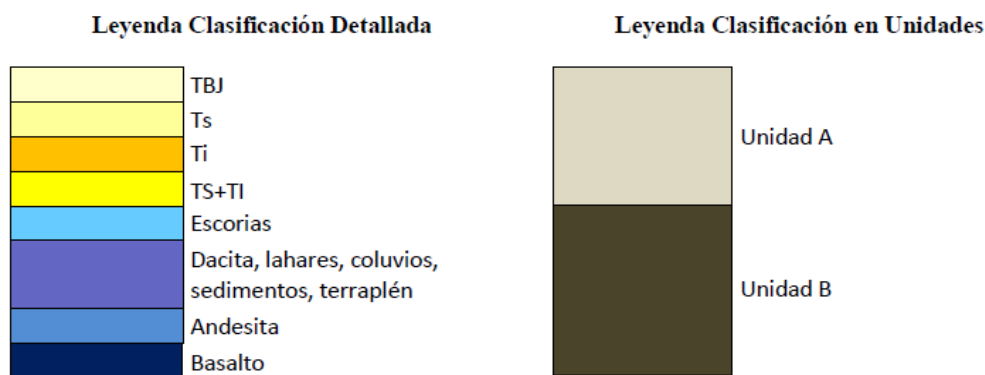


Figura 8. Reclasificación de unidades geológicas. Fuente: MARN. 2011

5.3. Reclasificación de espesores y asignación de Vs

Primeramente, se trabajó con los afloramientos de la Dr. Dolores Ferrés, los cuales se presentan por medio de columnas litológicas para la caracterización geológica del Volcán de Boquerón o San Salvador. Luego se procedió a trabajar con los perfiles geológicos de los 353 pozos contenidos en la plataforma SIHI, los cuales se encuentran distribuidos a nivel nacional (ver figura 2). Como se dijo antes, esta información contempla en su mayoría los pozos perforados por ANDA; enfocados para monitoreo de niveles de acuíferos.

Cabe mencionar, que en ambos trabajos se siguió el proceso de reclasificación de unidades geológicas planteada en el documento de MARN, 2012; con el objetivo de seguir la misma idea del comportamiento sísmico del suelo respecto a su origen eruptivo, definiendo las fronteras y profundidades límites de las unidades geológicas. Respecto a la asignación de un valor de velocidad de corte, se utilizarán los siguientes 3 criterios por nivel de importancia:

- Opinión de expertos sobre la geología de El Salvador, con conocimiento de valores de velocidad de cortes obtenidas en ensayos Down-hole de suelos específicos.
- Utilización de valores promedio de las velocidades de corte presentados por el NEHRP 2020 (Ver tabla 2) en consonancia con la descripción de los tipos de suelo mostrados. Estos serán discutidos con el geólogo para definir con mejor criterio el tipo de suelo que muestran las descripciones de los perfiles de pozos.

- Utilización de valores medios de las velocidades de corte presentados por Ignacio García, 2016e (Ver tabla 3) en conformidad con los grupos litológicos-geotécnicos mostrados. De igual manera, estos serán discutidos con el geólogo.

5.3.1. Afloramientos Dolores Ferrés

Un total de 100 columnas litológicas fueron extraídas de la tesis doctoral de la Dr. Ferrés, para llevar a cabo el proceso de reclasificación de espesores. A esta agrupación se le fueron añadidas las secuencias de depósitos piroclásticos del Volcán Boquerón que forman parte de la formación geológica Cuscatlán, las cuales se encuentran por debajo de la erupción TBJ. También, se agrega el depósito de oleadas piroclásticas y escorias más recientes que TBJ, denominado Talpetate; perteneciente a la formación geológica de San Salvador.

EDIF VOLC	DEPÓSITOS /ESTRATIGRAFÍA	COMPOSICIÓN	FLUJOS IDENTIFICADOS	MUESTRAS	EDAD	
ID) VOLCAN BOQUERON	LAVAS LADERA N y BOQUERONCITO	Andesítica	2	Fairbrothers, 1978 DF112, DF115-116	AD 1917.	
	III) LAVAS asociadas a EDIFICIOS MONOGENÉTICOS	Basáltico-andesítica y Andesítica	2	Sofield, 1998 y DF110-111.	AD 1658 ~1500	
	SECUENCIAS TALPETATE I y TALPETATE II	¿?	Caidas y Oleadas		AD 1031-1285	
	III) LAVAS asociadas a EDIFICIOS MONOGENÉTICOS		5	Sofield, 1998 y DF103-104	AD 1031 a 429	
	TBJ Caldera de Ilopango (cal. AD 429=107.; Dull et al., 2001)					
	III) LAVAS asociadas a EDIFICIOS MONOGENÉTICOS	Basáltico-andesíticas y andesíticas	4	Sofield, 1998 y DF121	1,500 – 3,000 BP	
	BOQUERON I	Basáltico-andesítica	Caida proximal y distal	DF77, DF79	>3,000 BP	
	SERIE SUPERIOR INTERIOR CRATER	Basáltico-andesíticas y andesíticas	7	Fairbrothers, 1978 DF105 a 109 y DF113	3,000 - 8,000 BP	
	SECUENCIA G2	Dacítica	Caida proximal y distal	DF75, DF206	8,000	
	PLANICIE ESTE (SAN SALVADOR)		2	DF131, DF132	8,000 - 16,000 BP	
	TB2 Caldera de Ilopango (16.000 BP; Kutterolf et al., 2008)					
	LAVAS LADERAS SE y NE VB	Basáltico-andesíticas	8	DF122, DF123 y DF125-130	16,000 - 36,000 BP	
	SERIE INFERIOR INTERIOR CRATER	Basálticas y basáltico-andesíticas	7	Fairbrothers, 1978 DF117- DF120		
	TB3 Caldera de Ilopango (30.000 BP; Kutterolf et al., 2008)					
PLANICIE ESTE (SAN SALVADOR)	Dacítica (?)	1	DF-136	30,000 - 36,000 BP		
SECUENCIA G1	Andesítica-Dacítica	Caidas G10 y G11	Sofield, 1998; DE10, 28G11, DF28G10			
TB4 Caldera de Ilopango (36.000 BP; Kutterolf et al., 2008)						
J) VSS	LAVAS BASE y LAVAS TECHO	Basáltico-andesíticas y andesíticas	6 (2-base, 1-Jabalí y 3-El Picacho)	Sofield, 1998; Misión Italiana, 1988 y DF101-102, DF135, DF134	36,000 - 72,000 BP	

Figura 9. Cronología de unidades geológica de complejo volcánico de San Salvador. Fuente: Ferrés, 2014

Entre los depósitos piroclásticos del Boquerón se encuentran la unidad más antigua descrita como depósito de pómez G1, el cual es un producto de la erupción pliniana asociada a la destrucción del antiguo Volcán de San Salvador y se encuentra inmediatamente por encima de la toba superior TB4. Le sigue el depósito gris y de desechos fluviales conocida como G2, la cual es un posible producto de una erupción en un periodo intermedio de formación del El Boquerón, encontrándose por encima de las tobas superiores de TB2 y TB3. Para mayor entendimiento, en la pasada figura 9 se muestra un extracto de las cronologías expuestas.

Adicionalmente en la figura9, únicamente se logran apreciar las edades asumidas de las tobas superiores (Ts), las cuales se asocian a tres episodios eruptivos de la caldera de Ilopango (TB4, TB3 y TB2), obteniéndose depósitos de pómez de caída. Por otro lado, en la figura no se visualiza las edades de las llamadas Tobas inferiores (Ti), las cuales corresponden a las unidades piroclásticas más antiguas registradas de la caldera de Ilopango, componiéndose de una diversidad de depósitos de caída, ignimbritas compactas que varían de colores rosado (Olocuilta), beige y blanco. Finalmente, se encuentra la formación San Salvador que se compone por los productos efusivos y explosivos más recientes del Volcán Boquerón y la Caldera de Ilopango; Talpetate y TBJ, respectivamente.

Asimismo, forman parte de esta formación las erupciones inmediatas que se encuentran por debajo de la erupción de TBJ, figuran tres erupciones bien conocidas. La más antigua (arriba del depósito de G2) se encuentra B1, el cual es un producto de una erupción de menor magnitud anterior a los 3,000 años B. P, asociado al antiguo edificio volcánico del Boquerón correspondiente al resto del paleovolcán después del colapso caldérico, llamada cerro El Picacho. Luego se encuentran flujos de lavas, asociados a erupciones posteriores a la formación del cráter del Boquerón y finalmente se encuentra la erupción del Plan de La Laguna, la cual se asocia a un edificio monogenético y comprende depósitos freatomagmáticos, aparentemente asociada al macizo volcánico del San Salvador, que al entrar en contacto con el agua subterránea produjo una gran explosión.

A continuación, a manera de ejemplo se muestran 5 puntos de afloramientos con su reclasificación y la figura 10 con sus ubicaciones.

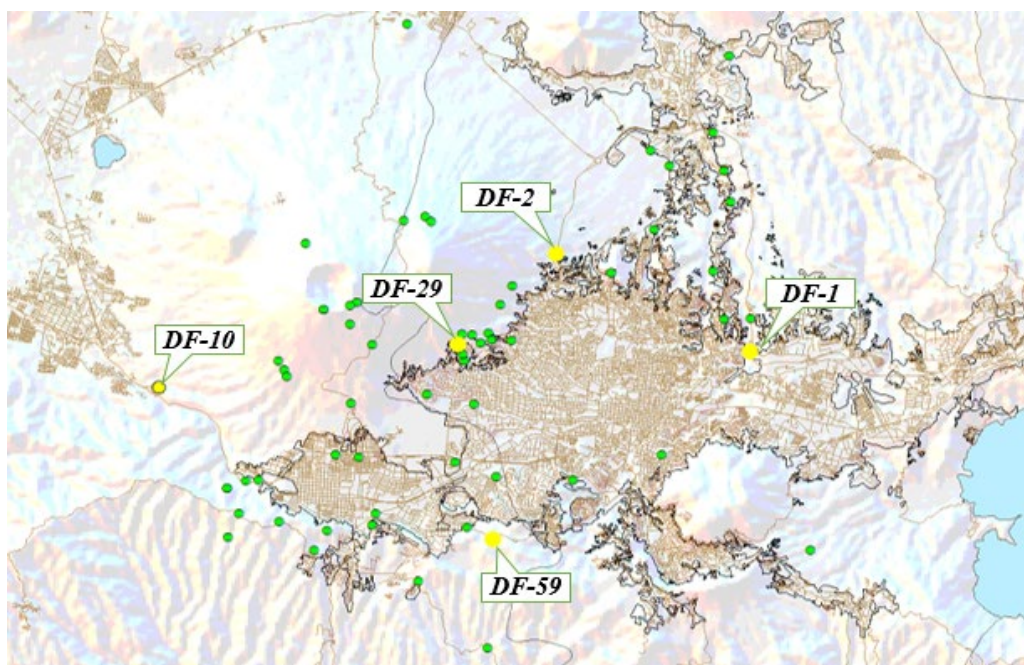
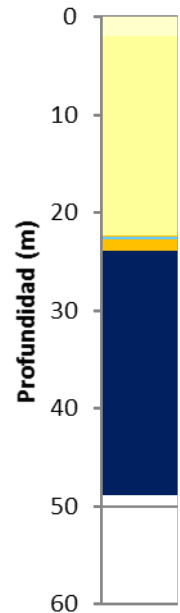


Figura 10. Mapa de ubicación de ejemplo de afloramientos de Dolores Ferrés. Fuente: MARN, 2020

Tabla 5. Descripción de unidades geológicas de afloramiento DF-2 extraído de Ferrés Dolors, 2014

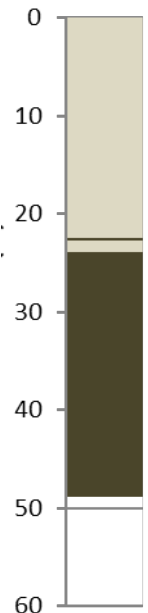
Desde	Hasta	Capa	Espesor	Clasificación	Unidad Italtেকna
0	2	TBJ	2	TBJ	Unidad A
2	3	Deposito G2	1	Ts	Unidad A
3	5.5	TB2	2.5	Ts	Unidad A
5.5	8.5	TB3	3	Ts	Unidad A
8.5	9	Paleosuelo	0.5	Ts	Unidad A
9	10.9	Depósitos G1-0 y G1-1	1.9	Ts	Unidad A
10.9	11.1	Oleada piroclástica G1 (?)	0.2	Ts	Unidad A
11.1	14.6	Deposito G1-1 freatomagnético	3.5	Ts	Unidad A
14.6	15.1	Oleada piroclástica o suelo(?)	0.5	Ts	Unidad A
15.1	15.55	Depósito de caída G1-2	0.45	Ts	Unidad A
15.55	16.05	Oleadas piroclásticas o suelo (?)	0.5	Ts	Unidad A
16.05	16.55	Depósito de caída G1-3	0.5	Ts	Unidad A
16.55	17.8	Paleosuelo	1.25	Ts	Unidad A
17.8	22.3	TB4	4.5	Ts	Unidad A
22.3	22.4	Paleosuelo	0.1	Ti	Unidad A
22.4	22.7	Deposito Apopa	0.3	Escorias	Unidad B
22.7	23.9	Paleosuelo	1.2	Ti	Unidad A
23.9	48.9	Colada de lava basalto-andesítica	25	Basalto	Unidad B

**DF-2
Clasificación**



Profundidad (m)

**DF-2
Unidad**

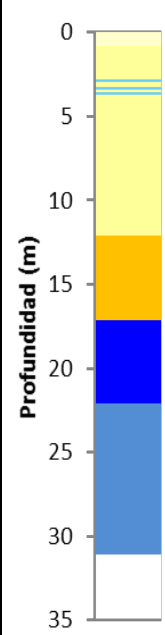


Profundidad (m)

Tabla 6. Descripción de unidades geológicas de afloramiento DF-10 extraído de Ferrés Dolors, 2014

De	Hasta	Capa	Espesor	Clasificación	Unidad Italtেকna
0	0.2	Suelo actual	0.2	TBJ	Unidad A
0.2	0.79	TBJ meteorizada (?)	0.59	TBJ	Unidad A
0.79	2.79	Deposito piroclástico de caída G2	2	Ts	Unidad A
2.79	2.94	Depósito de caída escoriáceo	0.15	Escoria	Unidad B
2.94	3.08	Depósito de ceniza	0.14	Ts	Unidad A
3.08	3.15	Depósito de caída	0.07	Ts	Unidad A
3.15	3.26	Depósito de ceniza	0.11	Ts	Unidad A
3.26	3.41	Depósito de caída escoriáceo	0.15	Escoria	Unidad B
3.41	3.59	Depósito de ceniza a suelo	0.18	Ts	Unidad A
3.59	3.69	Depósito de caída escoriáceo. Parecido a G1-0.	0.1	Escoria	Unidad B
3.69	3.86	Depósito de ceniza a suelo. G1-0.	0.17	Ts	Unidad A
3.86	4.21	Depósito de caída (freatomagmático). Parecido a G1-0	0.35	Ts	Unidad A
4.21	4.26	Oleada piroclástica. Surge G1	0.05	Ts	Unidad A
4.26	4.36	Depósito de caída pómez. Surge G1	0.1	Ts	Unidad A
4.36	4.47	Oleada piroclástica. Surge G1	0.11	Ts	Unidad A
4.47	4.51	Lentes discontinuos de lapilli de caída	0.04	Ts	Unidad A
4.51	5.41	Flujo piroclástico (?)	0.9	Ts	Unidad A
5.41	5.61	Depósito de caída de ceniza alterada	0.2	Ts	Unidad A
5.61	8.81	G1: Deposito de caída (G11)	3.2	Ts	Unidad A
8.81	12.11	TB4	3.3	Ts	Unidad A
12.11	17.11	Varios depósitos de caída y flujo/oleada piroclástico. El más superior es Apopa	5	Ti	Unidad A
17.11	22.11	Depósito de lavas en bloques o debris-flow	5	Lava o lahar	Unidad B
22.11	31.11	Lavas basáltico-andesíticas masivas	9	Andesita	Unidad B

**DF-10
Clasificación**



Profundidad (m)

**DF-10
Unidad**

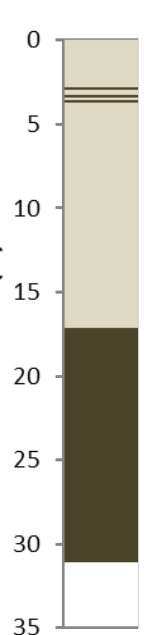


Tabla 7. Descripción de unidades geológicas de afloramiento DF-29 extraído de Ferrés Dolors, 2014

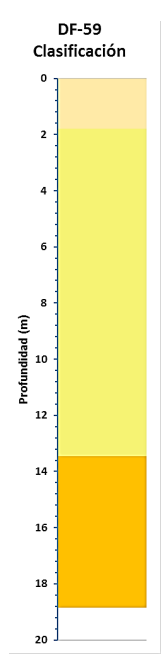
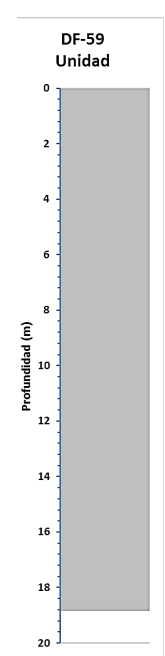
De	Hasta	Capa	Espesor	Clasificación	Unidad Italtেকna	DF-29 Clasificación	DF-29 Unidad
0	0.2	Suelo actual	0.2	TBJ	Unidad A		
0.2	1	TBJ	0.8	TBJ	Unidad A		
1	2	Depósito de ceniza	1	Ts	Unidad A		
2	2.85	Depósito de caída. G1-0	0.85	Ts	Unidad A		
2.85	6.35	Lava basáltica o basáltico-andesítica	3.5	Basalto	Unidad B		
6.35	6.85	TB4	0.5	Ts	Unidad A		
6.85	7.05	Depósito de caída alterado	0.2	Ti	Unidad A		
7.05	7.55	Flujo piroclástico	0.5	Ti	Unidad A		
7.55	12.55	Depósito de Debris - Avalanche	5	Lahar	Unidad B		

Tabla 8. Descripción de unidades geológicas de afloramiento DF-1 extraído de Ferrés Dolors, 2014

De	Hasta	Capa	Espesor	Clasificación	Unidad Italtেকna	DF-1 Clasificación	DF-1 Unidad
0	10	TBJ	10	TBJ	Unidad A		
10	10.2	Paleosuelo	0.2	Ts	Unidad A		
10.2	11.05	TB2	0.85	Ts	Unidad A		
11.05	11.55	Paleosuelo	0.5	Ts	Unidad A		
11.55	12.15	TB3	0.6	Ts	Unidad A		
12.15	12.55	Paleosuelo	0.4	Ts	Unidad A		
12.55	12.95	Deposito G1	0.4	Ts	Unidad A		
12.95	13.65	TB4	0.7	Ts	Unidad A		
13.65	13.85	Paleosuelo	0.2	Ti	Unidad A		
13.85	14.65	Flujo piroclástico	0.8	Ti	Unidad A		
14.65	15.15	Ignimbrita rosada	0.5	Ti	Unidad A		

Tabla 9. Descripción de unidades geológicas de afloramiento DF-59 extraído de Ferrés Dolors, 2014

De	Hasta	Capa	Espesor	Clasificación	Unidad Italtekna
0	0.1	Suelo actual	TBJ	A	0
0.1	1.8	TBJ	TBJ	A	0.1
1.8	2.5	G2	Ts	A	1.8
2.5	3.5	Depósito de caída	Ts	A	2.5
3.5	3.85	Paleosuelo	Ts	A	3.5
3.85	5.2	TB2	Ts	A	3.85
5.2	6.3	Colada piroclástica	Ts	A	5.2
6.3	6.9	Depósito de ceniza	Ts	A	6.3
6.9	7.55	Paleosuelo	Ts	A	6.9
7.55	8.25	Depósito de ceniza	Ts	A	7.55
8.25	8.75	G1	Ts	A	8.25
8.75	10.25	TB4	Ts	A	8.75
10.25	13.45	TB4	Ts	A	10.25
13.45	18.85	Depósito de caída	Ti	A	13.45

5.3.2. Plataforma de Sistema de Información Hídrica (SIHI)

Dado que todos los perfiles de la plataforma superan los 30 metros de profundidad, se hizo tanto la reclasificación de los espesores como la asignación de la velocidad de ondas de corte de acuerdo a los tres criterios mencionados anteriormente. Para el segundo criterio, donde se utilizaría la tabla No.2 de clasificación de suelos según la NEHRP 2020, se vio necesario construir de nuevo la tabla, tratando de colocar los espesores de los productos volcánicos propuestos por el MARN 2011 a los diferentes tipos de suelos.

Es importante destacar, que estas descripciones fueron asignadas mediante discusiones con geólogos y con la revisión inmediata de los resultados de velocidad de ondas de corte obtenidos en los 15 pozos perforados. Cabe mencionar, que un criterio importante para asignar las velocidades, es el **estado presión-dependiente en el que se encuentra una unidad geológica**, es decir, que entre más profundo se encuentre el estrato, más compactado estará, incrementando así el esfuerzo efectivo sobre los agregados del material; sumándole a esto, en los casos que el nivel freático coexista dentro de la columna de suelo.

Lo anterior, influye consecuentemente en la respuesta sísmica del suelo, donde sus propiedades dinámicas (amortiguamiento y módulo de corte) se verán modificadas; ya que es en el estado natural en el que se encuentra, con el que esperará las ondas sísmicas.

A continuación, en la tabla 10 se muestra la reclasificación de las clases de suelo junto a las descripciones detalladas asociadas con los orígenes eruptivos y productos volcánicos.

Tabla 10. Clasificación de suelos según NEHRP 2020 junto a reclasificaciones de MARN 2020

No	Reclasificación de Clases de Suelo					Criterio Geotécnico (prom. de 30 m superiores)		
	Tipo Clase 2016 NEHRP	Tipo Clase 2020 NEHRP	Descrip. general	Descripción detallada	Productos volcánicos DOMINANTES en columna de suelo según MARN 2011 y discusiones con geólogos	Vs30 (m/s)		Vs30 (m/s) Medi a GM
1	A	A	Roca dura	Roca dura	Lavas sanas (Andesíticas o Basálticas)	> 1500		
2	B	B	Roca media	Roca media	Lavas semi-fracturadas, Flujo de detritos con clastos lávicos, Lahares cementados con clastos lávicos, Avalancha de escombros	910	1500	1160
3		BC	Roca blanda	Roca blanda	Lavas fracturadas meteorizadas, Rocas Dacíticas, Lahares con escorias y fragmentos lávicos, Depósito de flujos piroclásticos, ej.: Boquerón (G1)	640	910	760
4	C	C	Suelo duro	Suelo muy denso o arcilla dura.	Lavas fracturadas bien meteorizadas, Ignimbritas máficas meteorizadas, Lahares con escorias, Tobas aglomeradas, Tobas inferiores (Productos antiguos de TB4 de la caldera de Ilopango)	440	640	530
5		CD	Suelo medio duro	Suelo de arena densa o arcilla muy rígida.	Coluvios en profundidad, Sedimentos Aluviales, Lahares meteorizados, Ignimbritas de Boquerón (G2), Tobas Superiores en profundidad (Productos antiguos de TBJ de la caldera de Ilopango)	300	440	360
6	D	D	Suelo poco duro	Suelo de arena medianamente densa o suelo arcilloso rígido	Tobas Superiores superficiales (Productos antiguos de TBJ de la caldera de Ilopango) Tierra Blanca Joven proximal en profundidad, Lahares bien meteorizados	210	300	250
7		DE	Suelo blando	Suelo de arena suelta o suelo arcilloso medio rígido	Tierra Blanca Joven distal superficial, Suelos plásticos meteorizados, Suelos retrabajados con bloques de rocas y granos gruesos.	150	210	180
8	E	E	Suelo muy blando	Suelos de arena suelta o suelo arcilloso blando	Rellenos no compactados, Suelos orgánicos, Sedimentos fluviales o retrabajados de grano fino y Suelos bien meteorizados con arcillas blandas	< 150		
9	F	F	Suelos que requieren análisis de respuesta de sitio					

Luego de realizada la reclasificación de los tipos de suelo, se ha procedido a obtener el valor de la velocidad de corte para los 30 metros superiores (Vs30) de cada pozo. Para ello fue necesario la profundidad de cada capa y la velocidad de corte de cada material. En ocasiones, por sugerencia del geólogo, se colocaba valores mínimos o máximos de Vs dentro del rango en que se encontraba el tipo de suelo y no necesariamente valores medios.

No obstante, debido a que en las descripciones de las unidades geológicas de los pozos de la plataforma SIHI no se reconocían con exactitud, específicamente los productos de la caldera de Ilopango: las Ts (Tobas superiores) y las Ti (Tobas inferiores), se procedió a localizar en el mapa geológico la edad en la que se encontraba cada uno de los pozos.

Junto a lo anterior, se auxilió con la base estratigráfica de las formaciones geológicas realizada por el geólogo Walter Hernández en el año 2004, por lo que se definió que las Ti serían aquellas que se encontrarán por debajo del miembro S3'a (Tobas color café, de la formación de San Salvador), mientras que las Ts serían aquellas que se encontrarán por encima de ella, es decir desde el miembro S3'a hasta Qf (depósitos sedimentarios del cuaternario, de la formación de San Salvador).

Sin embargo, no en todo el país las Ts se encuentran a una misma profundidad, al igual que la TBJ no se encuentra en todo el país y los espesores son diferentes. A continuación, se presentan las estimaciones de la profundidad de las Tobas superiores realizado a partir de cortes geológicos en perfiles sobre el área central del departamento de San Salvador, así como también con la ayuda de estudios que fueron realizados en dicha área.

5.3.2.1. Estimación de la Ts en el Valle de Zapotitán

La profundidad de las tobas superiores no es muy bien conocida con exactitud, ya que no hay información de estudios geológicos en la zona; por lo que al discutir con el geólogo Walter Hernández se sugirió estimar la profundidad a 60 metros (Ver figura 11). Lo anterior, debido a que en el valle es muy común la deposición de sedimentos, enterrando a esa profundidad, las primeras tobas consideradas como Ts.

Seguidamente, se procedió a realizar isolíneas representando curvas de nivel con igual profundidad. Para realizar las curvas se basó en los mapas geomorfológicos y en el mapa geológico 1:100 mil de El Salvador.

De igual manera, se utilizaron perfiles topográficos para conocer cómo cambian las elevaciones y el relieve en dirección hacia el valle. Respecto al noroeste del volcán San Salvador, también se decidió que las tobas superiores tuvieran esa profundidad de 60 metros considerando que por debajo estarían los depósitos potentes de la avalancha de escombros del Antiguo Volcán de San Salvador, correspondiente su primer colapso.

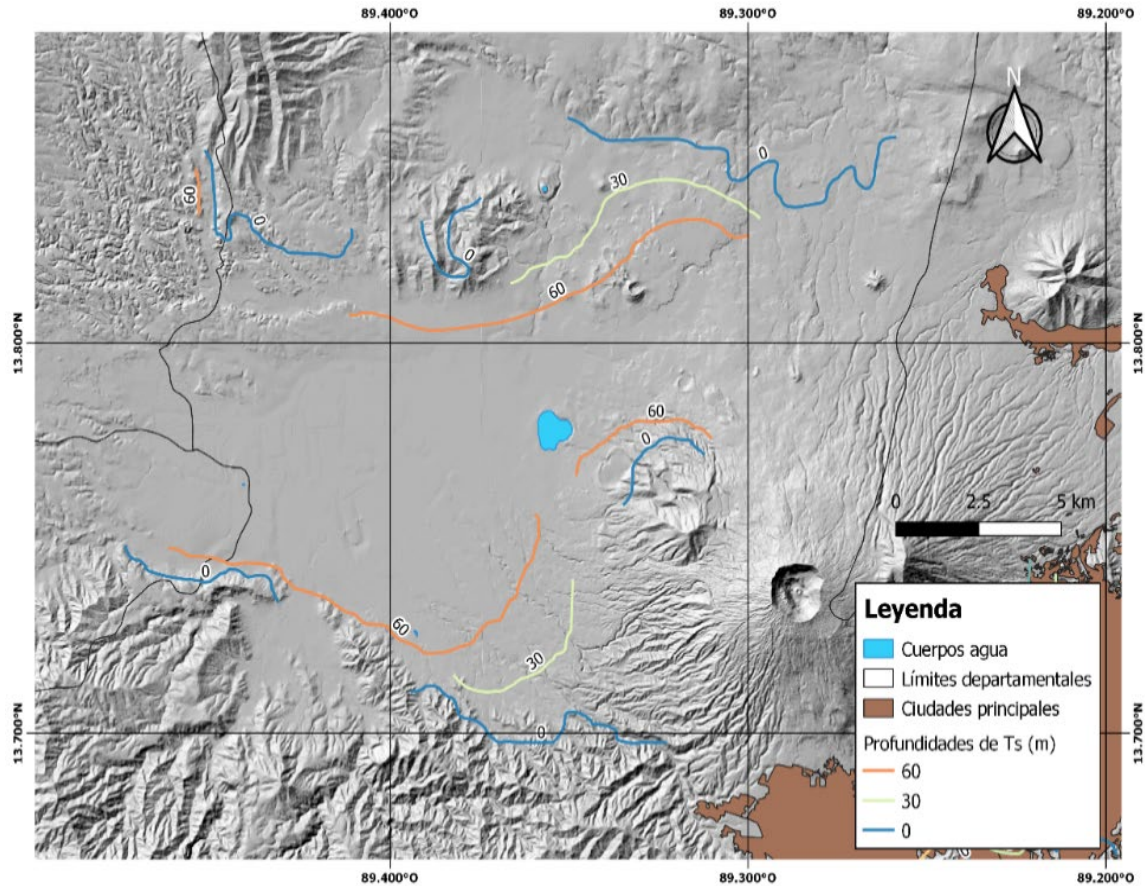


Figura 11. Estimado de iso-profundidades de tobas superiores en valle de Zapotitán. Fuente: MARN, 2020

5.3.2.2. Realización de cortes geológicos en el AMSS

Para estimar las profundidades a las tobas superiores con más precisión se recurrió a la elaboración de siete cortes geológicos (ver figura 12) a partir del nuevo mapa geológico del AMSS, a una escala de 1: 50,000, el cual fue levantado en el marco del proyecto “Gestión de Riesgos y Disminución de Vulnerabilidad Social en el Área Metropolitana de San Salvador” financiado por la Agencia Andaluza de Cooperación Internacional para el Desarrollo (AACID) y ejecutado por el COAMSS/OPAMSS, en el año 2019.

Las unidades litológicas con sus respectivas formaciones y subsecuentes miembros, siguen lo estipulado en la Guía Estratigráfica Internacional (1994). Este novedoso mapa geológico ha sido realizado principalmente para proporcionar información precisa para la evaluación de riesgos geológicos. La razón de ocupar este mapa es debido a que en ciertas zonas posee mejor detalle reuniendo lo necesario para poderse realizar el perfil de la sección transversal (ver figura 13). Las condiciones para realizar los perfiles son los siguientes:

- Los cauces de las quebradas y/o ríos corten las TBJ y tobas superiores, siendo este último, el que más interesa para determinar la profundidad de su base.

- Que existan un mínimo de dos cauces que corten las tobas mencionadas antes, dando una profundidad confiable para el inicio de las tobas superiores. Ts)
- De solo existir un cauce, debe de haber cerca afloramientos de tobas inferiores, marcando un límite en la estimación en lateral de la profundidad de las Ts.

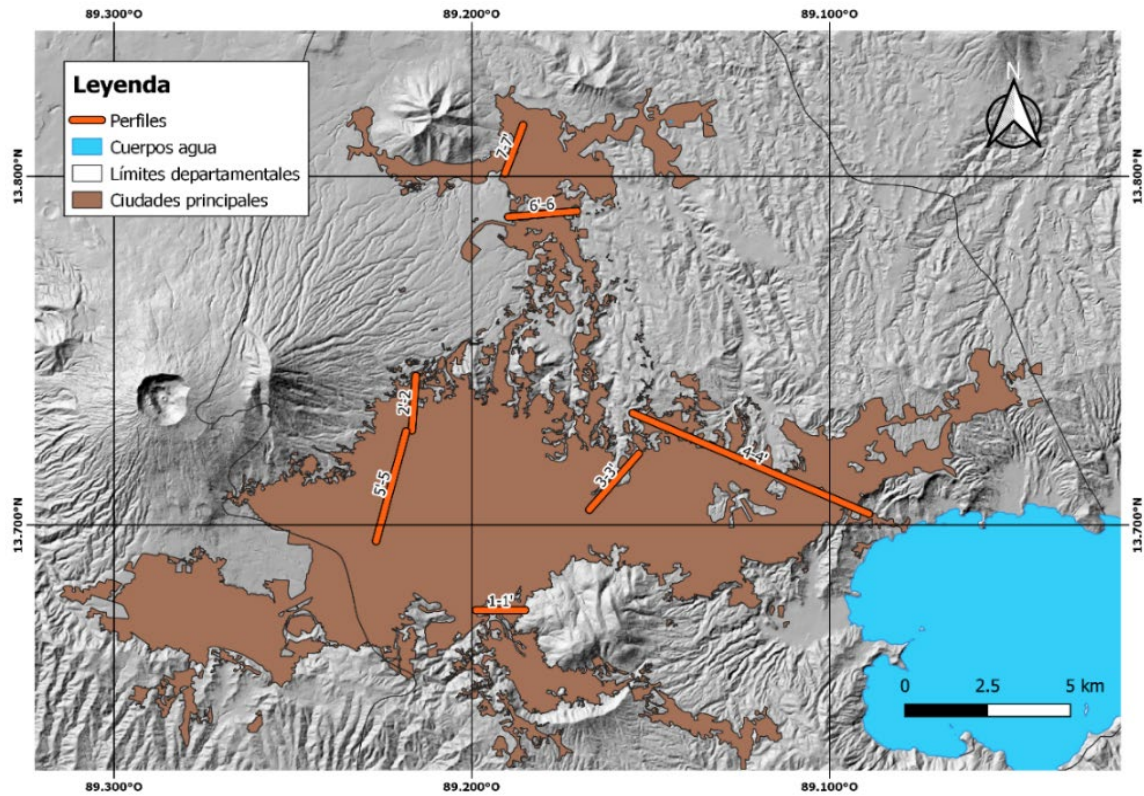


Figura 12. 7 Perfiles realizados en el Área Metropolitana de San Salvador. Fuente: MARN, 2020

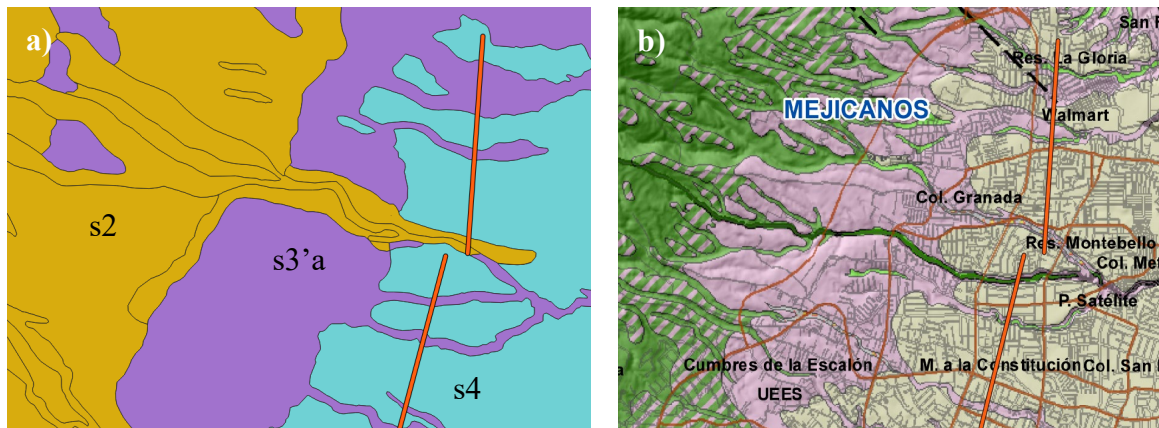









Figura 13. Comparación de mapas geológicos junto dos de los perfiles realizados. a) Mapa geológico “AMSS-15k” y b) Mapa geológico nuevo del AMSS 1:50,000 de OPAMSS. Fuente: MARN, 2020

La litología ocupada para identificar el techo de las tobas inferiores consistió en identificar cualquiera de las siguientes unidades mostradas en la figura 14.

Formación San Salvador/Cuscatlán

-  Volcán Antiguo de San Salvador (Picacho)-complejo estratovolcánico de la zona proximal
-  Volcán Antiguo de San Salvador (Picacho)-complejo estratovolcánico de la zona proximal cubierto por tefras y talpetate
-  Volcán Antiguo de San Salvador (Picacho)-complejo estratovolcánico de la zona proximal cubierto por tefras
-  Volcán Antiguo de San Salvador (Picacho)-aglomerados, escorias y tobas de la zona proximal y media
-  Aglomerados, escoria y tobas de Volcán Antiguo de San Salvador, cubierto por tefras y talpetate
-  Volcán Antiguo de San Salvador (Picacho)-flujos de lava de la zona media
-  Volcán Antiguo de San Salvador (Picacho)-predominan complejos epiclasticos de la zona media

Formación Cuscatlán


















-  Tobas antiguas de calderas de Ilopango y Coatepeque
-  Facies retrabajadas (fluvial/lacustrina) de tobas antiguas de calderas de Ilopango y Coatepeque
-  Brechas Subacuaticas y tobas pomez retrabajada
-  Cono de escoria dacitica San José
-  Cono efusivo andesítico/dacitico San Jacinto
-  Cono efusivo andesítico/dacitico San Jacinto cubiertos por tefras
-  Domos extrusivos rioliticos
-  Domos extrusivos daciticos-San Jacinto y equivalentes
-  Brechas relacionadas a domo dacitico
-  Flujos piroclasticos relacionados a domo dacitico
-  Domos extrusivos daciticos, cubiertos por tefras
-  Complejos de estratovolcanes/efusivos basalticos/andesíticos de la zona proximal-Loma Larga, Nejapa, Guaycume, Guazapa y El Carmen
-  Complejos de estratovolcanes/efusivos basalticos/andesíticos de la zona proximal cubiertos por tefras
-  Complejos de flujos de lava basaltica y brechas hialoclastitas-Loma Larga, Nejapa, Guaycume y Guazapa
-  Complejos de flujos de lava basaltica y brechas hialoclastitas cubiertas por tefras
-  Complejo de flujos de lava basalticas y brechas epiclasticas volcánicas de zona media
-  Ignimbritas de la caldera de Jayaque y rocas volcánicas epiclasticas relacionadas

Figura 14. Simbología de mapa geológico 1:50 mil de OPAMSS; para asociar las Ti. Fuente: MARN, 2020

En la figura 15 se muestran las unidades identificadas como Tobas superiores y TBJ.

Formación San Salvador
















-  Aglomerados/tobas de anillo basaltico-zona proximal;Plan de la Laguna y la Tabla
-  Facies distales de tefras freatomagmaticas con tefras Tierra Blancas y volcán Boquerón Subordinadas
-  Tobas freatomagmáticas del volcan Boquerón- Talpetate
-  Tefras de volcán Boquerón cubiertas por tobas freatomagmaticas (Talpetate)
-  Conos de escoria basaltica
-  Flujos de lava basaltica/andesítica y complejos/campos de flujo
-  Flujos de lava basaltica/andesítica y complejos/campos de flujo, cubiertos por tefras
-  Volcan Boquerón (San Salvador joven) predominantemente cono efusivo
-  Volcan Boquerón (San Salvador joven) predominantemente cono efusivo y flujos de lava, cubiertos por escoria/tefras
-  Volcan Boquerón (San Salvador joven) predominantemente cono efusivo, cubierto por tefras y talpetate
-  Tefras rioliticas Tierra Blanca Joven incluidas tefras de pomez
-  Facie retrabajada de tefras rioliticas iniciales de Tierra Blanca Joven
-  Tierra Blanca Joven depositos singeneticos de flujos de escombros
-  Tefras Tierra Blanca Joven yaciendo sobre tefras de volcán Boquerón y tobas de maar Plan de la Laguna
-  Tefras Tierra Blanca Joven yaciendo sobre tefras de volcán Boquerón y tobas de maar la Tabla

Figura 15. Simbología de mapa geológico 1:50 mil; para asociar Ts y TBJ. Fuente: MARN, 2020

La realización de los perfiles consistió en el siguiente procedimiento.

- Se trazaron alineamientos horizontales sobre el nuevo mapa geológico del AMSS y el modelo de elevación digital (obtenido de información LIDAR, a partir de curvas a cada 5 metros, del MARN); tratando de cumplir los criterios mencionados arriba. A partir de esto se obtuvo un perfil de sección transversal.
- Se tomó una captura de la zona, la cual se cortaba y giraba haciéndola encajar con las dimensiones del perfil de sección transversal, para luego ubicar la zona cortada por el río o quebrada que exhibe las tobas inferiores (específicamente con la zona en forma de “V” del perfil), esto debido a que no están las capas perfectamente una sobre otra, existiendo a veces pequeños desfases.
- Se comienza por proyectar los límites litológicos en el perfil de sección transversal por medio de líneas (ver figura 16), para luego considerar la litología cercana (a partir de los pozos de SIHI) para poder estimar cómo sería su extensión lateral y tomando en cuenta la litología que atraviesa el río como un punto de control.
- Se colocan las etiquetas, profundidades y fallas inferidas en el mapa (Ver figura 17).

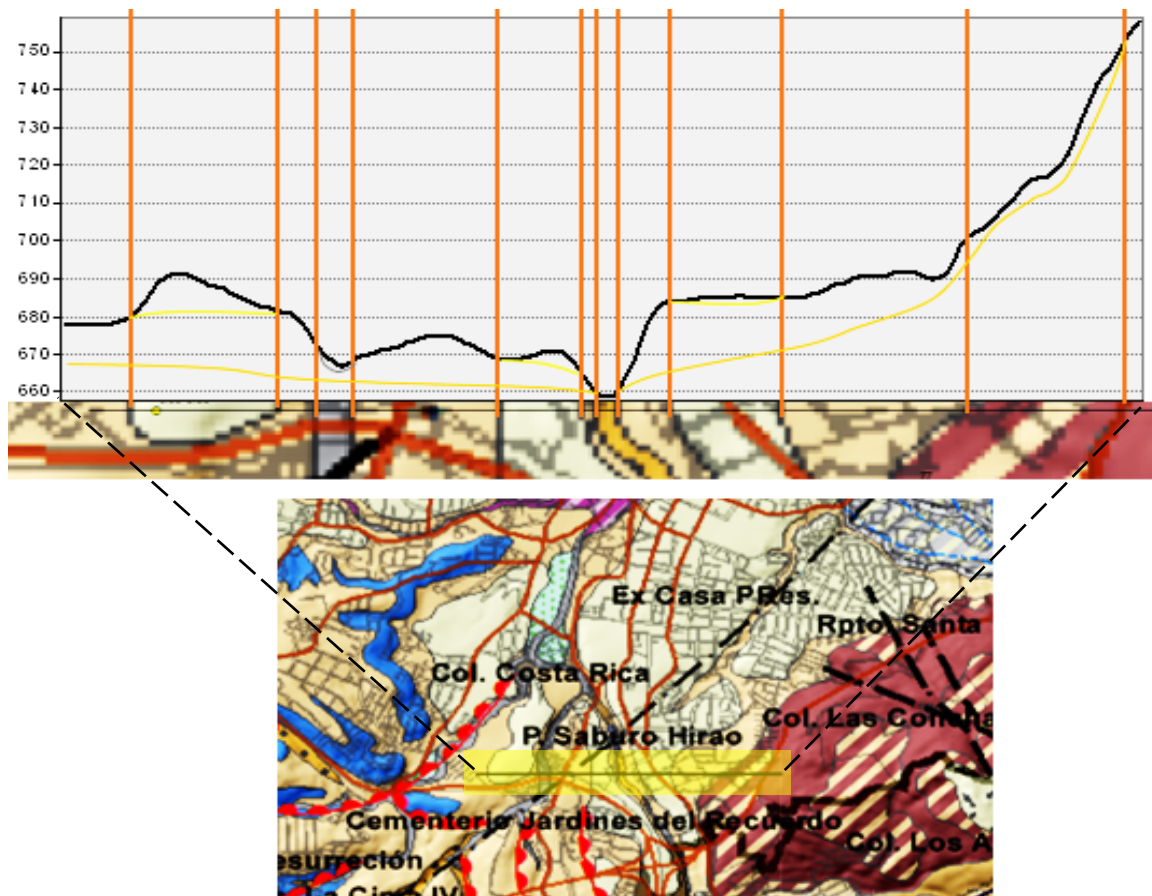


Figura 16. Ejemplo de realización de un corte geológico y Construcción del perfil. Fuente: MARN, 2020

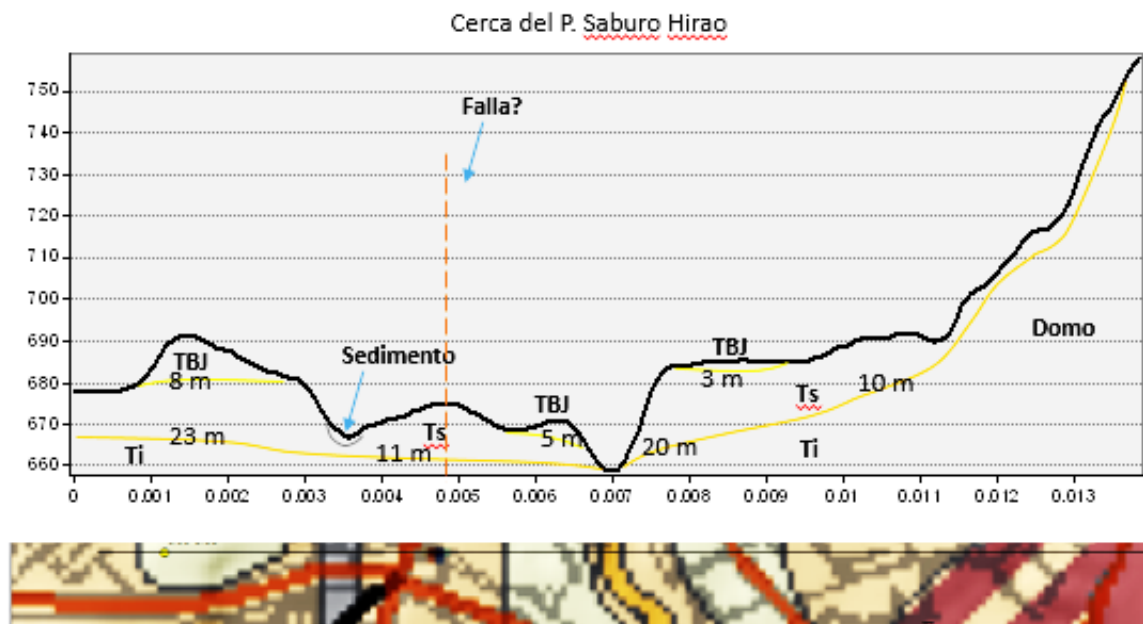
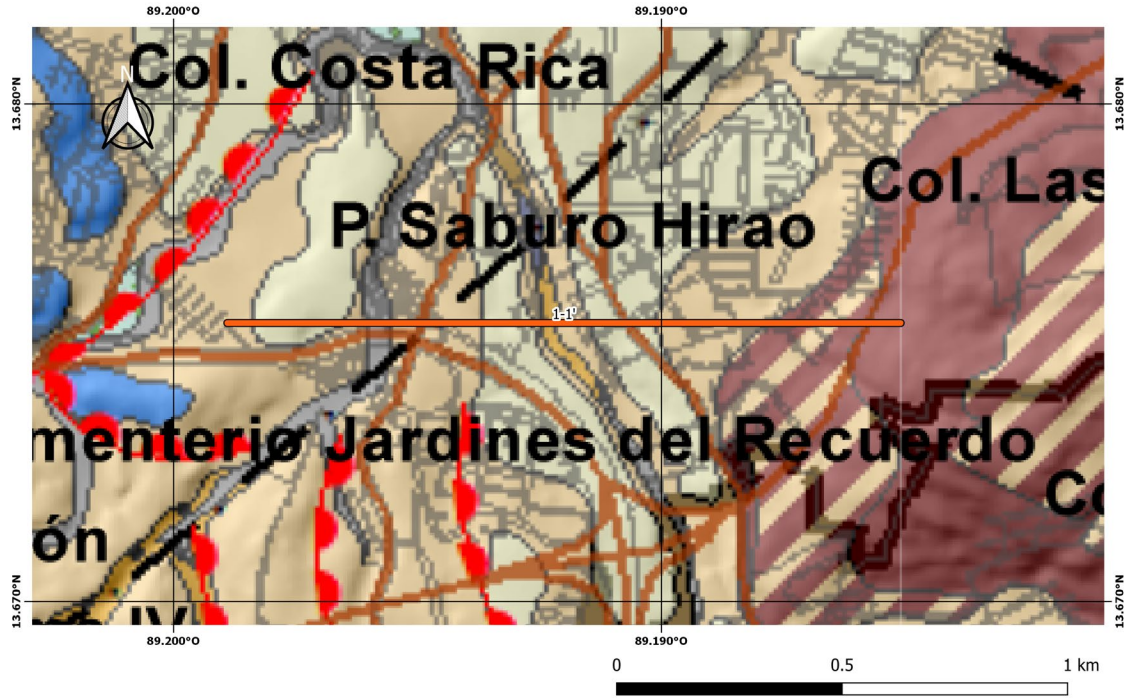


Figura 17. Ejemplo de realización de un corte geológico y perfil con su simbología. Fuente: MARN, 2020
 Los contactos litológicos se representaron de la siguiente forma: Contacto TBJ-Ts (Amarillo), contacto Ts-Ti, Ti y sedimentos, domos (Anaranjado) y contacto Sedimento- TBJ, Ts, Ti (Gris).

En la siguiente tabla No 11 se muestra información general de los siete perfiles hechos, como su ubicación, coordenadas (iniciales y finales) y la longitud del perfil (en metros). Desde la figura 18 hasta la 24, se muestran los siete cortes geológicos y sus respectivos perfiles de sección transversal, junto a una breve explicación de las profundidades promedio de las Ts.

Tabla 11. Información general de los siete cortes geológicos dentro del AMSS

Perfil	Ubicación	Inicio		Final		Longitud (m)
		LAT (°)	LONG (°)	LAT (°)	LONG (°)	
1-1'	Cerca del parque Saburo Hirao	13.6756	-89.1989	13.6756	-89.1851	1540.54433
2-2'	Almacén Walmart, La Gloria	13.7428	-89.2157	13.7271	-89.2168	1745.89194
3-3'	Cerca de FUNDASAL	13.7044	-89.1672	13.7205	-89.1534	2367.57138
4-4'	Reparto Las Cañas	13.7321	-89.1551	13.7032	-89.089	8030.1718
5-5'	Miralvalle	13.7269	-89.2184	13.6954	-89.2268	3625.08915
6-6'	Apopa Sur	13.79	-89.1706	13.7884	-89.1899	2148.68164
7-7'	Apopa Norte	13.8008	-89.1908	13.8147	-89.1857	1641.60248



Perfil 1-1'

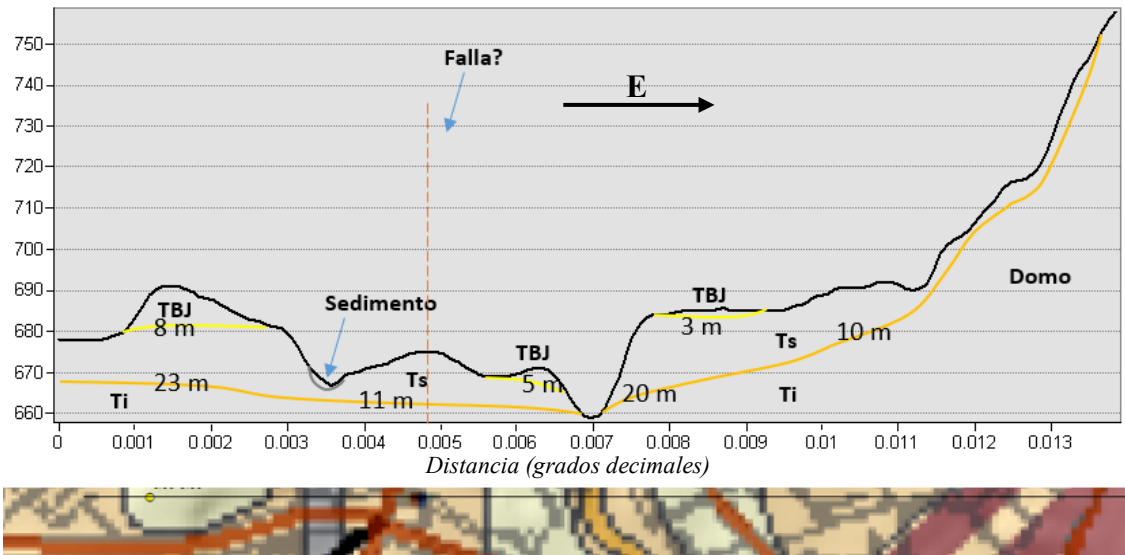
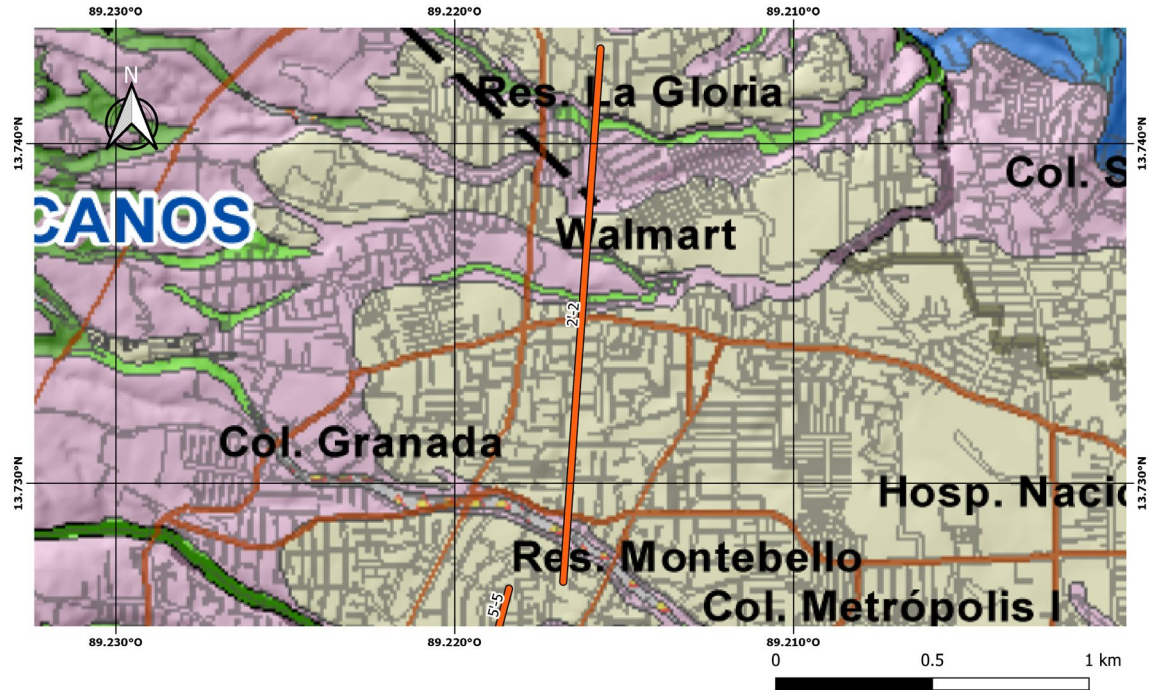


Figura 18. Perfil 1-1' cerca del parque Saburo Hirao, San Salvador. Fuente: MARN, 2020

Los contactos litológicos se representaron de la siguiente forma: Contacto TBJ-Ts (Amarillo), contacto Ts-Ti, Ti y sedimentos, domos (Anaranjado) y contacto Sedimento- TBJ, Ts, Ti (Gris).

El primer perfil 1-1' ubicado cerca del parque Saburo Hirao demuestra que en promedio las tobas superiores se podrían encontrar desde los 15-20 m de profundidad, las litologías con bandeo se consideraron que es una cubierta de tobas con poco espesor. La TBJ posee espesores de menos de 5 m excepto en un punto de ITALTEKNA al oeste del perfil que menciona que hay 8 m de esta última gran erupción del Ilopango.



Perfil 2-2'

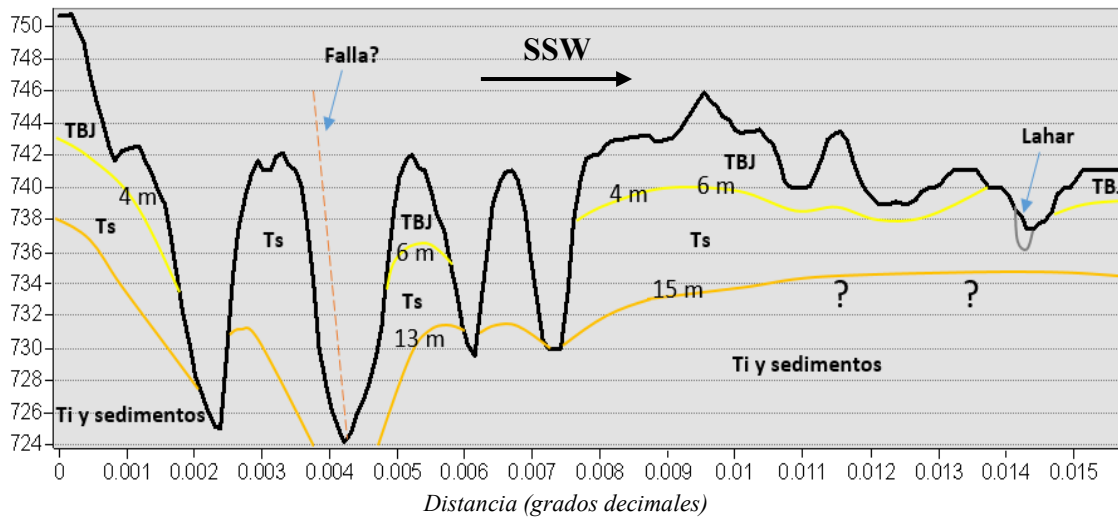


Figura 19. Perfil 2-2' cerca del almacén Walmart-La Gloria, Mejicanos. Fuente: MARN, 2020
 Los contactos litológicos se representaron de la siguiente forma: Contacto TBJ-Ts (Amarillo), contacto Ts-Ti, Ti y sedimentos, domos (Anaranjado) y contacto Sedimento- TBJ, Ts, Ti (Gris).

El perfil 2-2' ubicado en "Walmart, La Gloria" demuestra que en promedio las tobas superiores se podrían encontrar desde los 15 m de profundidad y que la Tierra Blanca Joven tendría un espesor de 4-6 m. En la zona existiría una falla normal en el que las Ts contarían con un menor espesor en el bloque más elevado.

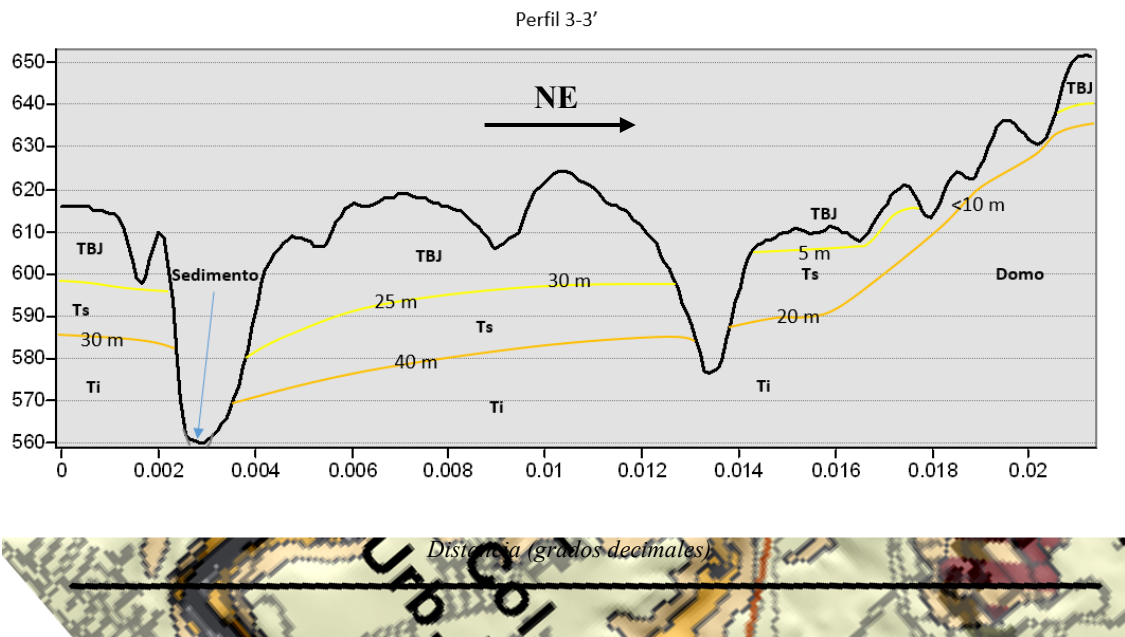
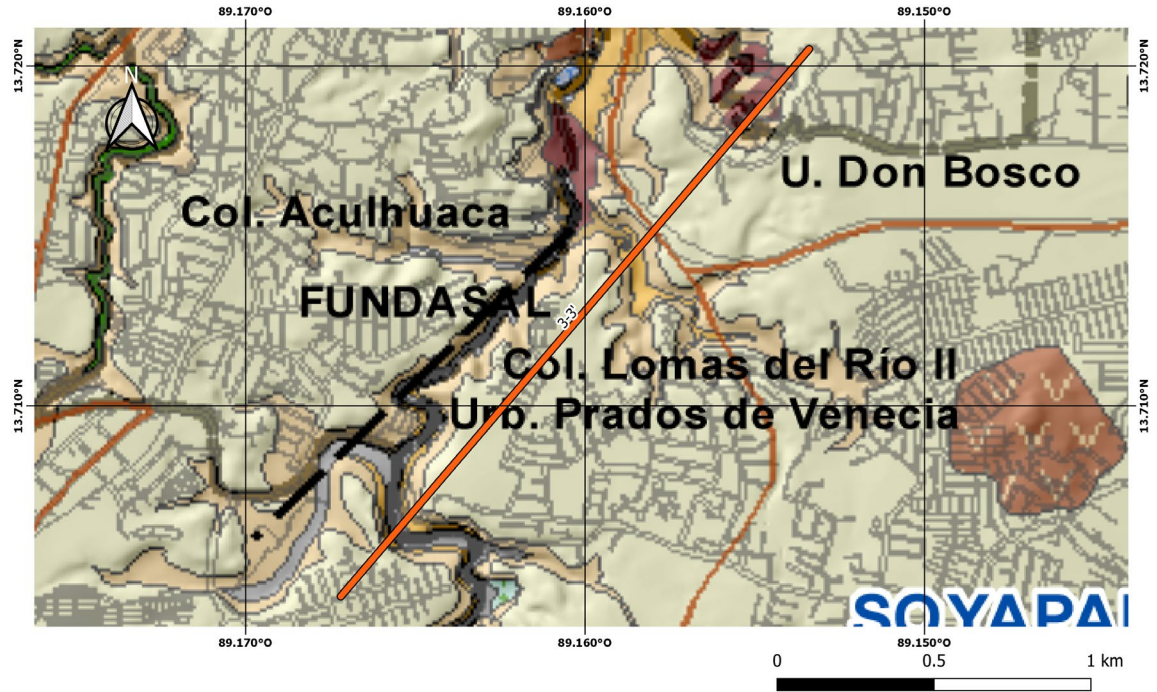


Figura 20. Perfil 3-3' cerca de FUNDASAL, Soyapango. Fuente: MARN, 2020

Los contactos litológicos se representaron de la siguiente forma: Contacto TBJ-Ts (Amarillo), contacto Ts-Ti, Ti y sedimentos, domos (Anaranjado) y contacto Sedimento- TBJ, Ts, Ti (Gris).

El perfil 3-3' ubicado cerca de "FUNDASAL" indica que en promedio las tobas superiores se encontrarían desde los 30 m a la izquierda del corte, a 40 m de profundidad en el centro y con disminución de 20->10 m hacia la derecha, producto de un domo de lava, la Tierra Blanca Joven poseería un espesor de 25-30 m y disminuyendo a la derecha hasta 5-10 m.

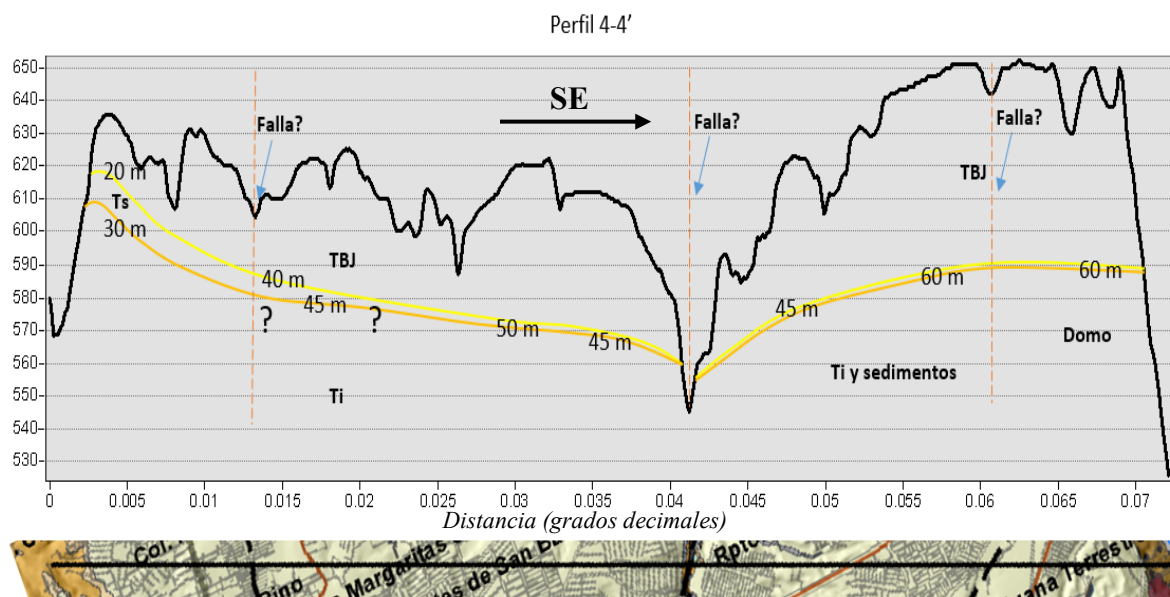
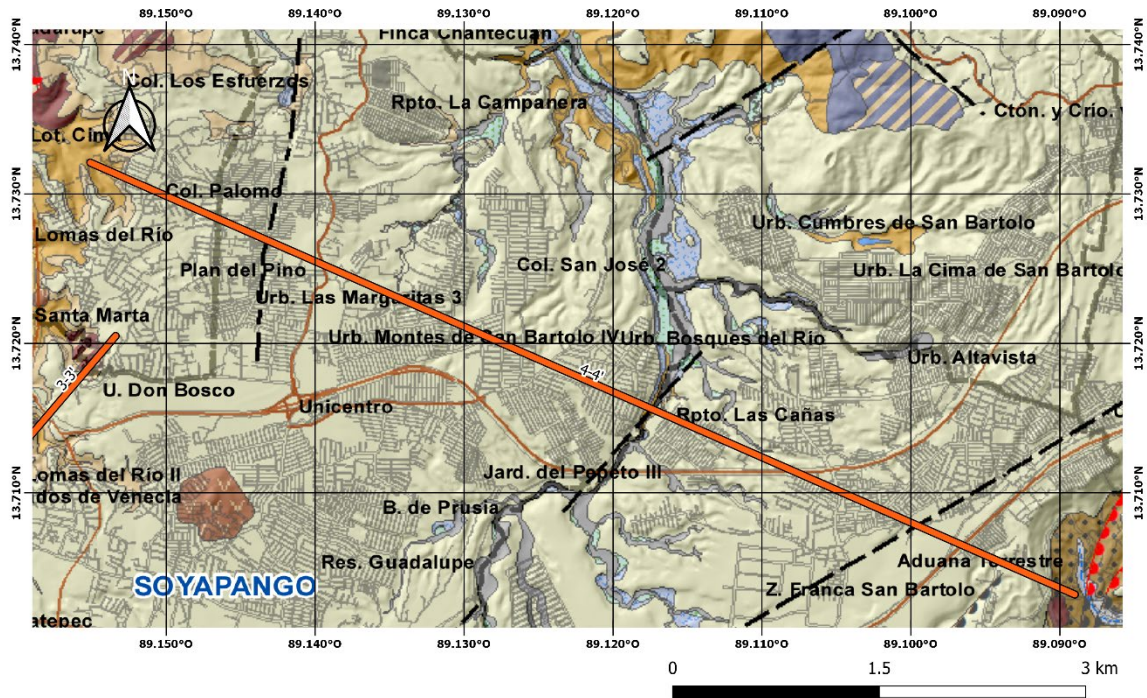


Figura 21. Perfil 4-4' en Reparto Las Cañas Ilopango. Fuente: MARN, 2020

Los contactos litológicos se representaron de la siguiente forma: Contacto TBJ-Ts (Amarillo), contacto Ts-Ti, Ti y sedimentos, domos (Anaranjado) y contacto Sedimento- TBJ, Ts, Ti (Gris).

El perfil 4-4' ubicado en "Reparto Las Cañas" muestra que en promedio las Ts se encontrarían a 30 m de profundidad haciéndose más profundas (disminuyendo su espesor) hasta que en el resto del perfil ocuparía un espesor pequeño o nulo, mientras que la TBJ aumentaría de 20 hasta 60 m, que correspondería con el borde de la caldera de Ilopango.

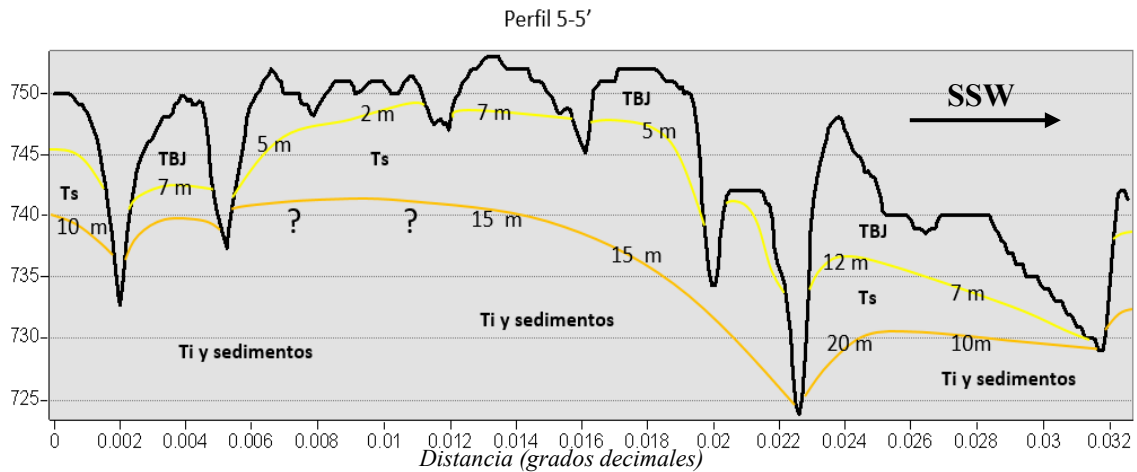
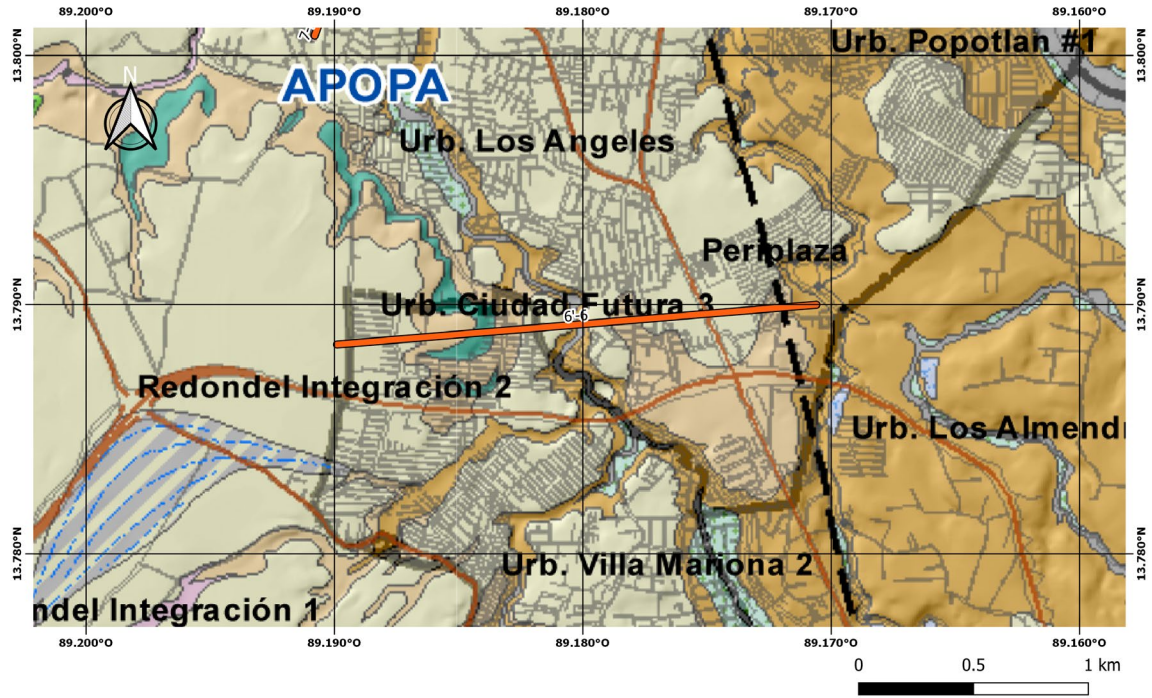


Figura 22. Perfil 5-5' en colonia Miralvalle, San Salvador. Fuente: MARN, 2020

Los contactos litológicos se representaron de la siguiente forma: Contacto TBJ-Ts (Amarillo), contacto Ts-Ti, Ti y sedimentos, domos (Anaranjado) y contacto Sedimento- TBJ, Ts, Ti (Gris).

El perfil 5-5' ubicado en "La Miralvalle", comienza cerca del final del perfil 2-2' y termina poco después de la plaza Salvador del Mundo, en ese mismo orden, las tobas superiores alcanzarían 7 m de profundidad con aumentos a 10 m en ambos lados del perfil, mientras en la zona de la Miralvalle (al centro) alcanza los 15 m. La TBJ poseería en general 5-7 m de espesor.



Perfil 6-6'

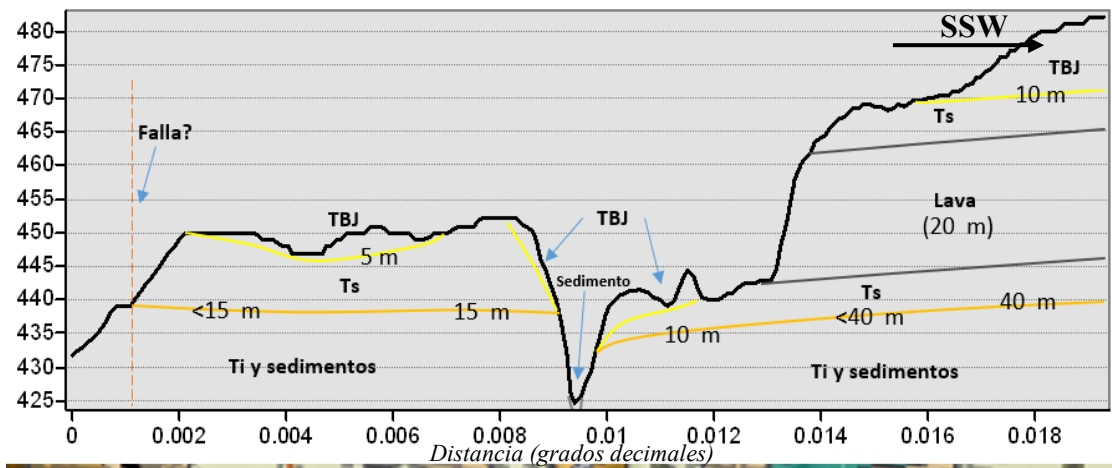
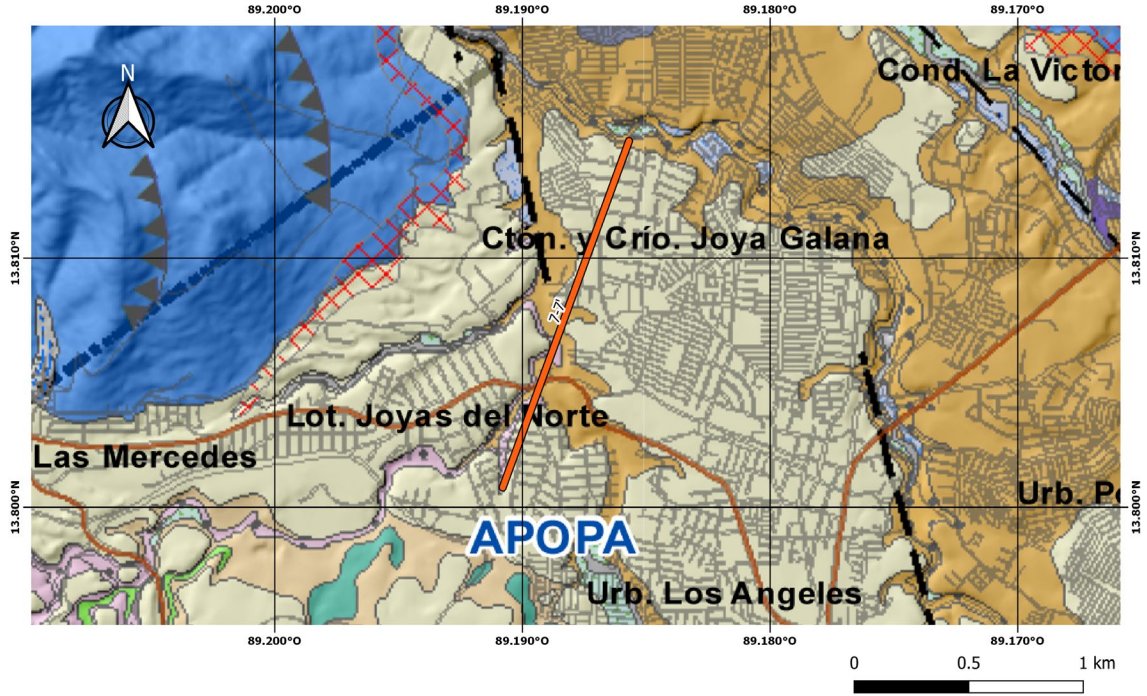


Figura 23. Perfil 6-6' en municipio de Apopa Sur en Ciudad Futura, San Salvador. Fuente: MARN, 2020
 Los contactos litológicos se representaron de la siguiente forma: Contacto TBJ-Ts (Amarillo), contacto Ts-Ti, Ti y sedimentos, domos (Anaranjado) y contacto Sedimento- TBJ, Ts, Ti (Gris).

El perfil 6-6' ubicado en "Apopa sur", en las cercanías de Urbanización Ciudad Futura, se muestra que en general las tobas superiores comenzarían a 15 m de profundidad en el lado izquierdo del río ubicado en el centro del perfil, en dirección opuesta a la flecha arriba del perfil ya no hay mapeadas Ts, en el lado derecho del cauce la base de las tobas superiores pasa de 10 m de profundidad a 40 m debido a un flujo de lava del volcán San Salvador de 20 m de espesor y al aumento de grosor en la TBJ.



Perfil 7-7'

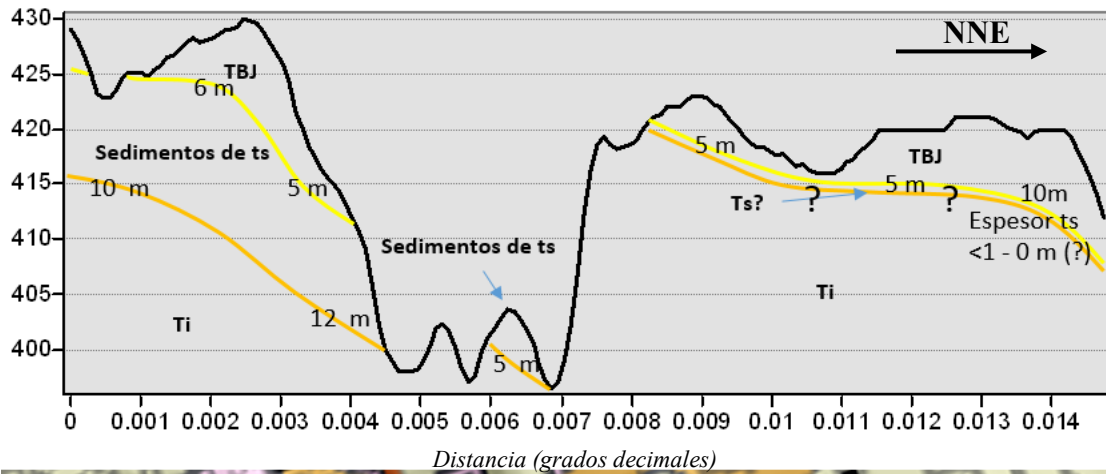


Figura 24. Perfil 7-7' en municipio de Apopa Norte en Joya Galana, San Salvador. Fuente: MARN, 2020
 Los contactos litológicos se representaron de la siguiente forma: Contacto TBJ-Ts (Amarillo), contacto Ts-Ti, Ti y sedimentos, domos (Anaranjado) y contacto Sedimento- TBJ, Ts, Ti (Gris).

El perfil 7-7' ubicado en "Apopa norte", muestra que los sedimentos de tobas superiores se podría situar su base de 10-12 m de profundidad mientras que en el lado derecho las tobas superiores podrían aparecer, pero con un espesor pequeño, esta zona se corresponde con el área izquierda del perfil 6-6' en el que desde ese punto va disminuyendo su espesor. La TBJ poseería, en general, un espesor de 5 m.

5.3.2.3. Estudios para la estimación de la profundidad de la Ts en la zona de Nejapa

Para poder estimar las profundidades con los datos obtenidos por los perfiles realizados, se correlacionaron los datos de diferentes estudios en los cuales se presentan modelos de las profundidades del miembro geológico de la formación San Salvador S3'a. A continuación, se presenta la reseña de cada uno estos.

- Hidrogeofísica de la cuenca del río San Antonio, El Salvador: Aplicación al modelo conceptual y vulnerabilidad intrínseca.** *Realizado por Laura Gil Urrutia & Mario E. Arias Salguero, publicado en la Revista Geológica de América Central en el año 2007. (Gil Urrutia & Arias Salguero, 2007)*

La zona de estudio se ubica en el municipio de Nejapa, departamento de San Salvador. La cuenca que contiene al río San Antonio tiene un área de 57.6 km², limitada al norte con el cerro Ojo de Agua y loma el Quebracho, al oeste con loma Las Mesas, al sur con el volcán de San Salvador, y al este con el cerro de Nejapa.

En la figura 6, se muestra la zona de estudio con la localización de los sondeos y los perfiles geo-eléctricos. El modelo conceptual del sistema acuífero de la cuenca del río San Antonio se determinó a partir de la interpretación de los sondeos eléctricos verticales (SEV). Se modelaron tomando en cuenta las variables como: geología aflorante, niveles estáticos medidos en los pozos, información de columnas litológicas de pozos perforados, pruebas de infiltración y observaciones realizadas en campo. El total de SEV utilizados fue de 40 realizados en el año 1975 y 15 en el 2006.

Se describen dos acuíferos; uno sobre depósitos piroclásticos del Cuaternario el cual es somero y libre y se encuentra desarrollado en los miembros s4 y S3'a de la formación San Salvador y en el miembro c1 de la formación Cuscatlán. El otro acuífero descrito se encuentra sobre lavas fracturadas del Cuaternario en el miembro s2 de la formación San Salvador.

Los perfiles de correlación geo-eléctrica, estratigráfica e hidrogeológica desarrollados son los siguientes:

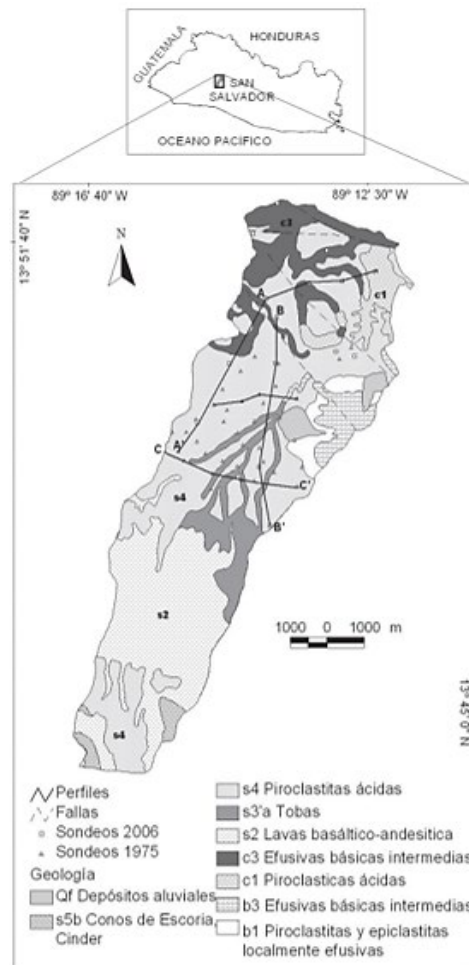


Figura 25. Distribución de perfiles en el área de Nejapa. Fuente: MARN, 2020

- El perfil A-A' fue realizado con rumbo SW en el sector norte de las lomas Las Mesas. Se determinó que en promedio el miembro s3'a se encuentra a 25 m.
- El perfil B-B' fue realizado con rumbo S, de igual manera se determinó que el miembro s3'a se encontraba a una profundidad promedio de 25 m.
- El perfil C-C' presenta un rumbo hacia el E-SE, en superficie aflora la "Tierra Blanca" entre profundidades de 3 y 6 m, y el miembro s3'a se encontraba a una profundidad promedio de 25 m.

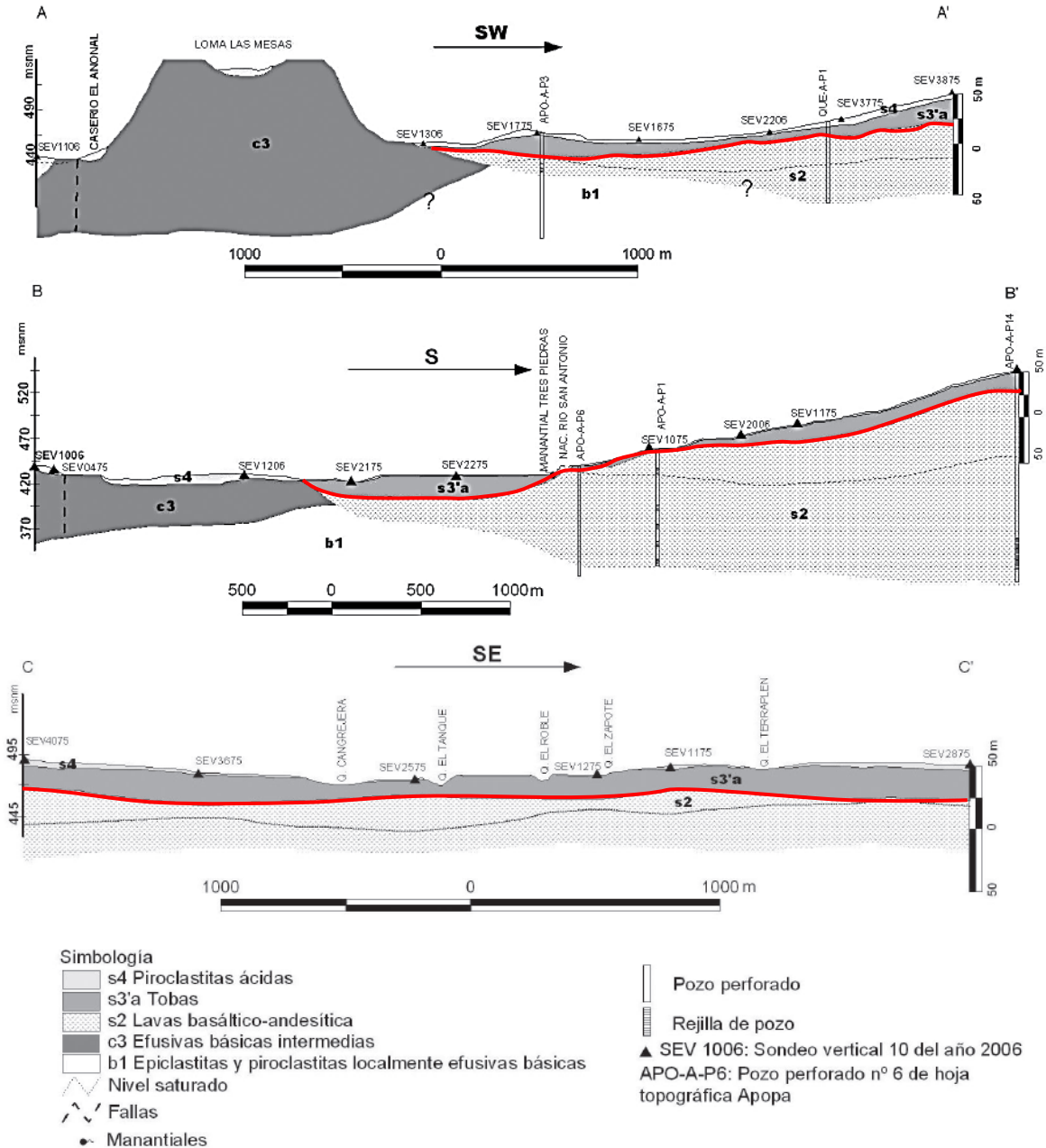


Figura 26. 3 perfiles realizados en el área de Nejapa. Línea roja: frontera de S3'a. Fuente: MARN, 2020

Para tratar de correlacionar las profundidades extraídas de los perfiles del trabajo de Gil y Salguero, 2007; con los pozos de la plataforma SIHI, se georreferenció la figura 25 donde se muestran los perfiles realizados y se montó los puntos de los pozos de SIHI. En la figura 27 se observa este procedimiento. Un total de 15 pozos fueron caracterizados usando esta información.

Figura 27. Mapa georreferenciado del estudio “Hidrogeofísica de la cuenca del río San Antonio, El Salvador: Aplicación al modelo conceptual y vulnerabilidad intrínseca”. Fuente: MARN, 2020

- **Modelo Conceptual y Vulnerabilidad Intrínseca del Acuífero Quezaltepeque-Nejapa.** *Realizado por Cesar Alvarado Batres, en la Universidad de El Salvador en 2007. (Alvarado Batres, 2007)*

La investigación se realizó en el área del acuífero de Quezaltepeque-Nejapa, ubicado entre los departamentos de La Libertad y San Salvador, el cual pertenece a la cuenca del río San Antonio. El propósito del estudio era poder determinar un modelo hidrológico conceptual del acuífero y su vulnerabilidad intrínseca debido a la importancia que presenta el recurso hídrico para el AMSS y los municipios de Nejapa y Quezaltepeque.

Se realizaron sondeos eléctricos verticales con el arreglo Schlumberger, cuya correlación con la litología de los pozos en la zona de estudio y la geología superficial, sirvieron de base para la aplicación de la metodología GOD (por sus siglas en inglés *Groundwater Hydraulic confinement, Overlaying strata, Depth to groundwater*) con la cual se podría determinar el grado de vulnerabilidad del acuífero. También se realizaron pruebas de infiltración y se correlacionaron con los registros de evapotranspiración y precipitación, sirviendo de base para la determinación del modelo conceptual.

Con la aplicación de todas las herramientas se determinó que el acuífero de Quezaltepeque-Nejapa se puede clasificar tres unidades hidrogeológicas:

- **Unidad hidrogeológica de depósitos piroclásticos.** Los que se depositan sobre la unidad b1 de la formación Bálsamo
- **Unidad hidrogeológica de lavas fracturadas.** Las cuales pertenecen a las formaciones Cuscatlán y Bálsamo
- **Unidad hidrogeológica de tobas endurecidas.**

Como se menciona, parte de la metodología utilizada para el desarrollo de la investigación, fue realizar diferentes perfiles geológicos del área, esto se efectuó con la ayuda de 75 pozos de los cuales 55 fueron excavados y 20 fueron perforados; esta información fue de utilidad para poder realizar la estimación de la profundidad de las Tobas Superiores y TBJ en el área que se muestra en la figura 28.

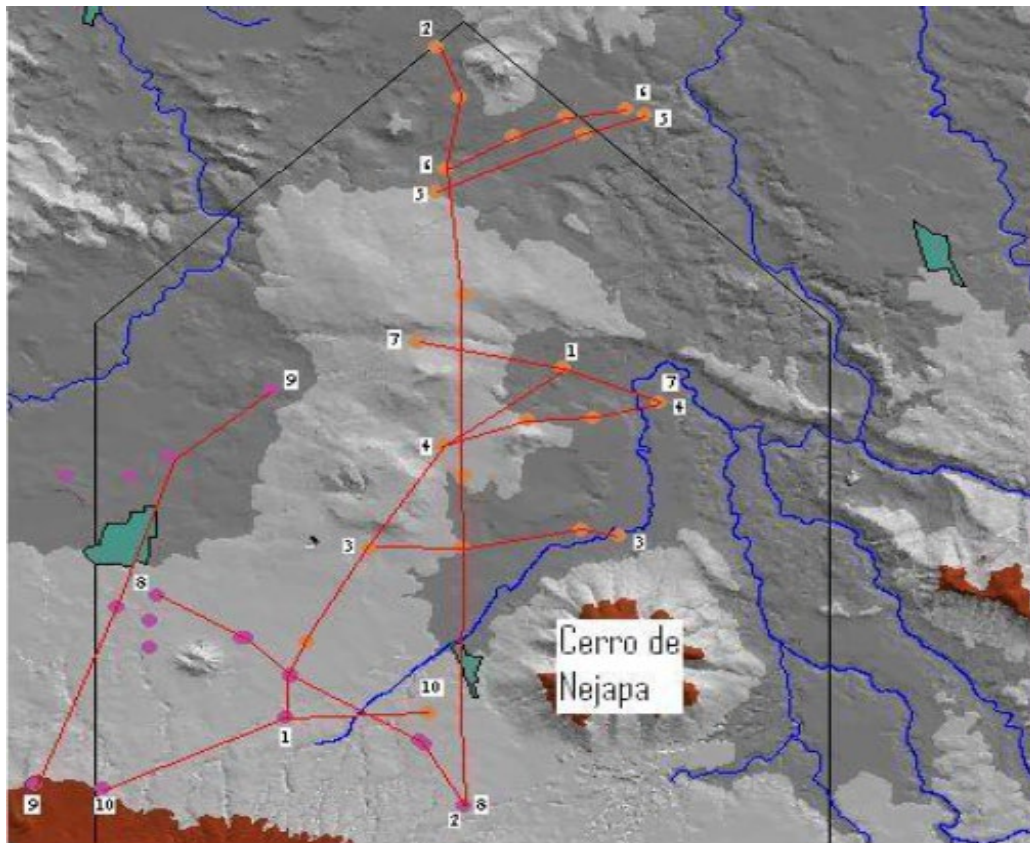


Figura 28. Distribución de perfiles estudiados en área de Quezaltepeque-Nejapa. Fuente: Batres, A. 2007

Para poder definir la profundidad de las Tobas Superiores se tomó la unidad S3'a como la base y se tomó el miembro S4 para definir la profundidad de la TBJ. De los 10 perfiles realizados en el estudio, los que se utilizaron para estimar la profundidad de las Tobas Superiores fueron 7. A continuación, en las figuras 29 hasta 35 se muestran dichos perfiles.

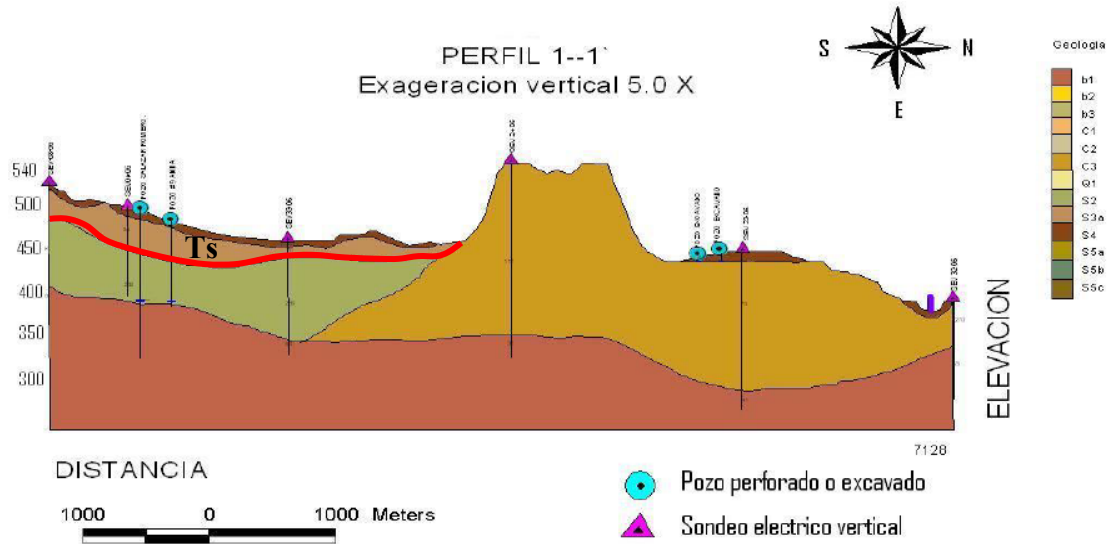


Figura 29. Perfil 1-1' en el área de Quezaltepeque-Nejapa. Fuente: Batres, A. 2007

En el perfil 1-1' la profundidad que se estima para las Ts está entre los 30-40m al sur del perfil y se observa que entre 5-10 m de profundidad aproximadamente al sur y norte se tiene TBJ, según el rumbo SN del perfil.

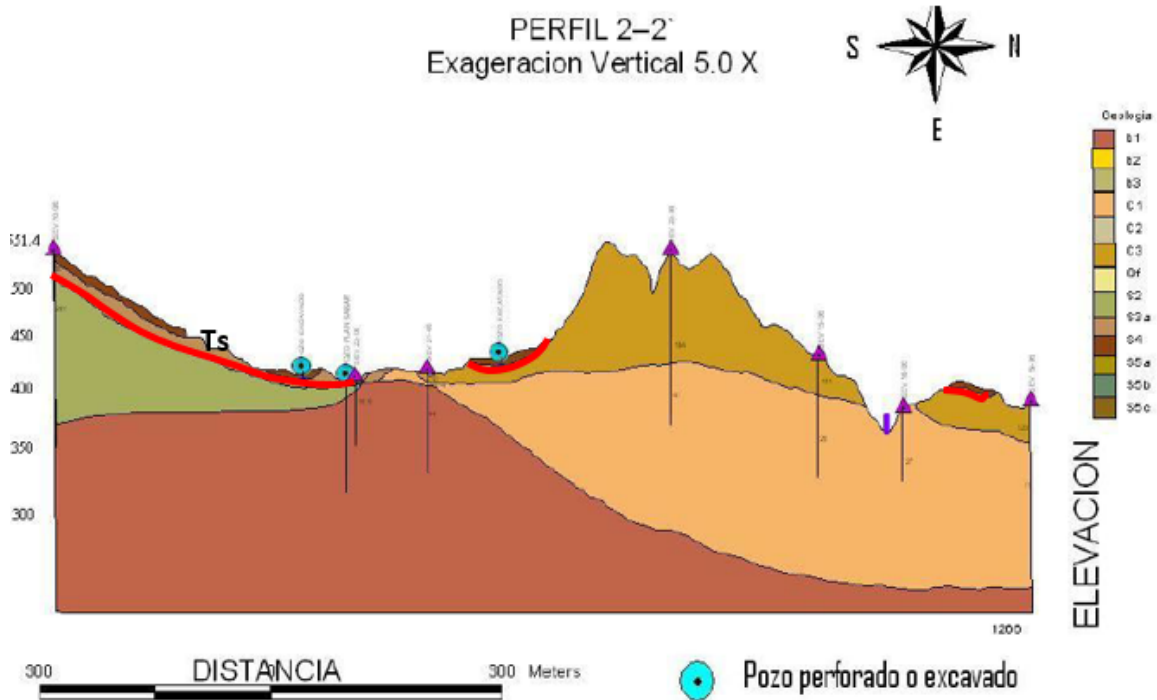


Figura 30. Perfil 2-2' en el área de Quezaltepeque-Nejapa. Fuente: Batres, A. 2007

En el perfil 2-2' la profundidad estimada para las Ts está entre los 20-30m al sur del perfil, mientras que se observa una profundidad de 5-10m aproximadamente tanto al sur como en el centro y norte de TBJ, siguiendo el rumbo SN del perfil.

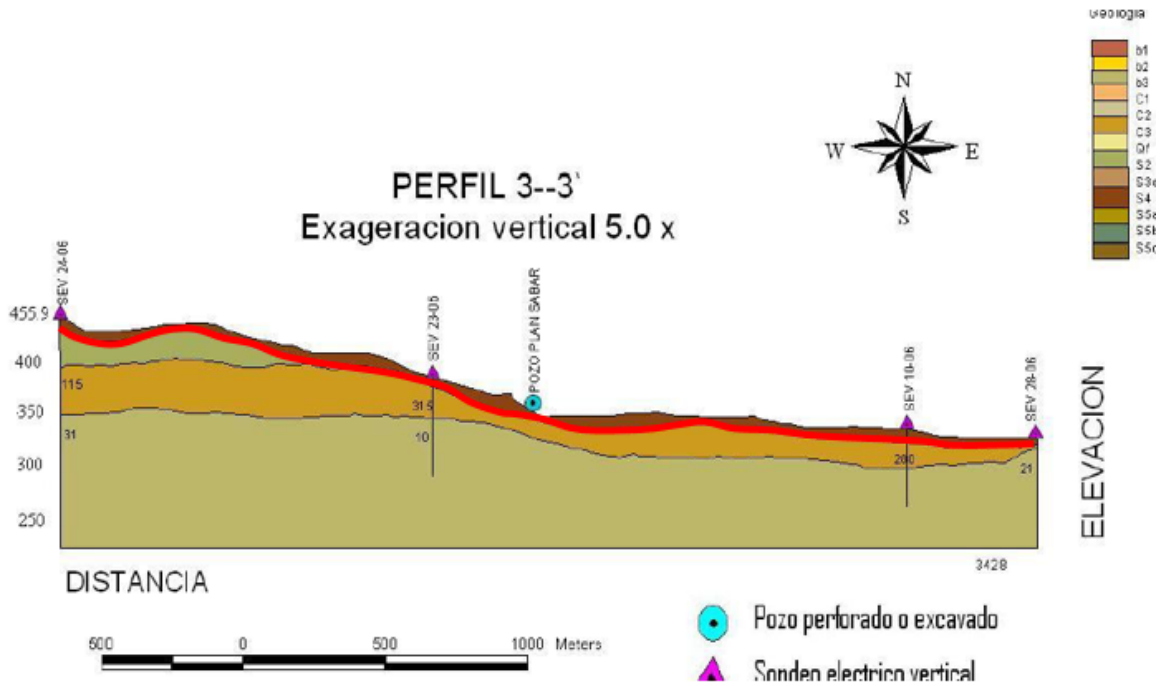


Figura 31. Perfil 3-3' en el área de Quezaltepeque-Nejapa. Fuente: Batres, A. 2007

En el perfil 3-3' de rumbo WE, la TBJ se encontraría entre los 15-20m de profundidad aproximadamente a lo largo del perfil.

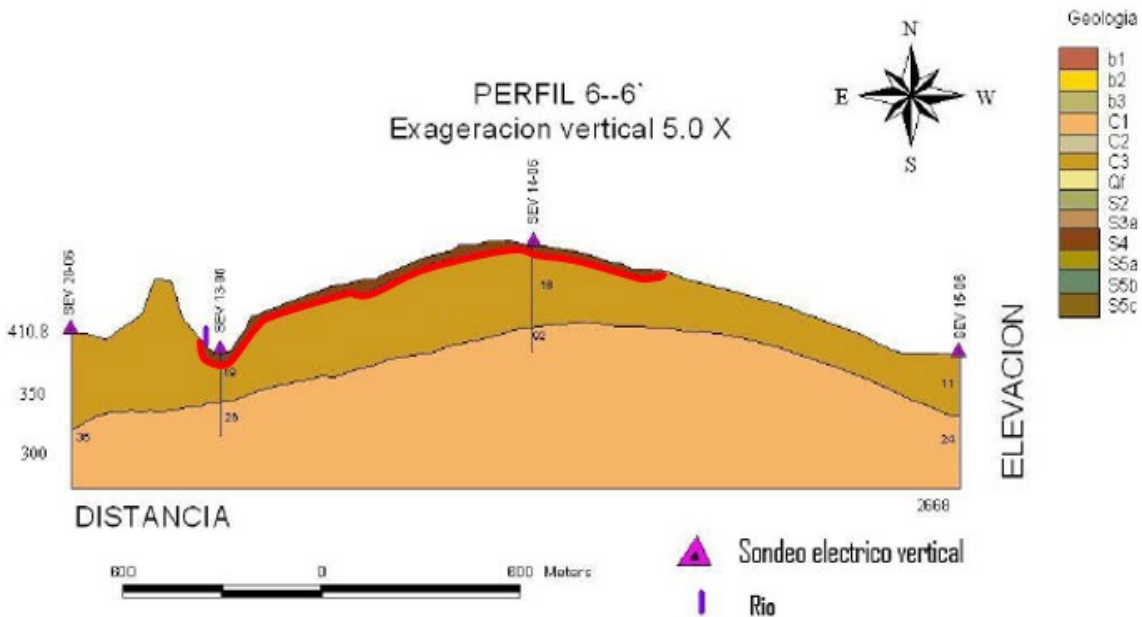


Figura 32. Perfil 6-6' en el área de Quezaltepeque-Nejapa. Fuente: Batres, A. 2007

En el perfil 6-6' de rumbo EW la TBJ se encontraría hasta aproximadamente 10m de profundidad en la parte central del perfil.

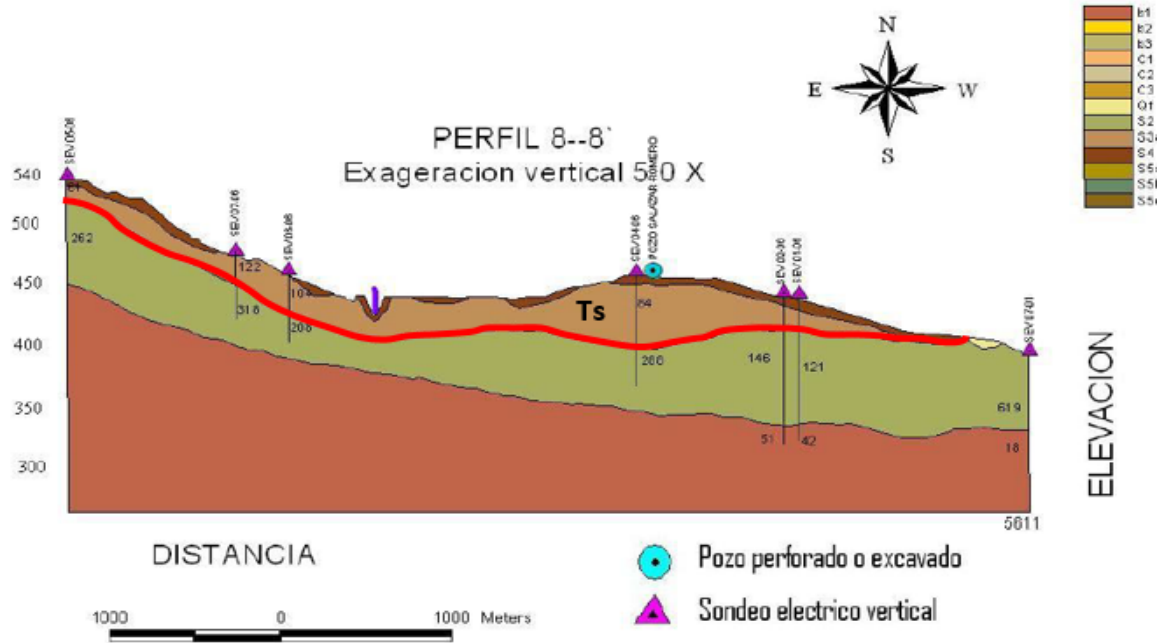


Figura 33. Perfil 8-8' en el área de Quezaltepeque-Nejapa. Fuente: Batres, A. 2007

En el perfil 8-8' la Ts se encontraría a 20m de profundidad en la parte Este del perfil, a 30m en la parte media y hasta 50m en la parte oeste, según el rumbo EW del perfil y la TBJ se presenta en los primeros 10m aproximadamente.

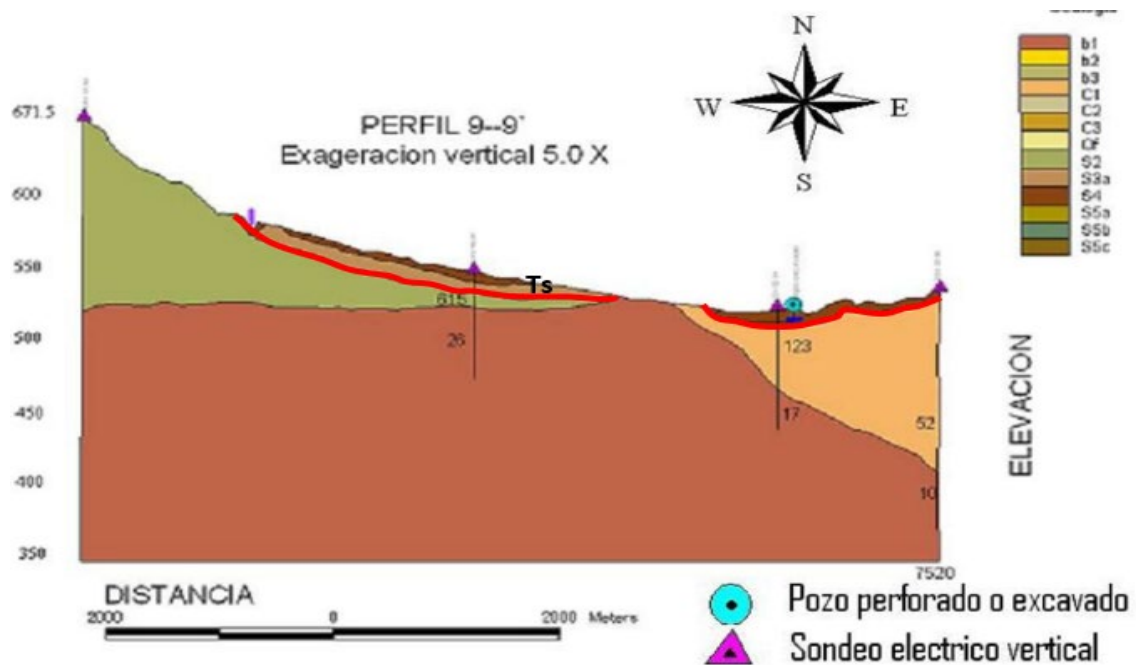


Figura 34. Perfil 9-9' en el área de Quezaltepeque-Nejapa. Fuente: Batres, A. 2007

En el perfil 9-9' con rumbo WE, las Ts se estiman que se encuentran entre los 20-25m de profundidad en la parte central del perfil, mientras que la TBJ al este del perfil se encuentran hasta 12m de profundidad y al centro hasta los 10m.

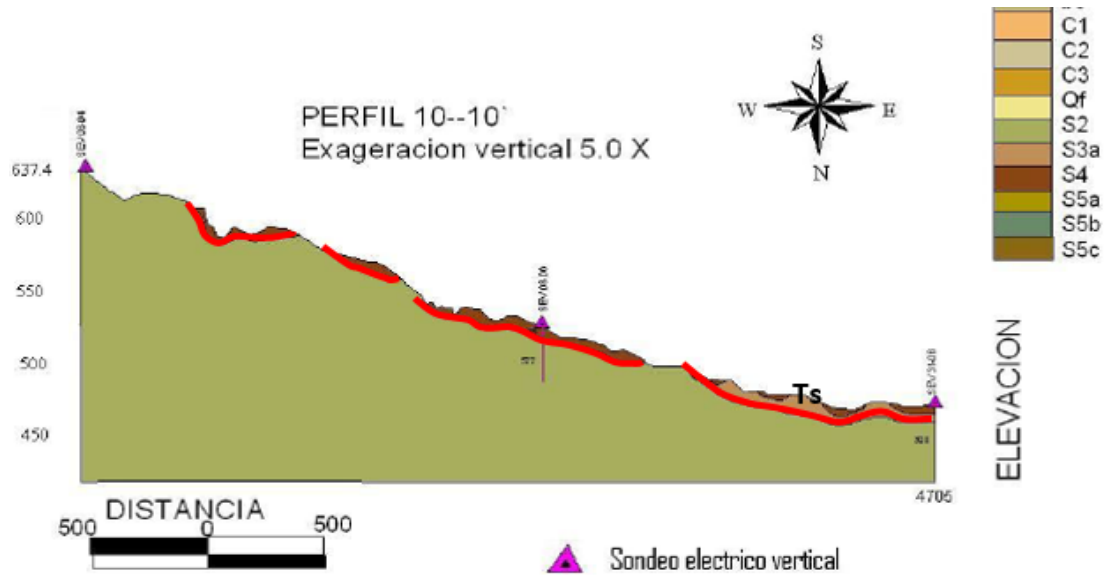


Figura 35. Perfil 10-10' en el área de Quezaltepeque-Nejapa. Fuente: Batres, A. 2007

En el perfil 10-10' con rumbo WE, las Ts se estima que se encuentran a una profundidad de 10-20m al este del perfil, mientras que la TBJ se observa hasta 10m aproximadamente. De igual manera al estudio anterior, se georreferenciaron los perfiles junto a los puntos de SIHI y así lograr realizar correlaciones para ayudar a la reclasificación de unidades geológicas.

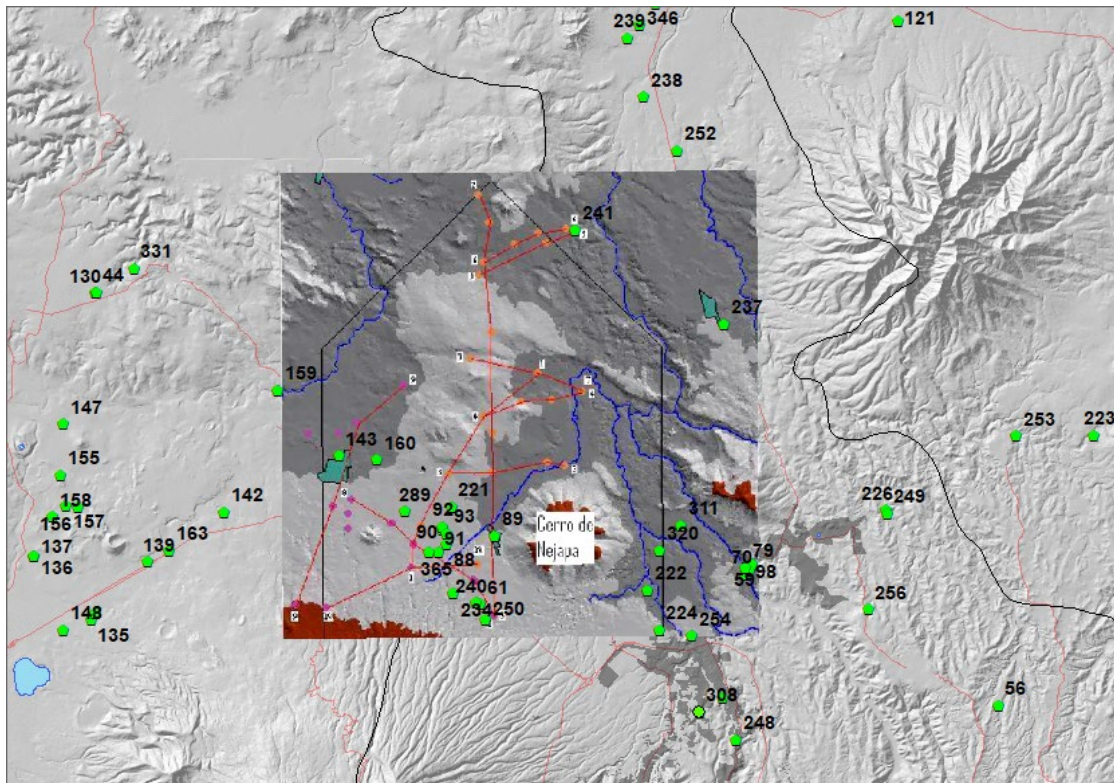


Figura 36. Mapa georreferenciado del estudio "Modelo conceptual y vulnerabilidad intrínseca del acuífero Quezaltepeque-Nejapa". Fuente: MARN, 2020

5.3.2.4. Estudios para la estimación de la profundidad de la Ts en la zona de San Salvador

- The Ilopango Caldera Complex, El Salvador: Origin and Early Ignimbrite-forming eruptions of a Graben/Pull-Apart Caldera Structure. *Realizado por Sunye Puchol, Ivan & Aguirre-Díaz, Gerardo & Davila-Harris, Pablo & Miggins, Daniel & Pedrazzi, Dario & Ortega-Obregon, Carlos & Lacan, Pierre & Hernández, Walter & Gutierrez, Eduardo, para revista Journal of Volcanology and Geothermal, publicado el año 2018.*

En el estudio se presentan las características geológicas de las tres primeras erupciones formadoras de Ignimbrita de la caldera de Ilopango, en el cual se proporciona una interpretación del origen y las etapas iniciales de la evolución volcánica.

Estas erupciones se asocian a los Tobas Inferiores. El trabajo geológico de este estudio se basó inicialmente en el mapa geológico El Salvador 1: 500,000 de Weber (1974), así como en mapas anteriores y estratigrafía informados por otros autores (Hernández, 2008; Lexa, 2011; Garibaldi, 2016).

En el estudio se menciona que el último evento volcánico de la caldera de Ilopango ha sido la erupción de las Islas Quemadas en 1879-1880, el cual formó un domo de lava dentro del lago. Posteriormente, se explica que la erupción de Ignimbrita más joven formada por la caldera ha sido la erupción histórica Tierra Blanca Joven (TBJ), hace aproximadamente 1500 años y que aparentemente devastó la civilización maya en la región (Dull et al., 2010).

Seguidamente, antes de la erupción de TBJ, hubo 3 grandes erupciones explosivas durante los últimos 57 mil años: TB4, TB3 y TB2; las cuales corresponden a las Tobas Superiores, mencionadas en este documento.

Finalmente, y el objetivo de los autores, es caracterizar las más antiguas y múltiples erupciones de ignimbrita de gran volumen asociadas con episodios de colapso ocurridas en la Caldera de Ilopango, las cuales no han sido descritas e interpretadas previamente. Estas son: Olocuilta Ignimbrita (OI), Colima Ignimbrita (CoI) y Apopa Ignimbrita (ApI). Los autores argumentan que toda esta actividad volcánica se relaciona con la estructura regional de fallas preexistente asociada con la zona de falla de El Salvador (ESFZ).

A través de estos datos y varias campañas de trabajo de campo entre 2015 y 2017, se construyó un nuevo marco estratigráfico y un nuevo mapa geológico para el complejo de la caldera de Ilopango y sus alrededores.

El nuevo mapa geológico, mostrado en la figura 37, cubre un área aproximada de 3000 km² de la parte central de El Salvador, desde la Cordillera Fronteriza en el norte, hasta la costa del Pacífico en el sur y desde la cuenca del río Lempa en el este hasta la caldera de Jayaque en el oeste. Los afloramientos fueron descritos geológicamente y se midieron las secciones estratigráficas.

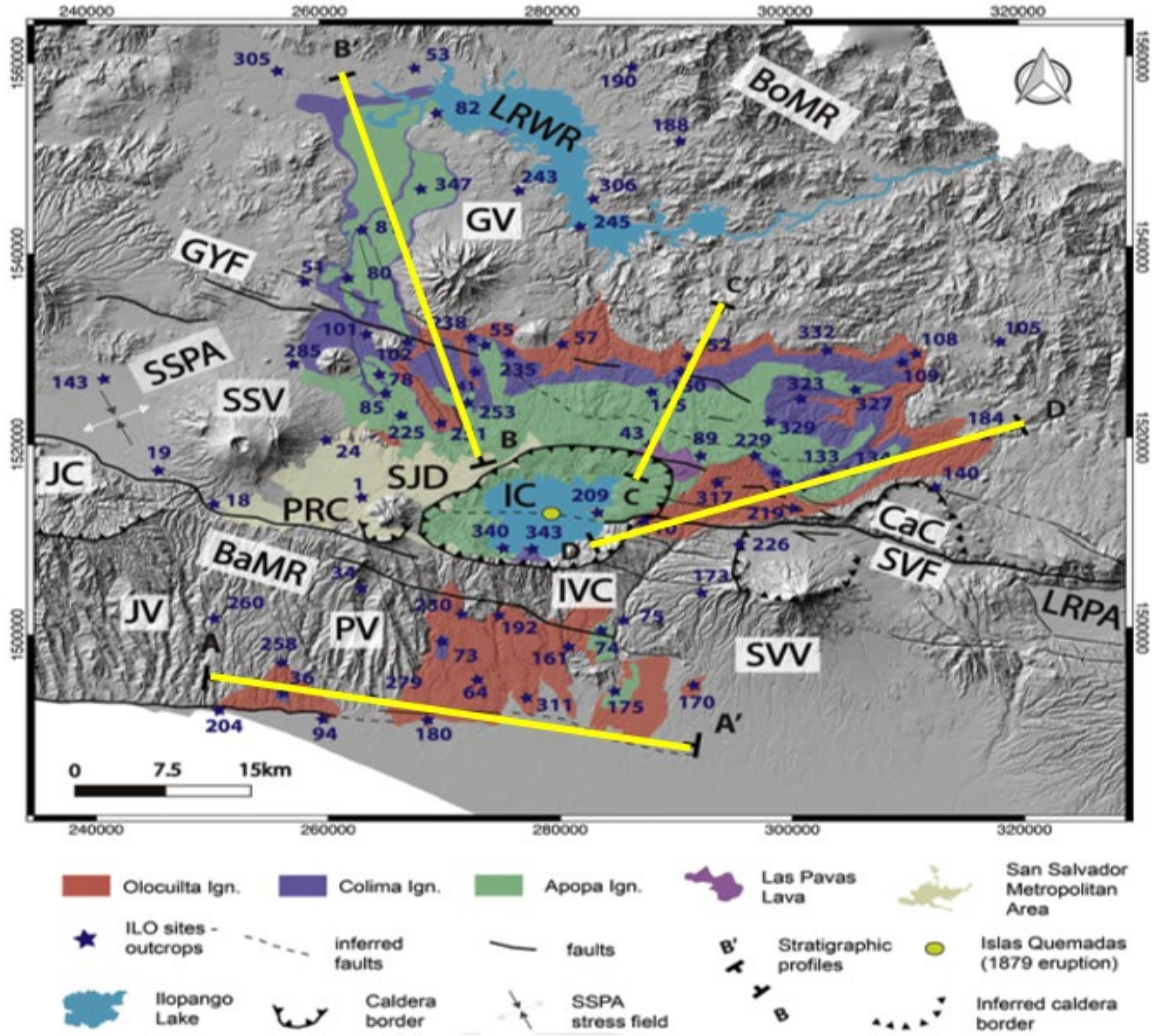


Figura 37. Distribución de las primeras tres ignimbritas de la caldera de Ilopango (IC), Olocuilta Ignimbrita (OI, rojo), Ignimbrita de Colima (CoI, azul) y Ignimbrita de Apopa (ApI, verde). Sitios de ubicación de muestreo (estrellas azules, muestra número con prefijo ILO-), Las Pavas Lava (LPL, violeta), IVC: Complejo Volcánico Ilopango, JC: Caldera Jayaque, JV: Volcán Jayaque, PV: Volcán Panchimalco, PRC: Caldera Planes de Renderos, SJD: San Jacinto Dome, SSV: Volcán San Salvador, SVV: Volcán San Vicente, CaC: Carbonera caldera, GV: Volcán Guazapa, SSPA: San Salvador Pull-Apart, BaMR: Cordillera Bálsamo, BoMR: Cordillera Fronteriza, LRPA: Río Lempa Pull-Apart, SVF: San Falla Vicente, GYF: Falla Guaycume; LRWR: Embalse del Río Lempa. Fuente: Puchol, et al., 2018

Los autores utilizaron el modelo de elevación digital (DEM) con una precisión de 10 m como mapa base. Se midieron aproximadamente 85 secciones estratigráficas y se recolectaron varias muestras para la datación $^{40}\text{Ar}/^{39}\text{Ar}$ y U-Pb para el análisis petrográfico y para el análisis de elementos principales y trazas.

Luego de las descripciones estratigráficas, el mapeo geológico y los resultados analíticos, los autores realizaron una correlación espacial y temporal para las primeras tres ignimbritas de la Caldera de Ilopango. Por lo que se realizaron cuatro perfiles (A, B, C Y D) los cuales contienen medidas de secciones estratigráficas representativas (registros) que muestran el grosor y la distribución de las primeras ignimbritas del Ilopango.

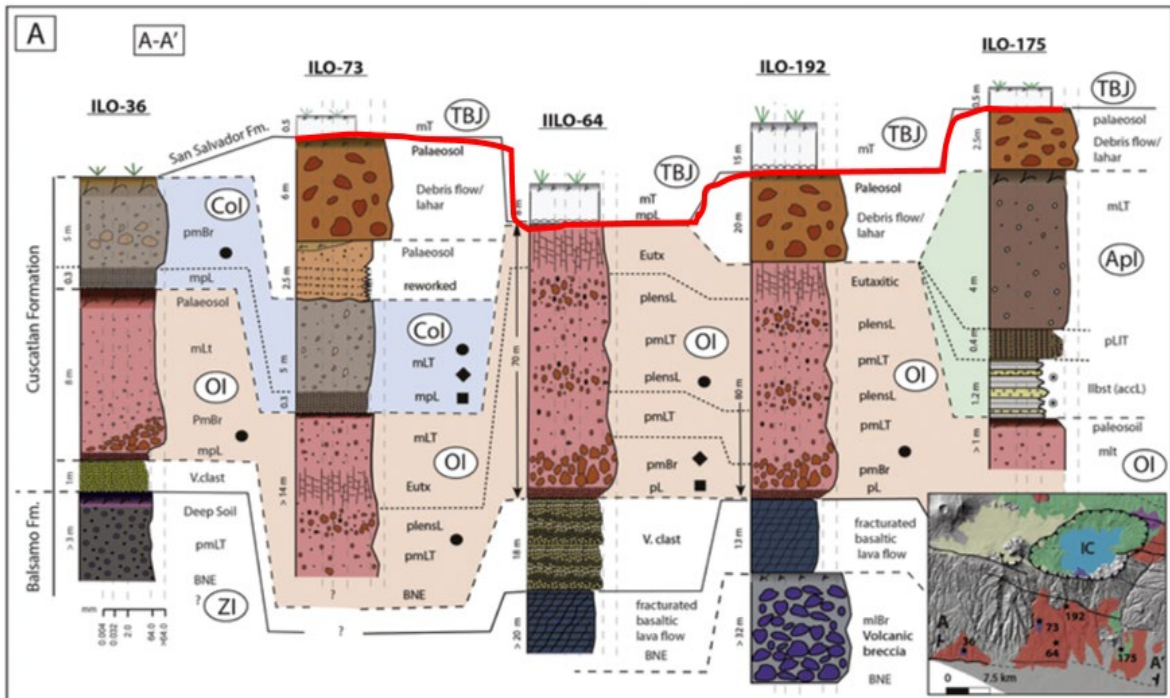


Figura 38. Perfil A-A' ubicado al sur de San Salvador y Caldera de Ilopango. Fuente: Puchol, et al., 2018

En el perfil A-A' se determinó un espesor de la TBJ de entre 0.5 m a 8 m y en ILO-192 hasta de 15 m de espesor el cuál es el más cercano a la caldera de Ilopango, luego se tiene la formación Cuscatlán.

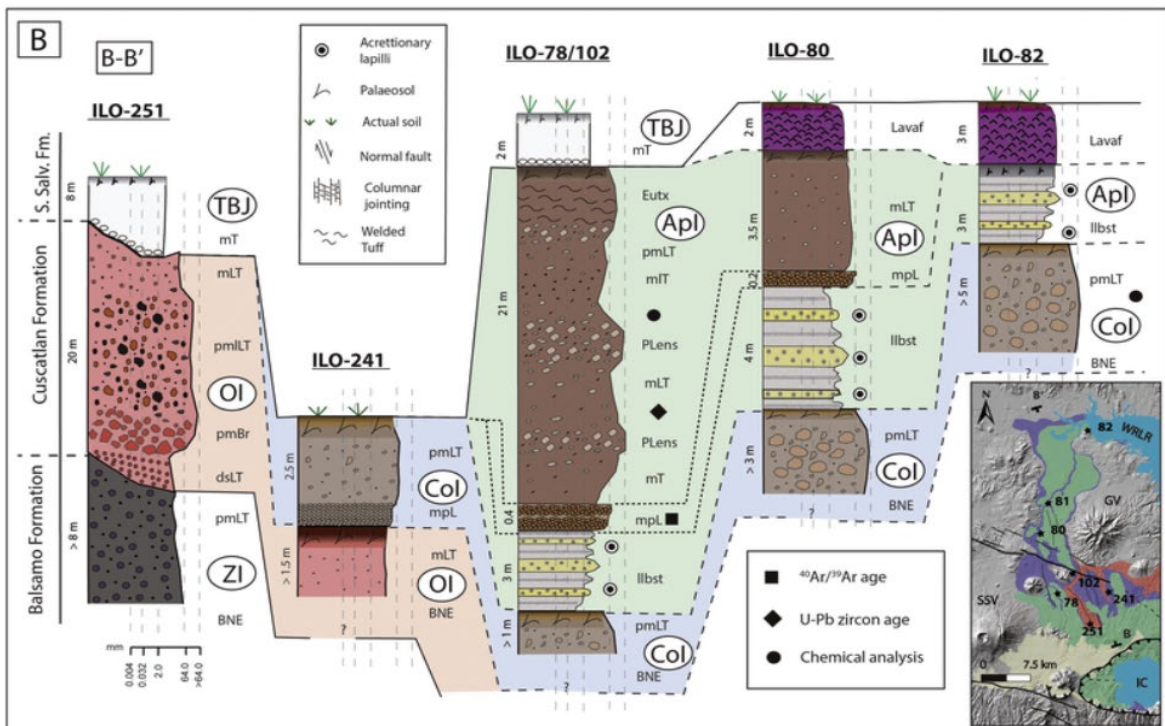


Figura 39. Perfil B-B' ubicado al norte de San Salvador, zona de Apopa. Fuente: Puchol, et al., 2018

Para el perfil B-B' de igual manera se determinaron los espesores de TBJ entre 2 y 8 m, siendo el mayor de 8 m que se encuentra más cercano a la caldera.

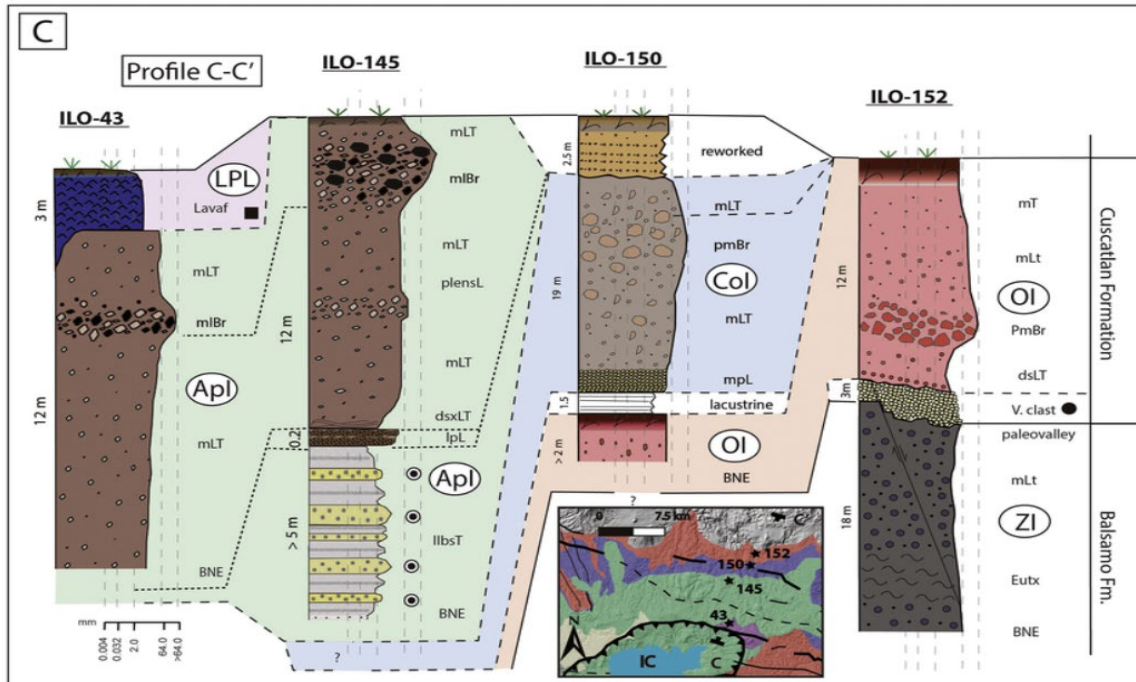


Figura 40. Perfil C-C' ubicado al noreste de San Salvador, zona de Ilobasco. Fuente: Puchol, et al., 2018

Para el perfil C-C' no se encontró TBJ, está representado mayormente por la formación Cuscatlán y sus respectivas Tobas Inferiores correspondientes a la caldera de Ilopango.

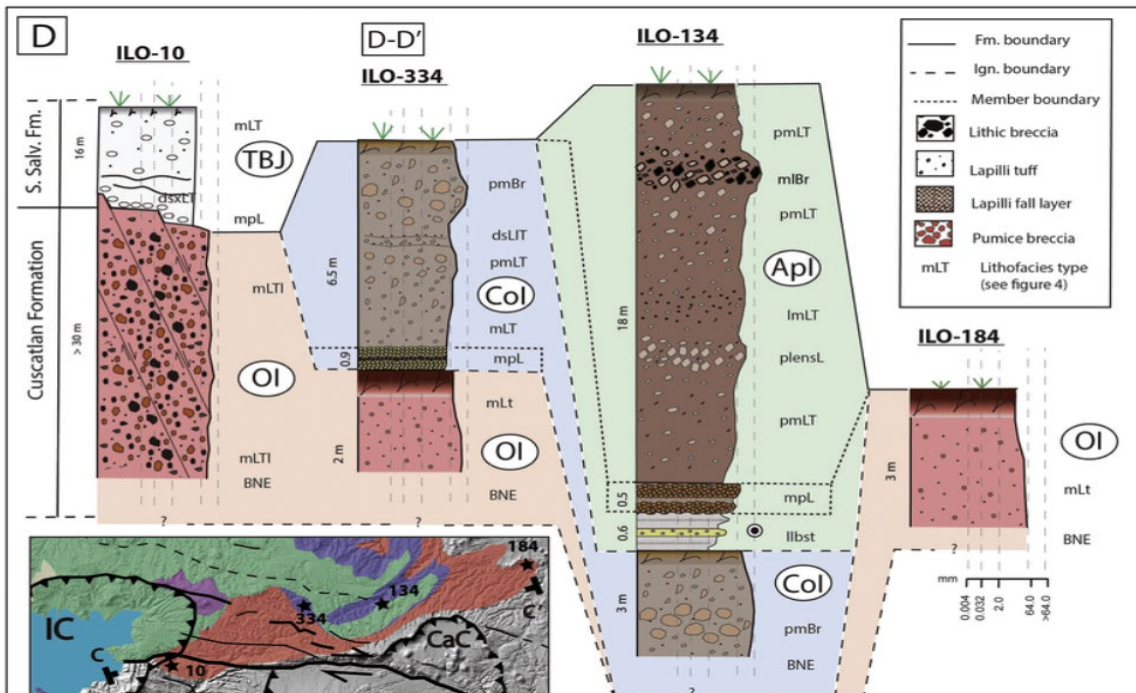


Figura 41. Perfil D-D' ubicado al este de la caldera de Ilopango. Fuente: Puchol, et al., 2018

Para el perfil D-D' se determinó un espesor para la TBJ de 16 m en el punto ILO-10 el cual está ubicado en los límites de la caldera de Ilopango.

Finalmente, se utilizaron los perfiles B-B' y D-D' por la mayor concentración de puntos SIHI que se tenía en el área. Según la descripción de estos dos perfiles no se encuentran las tobas superiores (Ts) solamente se cuentan con TBJ con un promedio de profundidad de entre 2 a 8 m en el perfil B-B' y en el perfil D-D' en la zona cercana a la caldera de Ilopango con una profundidad de 16 m, luego de la TBJ se tiene la formación Cuscatlán a las que pertenecen las tobas inferiores (Ti).

- **Caracterización Hidro-geoquímica e Isotópica de Áreas de Recarga en el Acuífero de San Salvador.** *Realizado por Marcia Lizeth Barrera de Calderón, en la unidad de posgrados de la Universidad de El Salvador en el año 2010. (Barrera Calderón, 2010)*

El área de estudio se centra en el acuífero de San Salvador (ver figura 42) delimitada por la microcuenca alta del río Acelhuate, fue estudiada por la autora de manera hidrogeológicamente y caracterizada mediante técnicas hidro-geoquímicas e isotópicas.

La unidad acuífera fue identificada en lavas basálticas fracturadas y lavas andesíticas con intercalaciones de piroclastos el cual estaba conectada con otra unidad acuífera porosa de piroclastos intercalados con sedimentos y tobas, fueron identificados y delimitados en un mapa a partir de los resultados de un modelo litológico y del análisis de informes de perforación de pozos de producción ubicados en la zona.

La información geológica utilizada por la autora se basó en una recopilación de dos fuentes principales las cuales son informes de perforación proporcionados por el Proyecto FORGAES que fue coordinado por el Ministerio de Medio Ambiente y Recursos Naturales y la Unión Europea; y la segunda fuente fueron registros litológicos propiedad de ANDA.

Obtuvo información geológica de 69 pozos, de los cuales 54 fueron los que utilizó para la realización del modelo geológico.

Los sitios de muestreo fueron seleccionados del inventario de ANDA y un inventario de 62 fuentes realizado para la investigación. El criterio que la autora tomo para seleccionar los puntos de muestreo fue la distribución espacial y en el caso de los pozos, tomó la consideración de la rejilla o el grupo de rejillas se ubicara en un estrato litológicamente similar.

A continuación, desde la figura 44 hasta la 55, se presentan los perfiles con la estimación de la profundidad de las Tobas superiores (Ts) el cual se corroboró con la ayuda de los perfiles que se realizaron utilizando el DEM de 10m. Para poder establecer que se trata de las Ts se guio a partir de la descripción litológica que se hace en la descripción de cada perfil y se comparó con los perfiles que se realizaron.

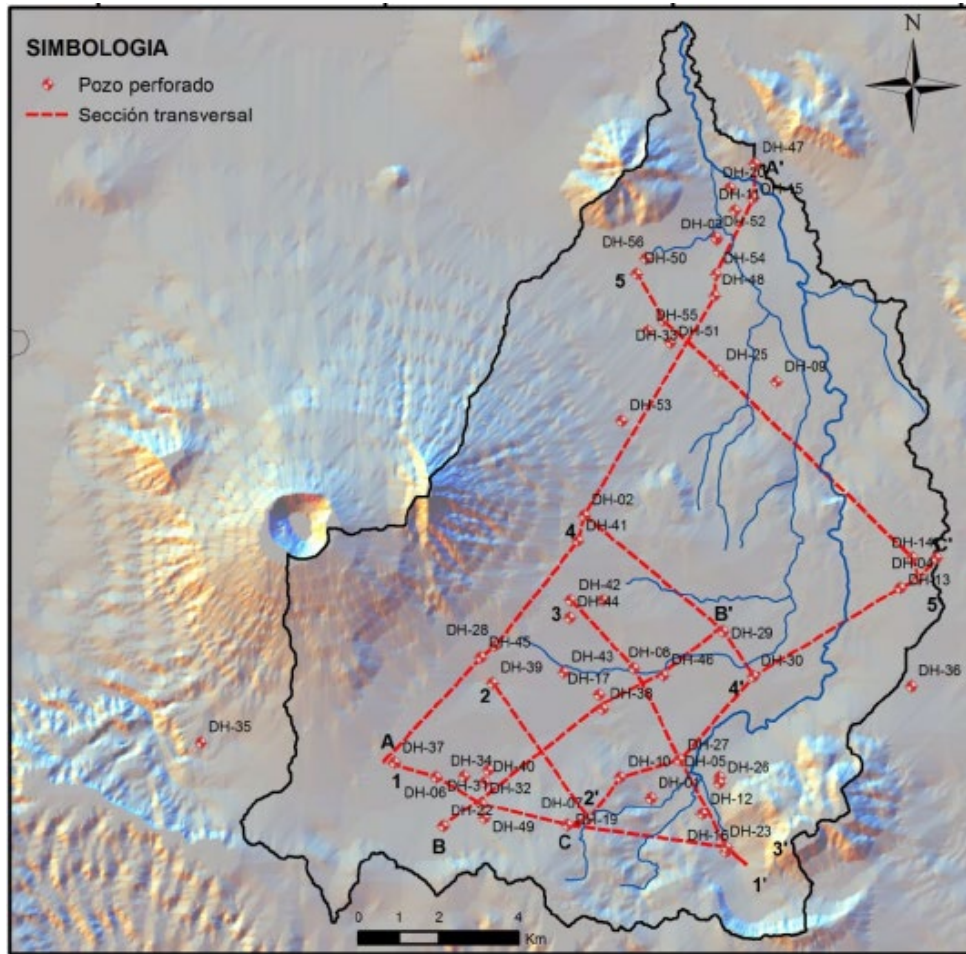


Figura 43. Distribución de perfiles en el área de San Salvador. Fuente: Barrera Calderón, 2010

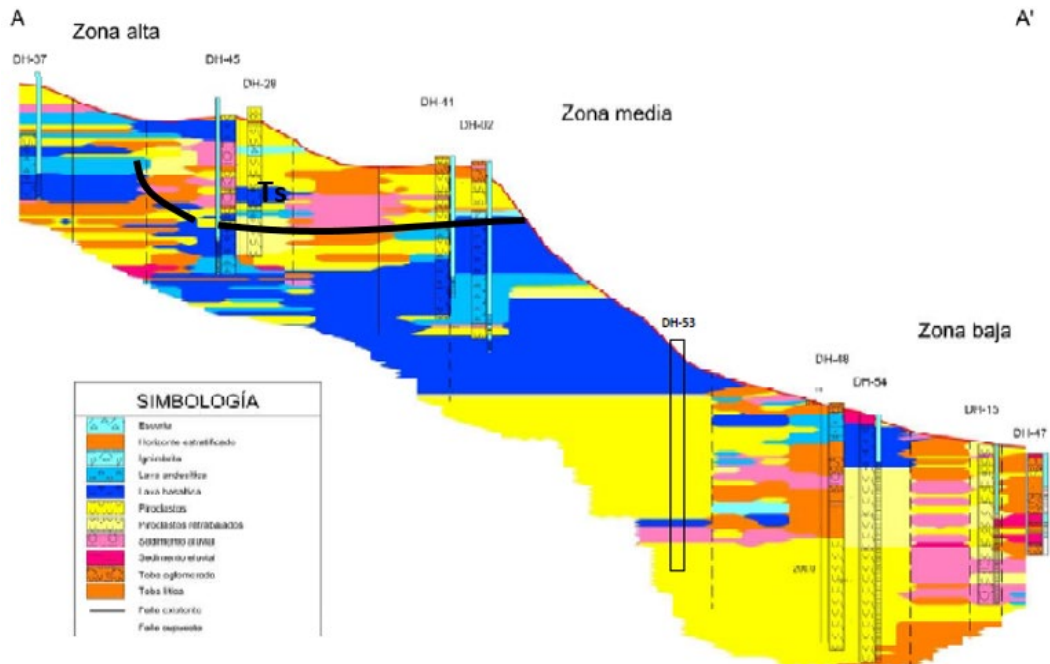


Figura 44. Perfil A-A' Urbanización Jardines del Volcán hasta Apopa. Fuente: Barrera Calderón, 2010

En perfil A-A' que se extiende desde la Urbanización Jardines del Volcán hasta Apopa, las Ts en la zona alta se encontrarían a una profundidad aproximada de 90m debido a que se encuentran grandes espesores de la avalancha de escombros del Volcán San Salvador.

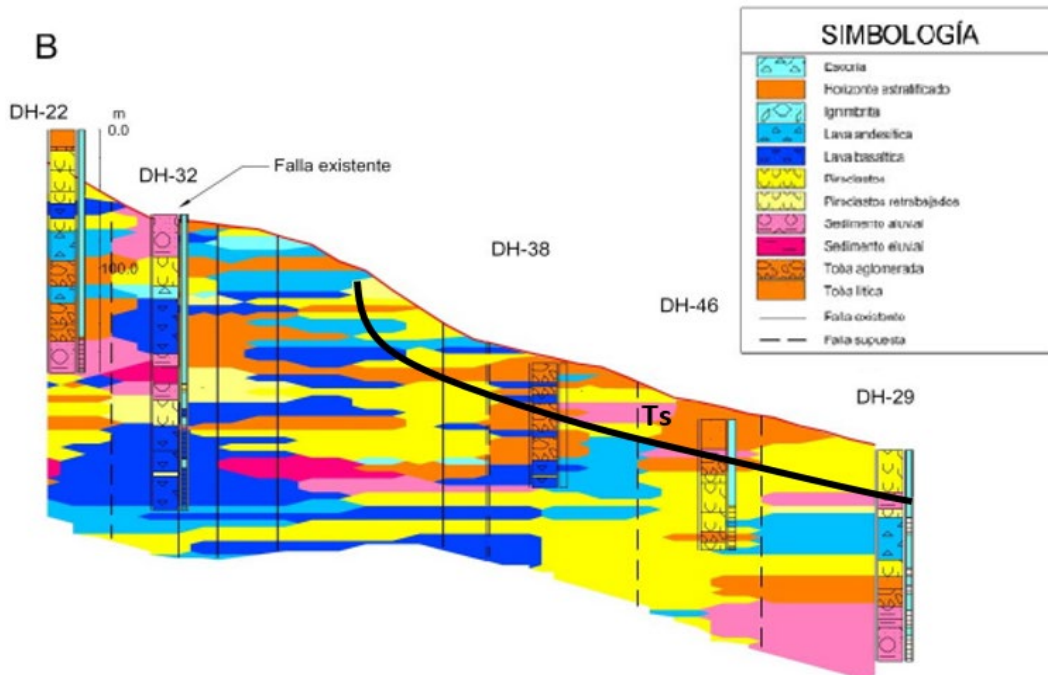


Figura 45. Perfil B-B' Antiquo Cuscatlán hasta la colonia Rábida. Fuente: Barrera Calderón, 2010

En el perfil B-B' que se extiende desde Antiquo Cuscatlán hasta la colonia Rábida, la Ts se localiza hasta 30m de profundidad aproximadamente.

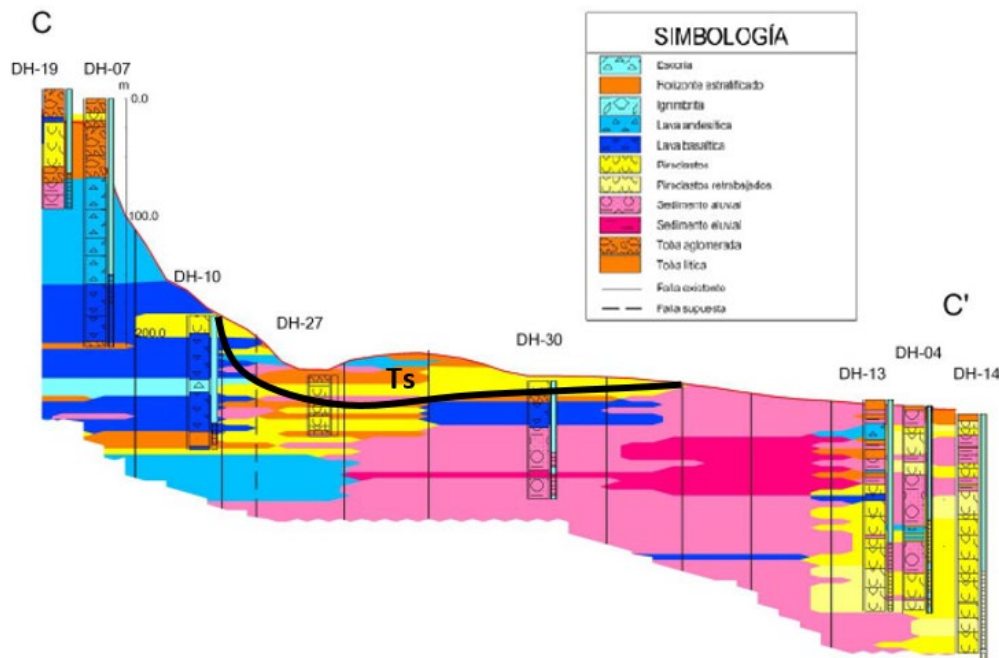


Figura 46. Perfil B-B' Cementerio La Resurrección hasta Plan del Pino. Fuente: Barrera Calderón, 2010

El perfil C-C’ el cual se extiende desde el Cementerio La Resurrección hasta Plan del Pino, la Ts se estima que se encuentra entre los 20-10m de profundidad aproximadamente en la zona de la Industria La Constancia.

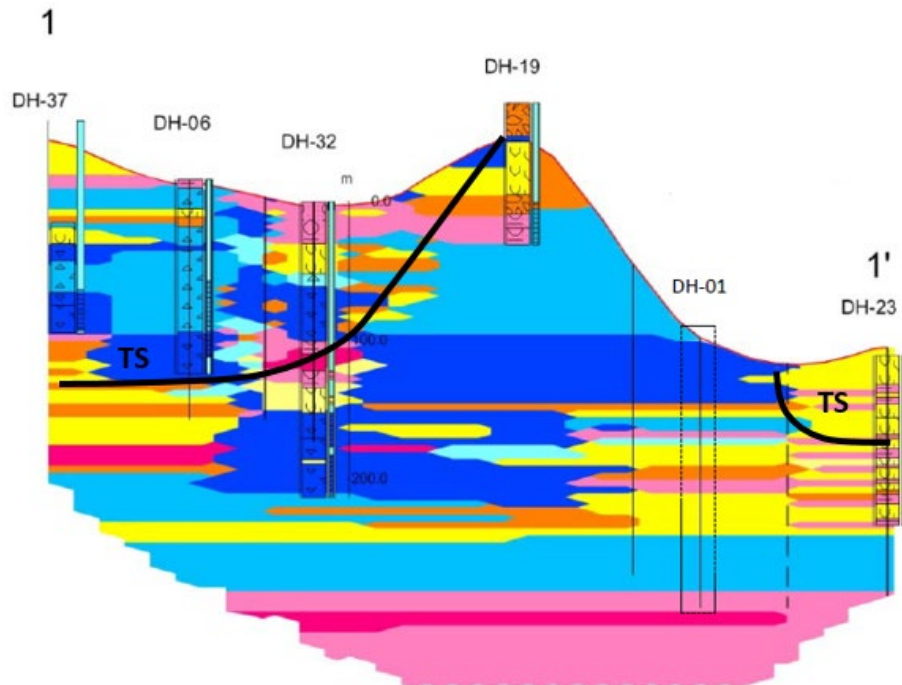


Figura 47. Perfil 1-1' Urb. Jardines del Volcán hasta San Marcos. Fuente: Barrera Calderón, 2010

El perfil 1-1' se extiende desde la Urbanización Jardines del Volcán hasta San Marcos, en la zona de San Marcos la Ts llegaría aproximadamente hasta los 50 m. En el plan de la Laguna la Ts tendría 100m de profundidad.

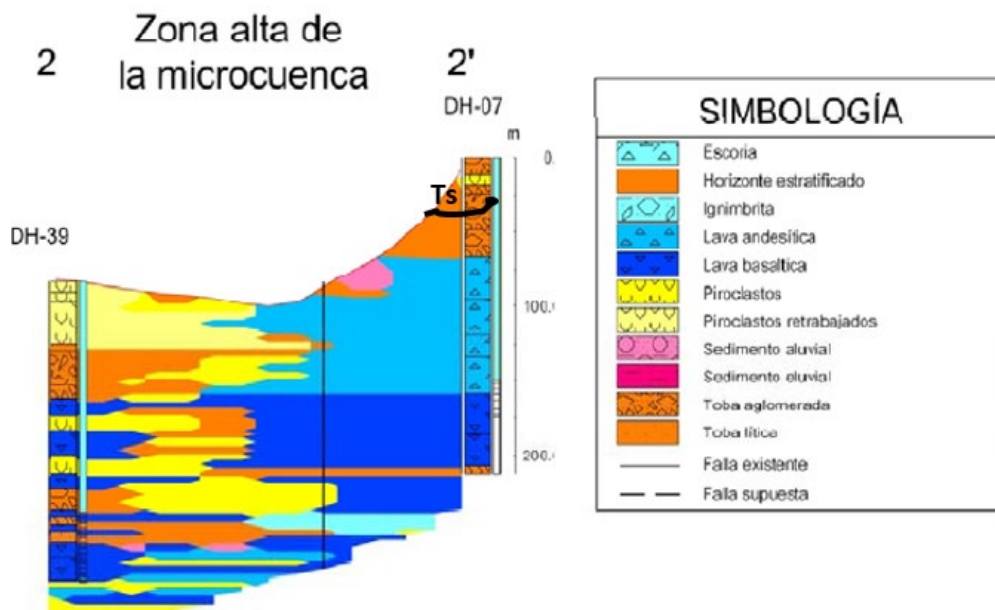


Figura 48. Perfil 2-2' Club Campeste hacia el Cementerio Resurrección. Fuente: Barrera Calderón, 2010

El perfil 2-2' se extiende desde el Club Campestre hacia el Cementerio La Resurrección, en la zona cerca de la colonia San Benito (a la izquierda en el perfil) la Ts se encuentra aproximadamente a una profundidad de 25m y 20 m (a la derecha).

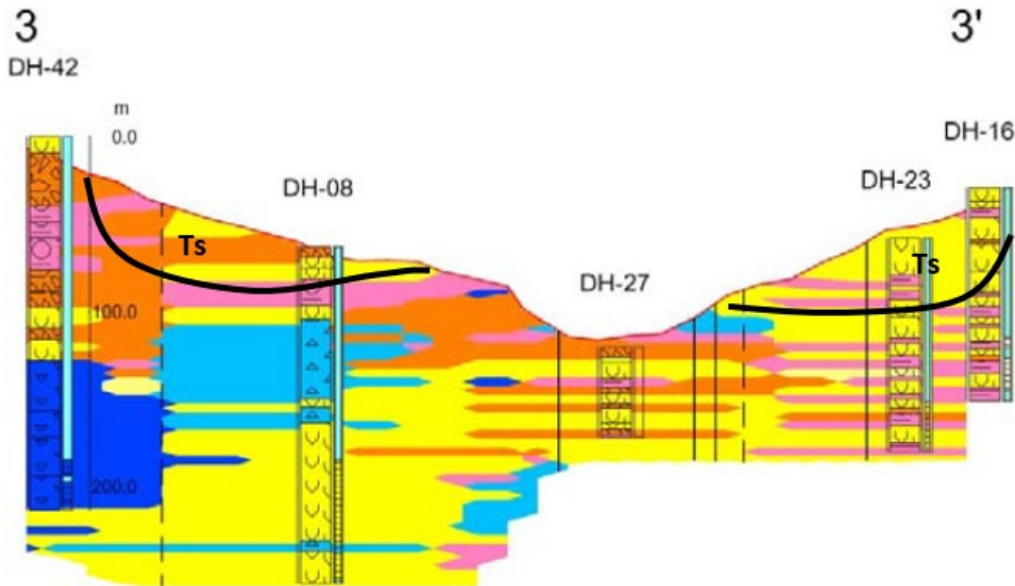


Figura 49. Perfil 3-3' Monumento a la Constitución hacia San Marcos. Fuente: Barrera Calderón, 2010

El perfil 3-3' se extiende desde el Monumento a la Constitución hacia San Marcos, la Ts en la zona cerca de la Miralvalle se podría encontrar a aproximadamente 25m de profundidad (a la izquierda en el perfil), mientras que, en San Marcos se podría encontrar aprox. a 50m (a la derecha).

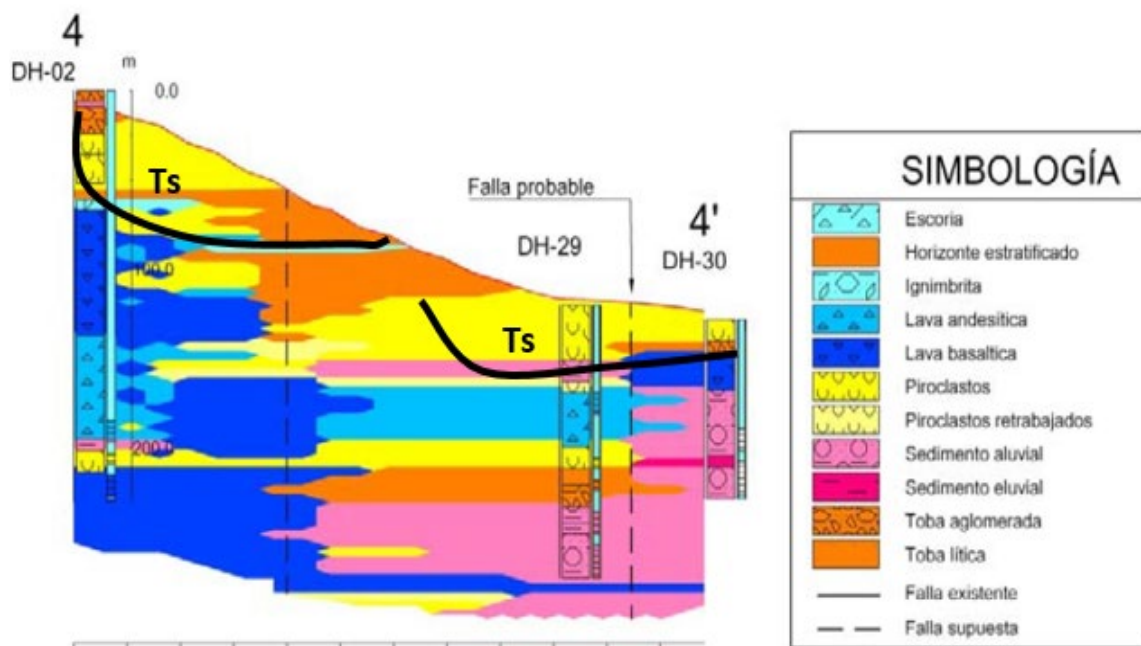


Figura 50. Perfil 4-4' Mejicanos hasta Industrias La Constancia. Fuente: Barrera Calderón, 2010

El perfil 4-4' se extiende desde Mejicanos hasta Industrias La Constancia, cerca de la colonia La Gloria la Ts se localizaría aprox. A 10m de profundidad (izquierda) y en la parte oriental en la zona de La Constancia a aprox. 30m de profundidad.

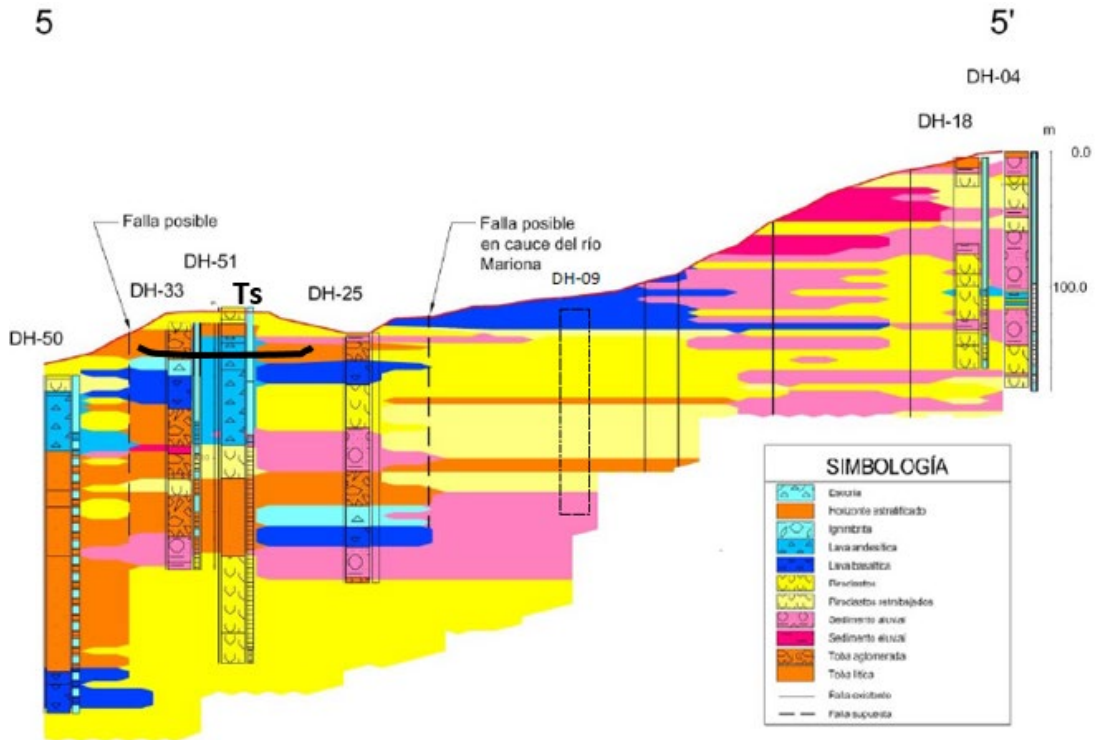


Figura 51. Perfil 5-5' Metrópoli San Gabriel hasta Plan del pino. Fuente: Barrera Calderón, 2010

El perfil 5-5' se extiende desde Metrópoli San Gabriel hasta Plan del pino, en él la Ts se encontraría a aproximadamente de 5-10m de profundidad al noreste de Apopa (a la izquierda en el perfil).

Seguidamente, el mapa de la zona de estudio con la distribución de los perfiles se georreferenció para poder tener una vista de los puntos de los pozos de SIHI, los cuales fueron revisados y caracterizados con la ayuda de la estimación de la Toba superior (Ts).

5.3.2.5. Mapa de isolíneas de profundidades de la Ts en la zona del AMSS

Una vez analizada la información de los perfiles anteriores, se procedió a hacer un mapa de isolíneas que representara de manera simplificada las profundidades en la que se estima la base de las tobas superiores (Ts) en la zona de San Salvador. Este procedimiento se asemeja al efectuado con el mapa de isolíneas del valle de Zapotitán en la figura 11.

Entre algunos aspectos importantes a considerar, para ciertas ubicaciones específicas de San Salvador y Santa Tecla, dada información adicional, se encuentran los siguientes:

- En la parte oriental del cerro San Jacinto, se estima que TBJ tendría 25 m de espesor.
- En la Col. Escalón, el estudio del colapso del Antiguo Volcán de San Salvador por Walter Hernández (Hernández & Jicha, 2019) muestra que las Tobas Superiores estarían a una profundidad de poco más de 20 m.
- La perforación realizada en el parque Bicentenario, Antiguo Cuscatlán y Santa Tecla, arrojó que la base de las tobas superiores estaría a 88 m de profundidad, ya que luego inicia la Avalancha de Escombros del colapso del Antiguo Volcán de San Salvador.
- En Metrocentro, una quebrada cercana demostró que tenía 20 m de caída y no se alcanzaba a las Tobas Inferiores por lo cual se infiere que las Tobas Superiores podrían llegar como máximo hasta 30 m.
- En las inmediaciones de Industrias La Constancia, las profundidades de la base de las Tobas Superiores se estiman entre 50 a 55 m.

A continuación, en la figura 52 se muestra un mapa de isolíneas de profundidades estimadas de las Tobas Superiores, utilizando toda la información mostrada anteriormente. Con este mapa, se superpuso las ubicaciones de los pozos de la plataforma SIHI (en esta área) y se fue revisando uno por uno, realizando la reclasificación las diferentes unidades.

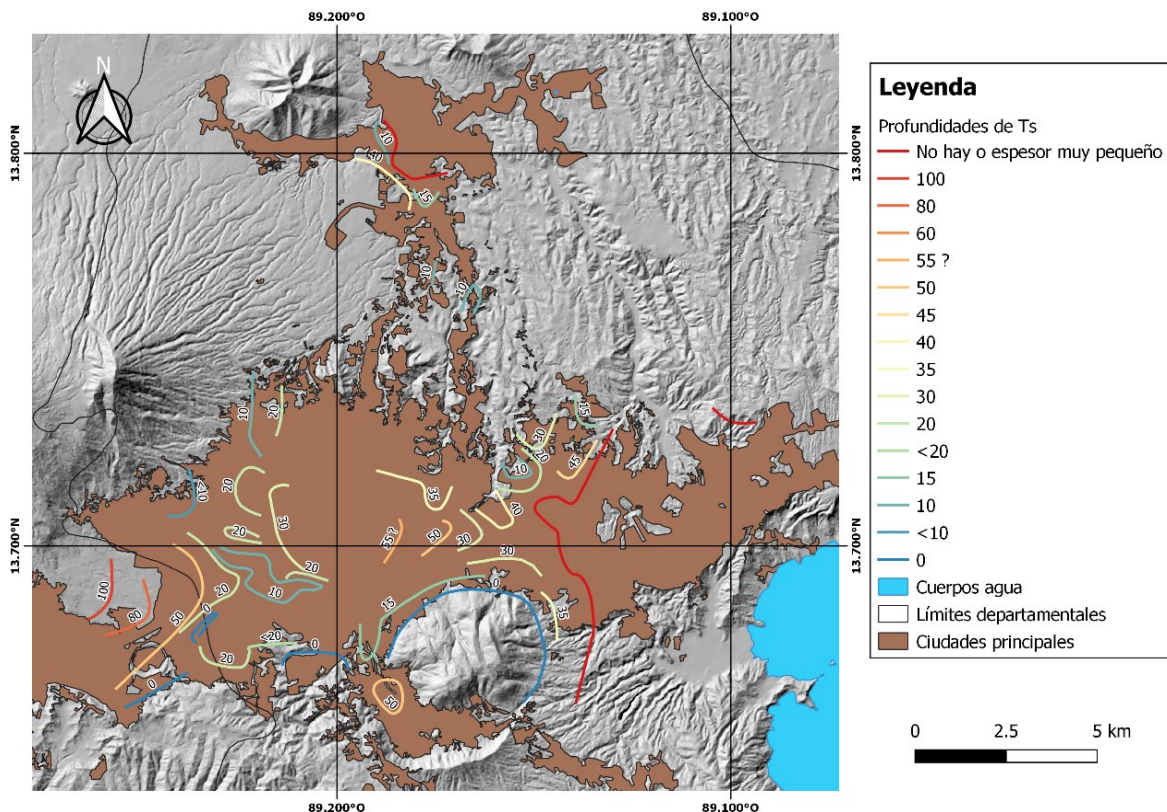


Figura 52. Profundidades de las tobas superiores en el Área Metropolitana de San Salvador (AMSS)
Fuente: MARN, 2020

6. Consolidación de base de datos

Una vez se fueron completando las reclasificaciones de cada uno de los perfiles geológicos de los pozos recabados, se fue asignando las velocidades de corte a cada una de las unidades geológicas identificadas. Con esto para cada punto se fue construyendo un perfil de velocidad unidimensional, con el fin de observar más claramente los contrastes de velocidades y en la determinación del valor de Vs30.

Como se mencionó anteriormente, esta práctica no es necesariamente una medida directa de la velocidad de corte, sin embargo, se trató de aprovechar en lo mayor posible la información geológica, la cual era abundante y abarcaba una buena distribución espacial; la cual, junto a un trabajo arduo con opiniones de expertos en la geología del país y con la comparación de los resultados de los ensayos Down-hole y su respectiva caracterización geológica (en lugares propicios a lo largo del país) se logró construir una clasificación de tipos de suelo que ayudó a ir construyendo los perfiles de velocidad.

Cabe mencionar que, utilizando las velocidades de propagación de las ondas P y de S, de una forma empírica, se han determinado algunas de las propiedades mecánicas del medio atravesado, como las siguientes en la tabla 12:

Tabla 12. Determinación empírica de propiedades mecánicas del suelo

No	Propiedad Mecánica: $[V_s \text{ y } V_p]$ en m/s	Fórmula
1	Coeficiente de Poisson	$\nu = \frac{\left[\left(\frac{V_p}{V_s}\right)^2 - 2\right]}{\left[\left(\frac{V_p}{V_s}\right)^2 - 1\right]} \times \frac{1}{2}$
2	Módulo de deformación de corte máximo (MPa)	$G_{m\acute{a}x} = \frac{\rho \times V_s^2}{1000}$
3	Módulo de Young (MPa)	$E = 2 \times G_{m\acute{a}x} \times (1 + \nu)$
4	Módulo de compresibilidad volumétrica o Bulk (MPa)	$E_v = \frac{\rho \times \left(V_p^2 - \frac{4}{3} \times V_s^2\right)}{1000}$
5	Densidad (Todo tipo de suelos) (g/cm^3)	$\rho_{todos} = 0.412 \times V_s^{0.262}$
6	Densidad (Suelo de grano fino) (g/cm^3)	$\rho_{finos} = 0.742 \times V_s^{0.166}$
7	Densidad (Suelo de grano grueso) (g/cm^3)	$\rho_{gruesos} = 0.352 \times V_s^{0.283}$

A continuación, en la tabla 13 y figura 53 se presenta un ejemplo resumen de un perfil de velocidad de ondas de corte del ensayo Down-hole del pozo localizado en el eco parque de Bicentenario, en la zona de Antiguo Cuscatlán; junto a sus otras propiedades mecánicas.

Tabla 13. Resultados de perfil de velocidades de corte de pozo de Bicentenario

Prof m	Vp medio (m/s)	Vs medio (m/s)	Coeficiente Poisson	G máx (MPa)	E Young (MPa)	Ev Bulk (MPa)	Densidad (g/cm ³)		
							Todos los suelos	Grano Fino	Grano Grueso
0	372.9	177.4	0.35	59.03	159.82	182.11	1.60	1.75	1.52
1	372.9	177.4	0.35	59.03	159.82	182.11	1.60	1.75	1.52
3.1	372.9	177.4	0.35	59.03	159.82	182.11	1.60	1.75	1.52
3.1	806.8	331.4	0.40	223.91	626.27	1028.53	1.88	1.94	1.82
10.2	806.8	331.4	0.40	223.91	626.27	1028.53	1.88	1.94	1.82
10.2	1652.7	653	0.41	947.00	2665.93	4807.16	2.25	2.18	2.20
23.3	1652.7	652.8	0.41	947.00	2665.93	4807.16	2.25	2.18	2.20
23.3	1760	771.9	0.38	1342.29	3707.18	5188.57	2.35	2.24	2.31
43.9	1760	771.9	0.38	1342.29	3707.18	5188.57	2.35	2.24	2.31
43.9	1061.8	452.4	0.39	440.21	1222.99	1837.98	2.04	2.05	1.99
50	1061.8	452.4	0.39	440.21	1222.99	1837.98	2.04	2.05	1.99

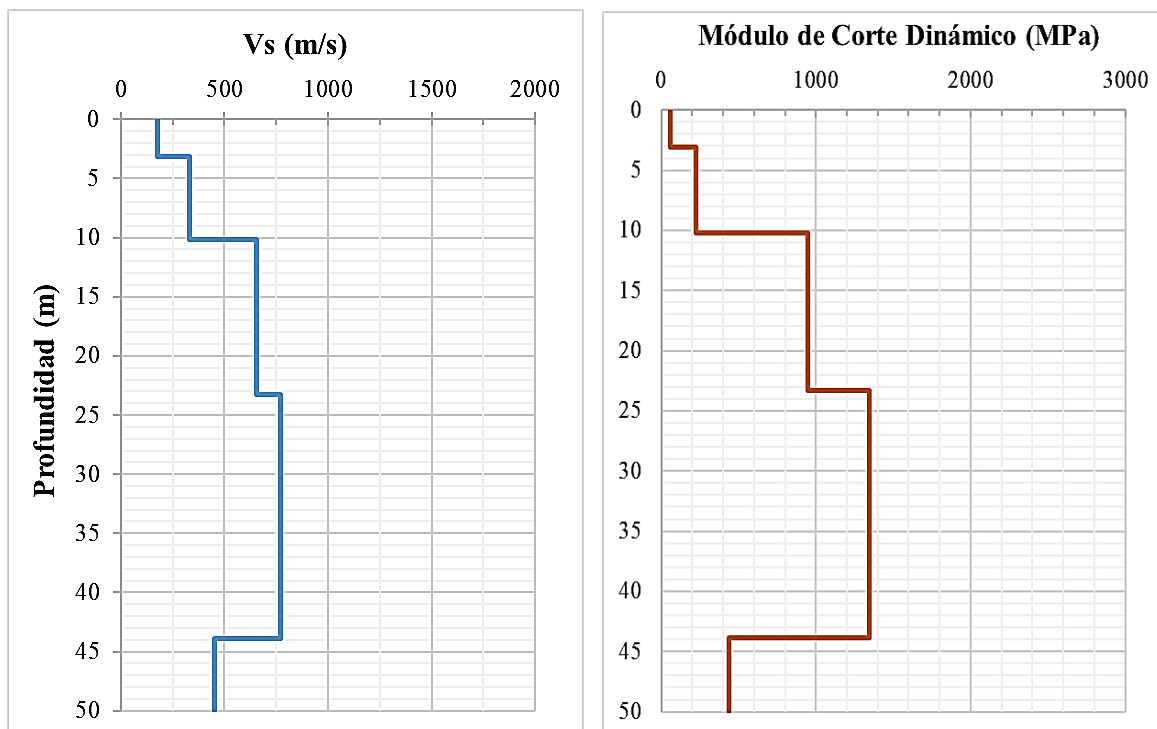


Figura 53. Perfiles de velocidad de corte y módulo de corte dinámico máximo. Fuente: MARN, 2020

Es importante mencionar, que la determinación de las densidades se ha tomado de un artículo donde se hacen regresiones numéricas de varios tipos de suelos, donde se ha medido en laboratorio el valor de la densidad. Por lo que, en ese trabajo, se compiló una base de datos y se obtuvieron ecuaciones en función de la velocidad de ondas de corte, las cuales fueron ajustadas a funciones exponenciales. Para el cálculo de las propiedades mecánicas, donde se utiliza la densidad, se fue determinando que tipo de suelo correspondía a cada cambio de estrato sísmico, utilizando la caracterización geológica definida para el perfil geológico.

Como se mencionaba anteriormente, cada cambio (o escalón) en el perfil de velocidades se le denomina estrato sísmico, el cual proviene de un promedio de velocidades (las cuales fueron midiéndose metro a metro en el ensayo Down-hole) donde se estima que no ha existido un cambio considerable en la pendiente de los tiempos de arribo de las ondas secundarias respecto a la profundidad. Por lo que el analista del ensayo, estima que podría calcularse un promedio de velocidades dentro de un estrato sísmico. Esto indica que no precisamente exista un estrato sísmico correspondiente a un estrato o unidad geológica; lo cual sería el deber ser; sin embargo, se podría inferir que, si bien podría existir cambios, pero no son tan abruptos.

Por otra parte, al momento de hacer la asignación de velocidades en los pozos geológicos, donde no se tienen mediciones directas de las velocidades (como los de la plataforma SIHI), cada estrato sísmico corresponderá esencialmente a un cambio de unidad geológica.

6.1. Determinación de Vs30 y Período Dominante

La base de datos se fue consolidando en una hoja de Excel, graficándose los perfiles de velocidad y demás propiedades. También, se determinaron los siguientes parámetros:

- El valor de Vs30 a lo largo de cada uno de los perfiles
- Estimación de profundidad donde se considera Roca (donde será aplicado un posible terremoto para propagación en los posteriores modelos analíticos 1D),
- Cálculo de periodos predominantes empíricos, y
- Asignación de tipo de suelo según NEHRP.

La expresión del cálculo de Vs30 se definió en el capítulo 3 de este documento. La estimación de la profundidad donde se consideró un estrato rocoso, se basó en dos criterios:

- a) El primero se basa en observar a que profundidad del pozo se alcanzaba una velocidad de corte de 760 m/s, la cual coincide con la velocidad con la cual se calcula la amenaza sísmica en condición de suelo firme. A este respecto, si observamos la tabla 10, esta velocidad se asocia con un tipo de suelo definido como Roca Blanda y clasificación BC, según la NEHRP 2020.

- b) Dado que principalmente en el área de San Salvador, existen intercalaciones de lavas y sedimentos, no solamente debemos de seleccionar aquellos estratos que alcancen velocidades del valor de 760 m/s, sin tomar en cuenta, la dimensión del espesor que representa tal estrato rocoso. Por ello, se definió una condición que consideraría un estrato rocoso cuando se alcance un espesor de estrato sísmico mayor a los 3 metros; ya que se estima que con ese espesor la onda sísmica se atenuaría lo suficiente para poder iniciar con las consecuentes amplificaciones de estratos más blandos hacia la superficie.

Seguidamente, se fue revisando detenidamente cada perfil de velocidades y asignando la profundidad al estrato rocoso. Lo anterior, fue clave poder realizar el cálculo de los períodos predominantes de cada perfil, utilizando la expresión simplificada siguiente sobre la estimación del período natural del sitio con una sola capa y un depósito de estrato rocoso.

$$T_g = \frac{4 \times H}{V_s}$$

donde, [Tg] es el período natural del suelo, se supone que es su período característico, [H] es el espesor de la capa blanda y [Vs] es la velocidad de corte en el suelo. A continuación, en la tabla 14 se muestra un ejemplo de ese cálculo para el perfil de velocidades del Bicentenario. Sombreado en amarillo se indica la profundidad del estrato rocoso.

Tabla 14. Determinación de valor de Vs30 y Período Dominante Tg en pozo de parque Bicentenario

Profundidad m	Vp medio (m/s)	Vs medio (m/s)		Vs referencia roca (m/s) Espesor roca > 3 (m)	760 3
0	372.9	177.4	espesor de Tg	espesor de Vs	Espesor de estrato sísmico
1	372.9	177.4	0.0225	0.0056	1
3.1	372.9	177.4	0.0474	0.0118	2.1
3.1	806.8	331.4	0	0	0
10.2	806.8	331.4	0.0857	0.0214	7.1
10.2	1652.7	652.8	0	0	0
23.3	1652.7	652.8	0.0803	0.0201	13.1
23.3	1760	771.9	0	0	0
43.9	1760	771.9	0.1067	0.0087	20.6
43.9	1061.8	452.4	0	0	0
50	1061.8	452.4	0	0	6.1
		SUMA	0.343	0.0676	
Valor de Vs30	443.48	m/s	30/[(1/177.4)+(2.1/177.4)+(7.1/331.4)+(13.1/652.8)+(6.7/771.9)]		
Período Dominante (Tg)	0.343	s	0.0056+0.0118+0.0214+0.0201+0.0087		
Tipo de suelo	C				

Finalmente, se elaboró una hoja de Excel resumen con los valores de Vs30 calculados, su fuente bibliográfica y demás propiedades. A continuación, en la tabla 15 se muestra un extracto de esta base de datos; el cual servirá para realizar las próximas interpolaciones y generación de modelos con aprendizaje supervisado para obtener mapas de Vs30 del país.

Tabla 15. Tabla resumen de Vs30 y Periodos Dominantes

DH: Down-hole, SGG: Empresa de Servicios Geológicos, LOTTI: Empresa de estudios de Las Colinas.

Nombre de pozo	Latitud N	Longitud O	vs30 (m/s)	Hrock (m)	Tg (s)	Tipo de suelo	Fuente
TR5-A	13.51210	-88.51404	756.181	0.18	0.1810	BC	Berlín
Acajutla-SGG	13.58010	-89.80430	674.685	5.20	0.0640	BC	DH-SGG
Atecozol-SGG	13.74630	-89.66590	690.558	9.30	0.0928	BC	DH-SGG
TR5	13.51200	-88.51432	691.739	11.00	0.0830	BC	Berlín
San Luis Talpa-SGG	13.46760	-89.09740	477.233	15.30	0.1934	C	DH-SGG
DIGESTYC-BID	13.71372	-89.16987	571.275	16.90	0.1523	C	DH-SGG
UNICAES-BID	13.98020	-89.54673	530.585	19.90	0.1940	C	DH-SGG
CIG	13.69870	-89.17330	447.986	20.10	0.2353	C	DH-ITAL
Perquín-BID	13.95881	-88.15773	464.501	22.20	0.2137	C	DH-SGG
La Laguna-BID	14.16456	-88.94320	646.820	22.20	0.1555	BC	DH-SGG
La Unión-SGG	13.34720	-87.87690	661.815	24.90	0.1628	BC	DH-SGG
ITCA-BID	13.67418	-89.27981	538.537	25.00	0.2018	C	DH-SGG
San Miguel-BID	13.43883	-88.15913	429.88	28.80	0.39	CD	DH-SGG
Chalchuapa-SGG	13.98470	-89.67290	532.068	29.00	0.2214	C	DH-SGG
BAL-11	13.66790	-89.30020	289.841	33.00	0.4277	D	DH-LOTTI
Ereguayquín-SGG	13.34490	-88.38950	386.939	40.80	0.3624	CD	DH-SGG
DH-García	13.66260	-88.90910	285.694	43.00	0.5270	D	DH-García
Bicentenario-SGG	13.68550	-89.25330	443.484	43.90	0.3426	C	DH-SGG
BAL-6	13.66450	-89.28740	207.677	44.00	0.6728	DE	DH-LOTTI
BAL-10	13.66500	-89.30080	157.439	44.50	0.8746	DE	DH-LOTTI
BAL-5	13.66300	-89.28490	209.976	44.50	0.6930	DE	DH-LOTTI
BAL-2	13.66220	-89.28870	241.361	48.00	0.6585	D	DH-LOTTI
BAL-7	13.66460	-89.28950	164.456	49.00	0.8626	DE	DH-LOTTI
AEIL-AAC-BID	13.69662	-89.11686	334.791	50.00	0.6955	CD	DH-SGG
BAL-3	13.66180	-89.29090	184.847	54.00	0.8997	DE	DH-LOTTI
BAL-1	13.66190	-89.28670	140.168	56.00	1.1026	E	DH-LOTTI
BAL-9	13.66540	-89.29830	160.595	75.00	1.0950	DE	DH-LOTTI
BAL-8	13.66560	-89.29540	194.892	80.00	1.0638	DE	DH-LOTTI
BAL-4	13.66400	-89.29340	188.095	81.50	1.0686	DE	DH-LOTTI

6.2. Correlación de geología con velocidades de corte

Como ejemplo, se mostrará la descripción geológica del pozo oscuro ubicado en el parque Bicentenario y la correlación existente con los valores de velocidad de ondas de corte.

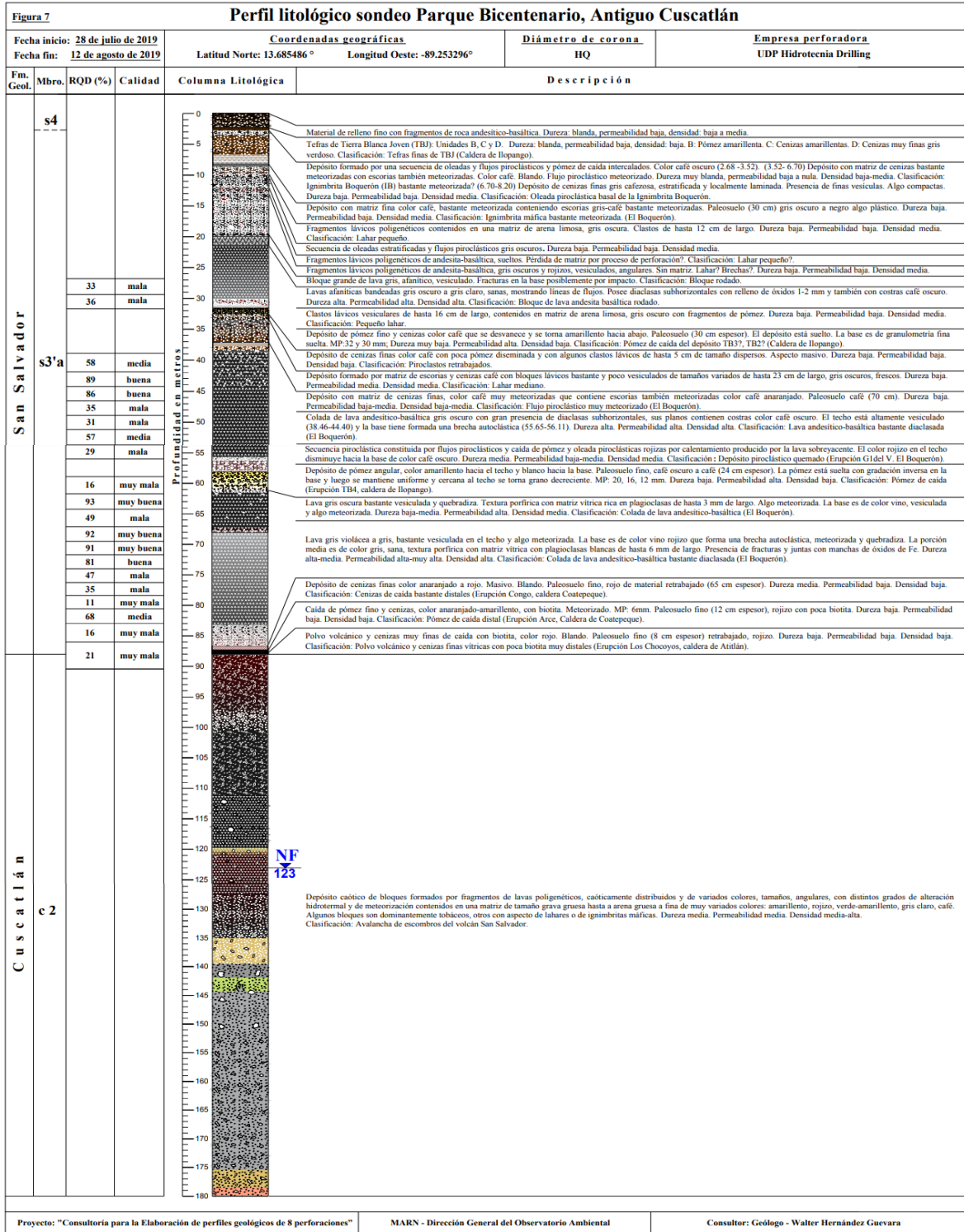


Figura 54. Caracterización geológica de perfil de pozo de parque Bicentenario. Fuente: MARN, 2020

Es importante mencionar, que la descripción y caracterización litológica a detalle de cada uno de los 15 pozos (donde se realizó el ensayo Down-hole) incluye una descripción física de los suelos y rocas identificadas, específicamente color, textura, grado de meteorización, grado de consistencia, presencia de vesículas y mineralizaciones, identificación de fracturas, juntas y fallas, así como también una estimación del grado de dureza, permeabilidad y densidad. En las secciones de la perforación donde se encuentran lavas o rocas blandas como ignimbritas se referencian medidas de *Rock Quality Designation* (RQD).

Con el objetivo de poder utilizar la información suministrada en la descripción litológica de cada uno de los perfiles estratigráficos, así como extraer muestras para realizar ensayos geotécnicos, se realiza un proceso de reclasificación sobre la base de criterios asociados a los parámetros que más influyen en las propiedades dinámicas del subsuelo. Estos son los siguientes:

- Perfil de espesores de velocidad de corte (Vs) obtenidos por ensayos Down-hole (en primera instancia, y las condiciones lo permitan) y ensayos de Microtremores.
- Edad y formaciones (miembros) geológicas de acuerdo al mapa geológico superficial 1:100,000, desarrollado por la Misión Geológica Alemana entre los años 1967-1971.
- Clasificaciones realizadas por cronología eruptiva; en profundidad.
- Descripciones físicas de las propiedades de los materiales que componen las unidades litológicas, tales como: Dureza, Densidad y Permeabilidad.
- Descripción visual de la granulometría de cada uno de los estratos identificados.

Lo anterior, servirá mucho en otro trabajo para obtener una caracterización más ajustada de la curva dinámica del módulo de reducción de corte normalizado y la curva de amortiguamiento (en función de la deformación por corte del suelo), al utilizar modelos empíricos ya preestablecidos por otros investigadores para tipos de suelos específicos.

Como se menciona anteriormente, los resultados obtenidos de los ensayos Down-hole, permiten diferenciar una secuencia de capas con diferentes comportamientos de las ondas sísmicas, lo cual obedece a una variación de los diferentes espesores de las unidades geológicas encontradas en el lugar de perforación.

Dado que en ocasiones los cambios en las velocidades de corte registradas durante el ensayo no son tan marcados, por practicidad, se realiza una simplificación del valor promedio entre profundidades específicas; por lo que discontinuidades en el perfil unidimensional del Down-hole, no indican específicamente a un cambio en el tipo de material o unidad geológica.

Sin embargo, en este ejercicio, dadas las descripciones litológicas de la perforación y de acuerdo a los criterios mostrados anteriormente, se buscó relacionar el comportamiento de la onda dependiendo el tipo de material, encontrándose en ciertos casos, variaciones claras y en otros casos (aunque se describen diferentes tipos de material) el comportamiento de la onda no varía significativamente, obteniéndose un comportamiento homogéneo.

Para cada uno de los sondeos se realizó esta correlación, en algunos casos, diferentes unidades litológicas fueron agrupadas en una sección, teniendo en cuenta que la curva Vs (m/s) no expresara un cambio abrupto en las velocidades.

Para el caso de la perforación del Bicentenario, se tiene un registro de Down-hole hasta una profundidad de 50 m por esta razón la correlación litológica se restringe hasta esta profundidad, aunque la profundidad del sondeo es de 180 m. Cabe mencionar, que el lugar exacto donde se podría extraer la muestra, se debe definir viendo los testigos en el sitio; sin embargo, este ejercicio proporciona un indicio para enfocarse en los estratos objetivo que se desean caracterizar, dada su importancia en el cambio de velocidad de ondas de corte. A continuación, en la figura 55 se muestra la correlación y reclasificación de las unidades geológicas del perfil del parque Bicentenario.

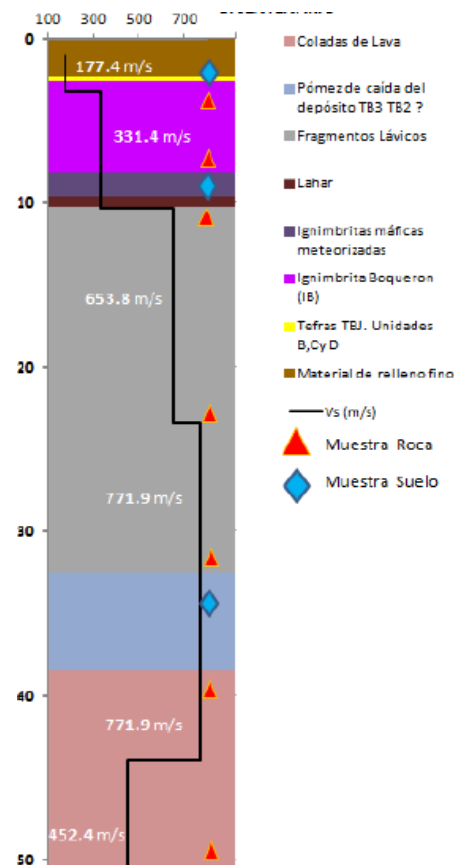


Figura 55. Correlación y reclasificación de perfil de pozo de parque Bicentenario. Fuente: MARN, 2020

6.3. Integración de datos de USGS a base de datos

Un total de 519 puntos se lograron registrar a nivel nacional, los cuales se concentran mayormente en el Área Metropolitana de San Salvador (OPAMSS). A continuación, en la figura 56 se muestran todos los 519 puntos recabados, los cuales han sido clasificados de acuerdo a los tipos de suelo según la NEHRP 2020.

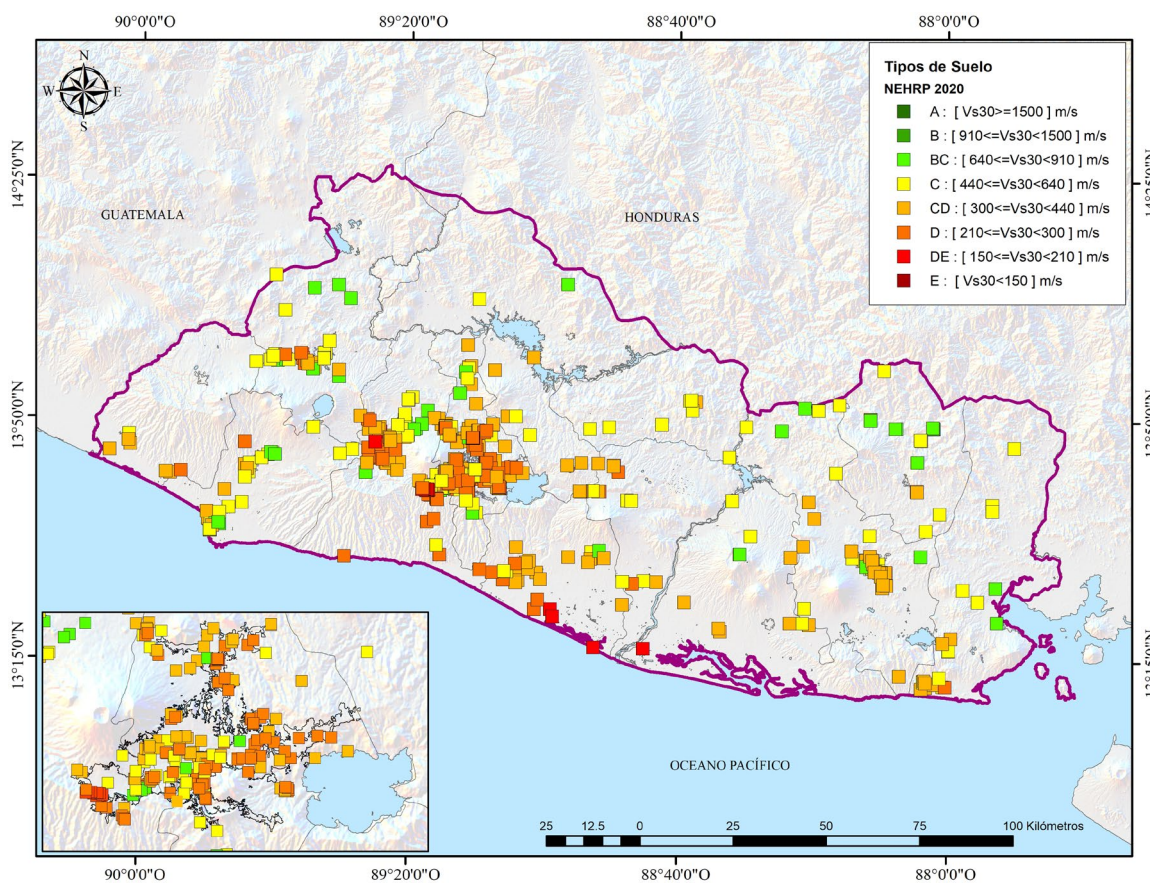


Figura 56. Ubicación geográfica de 519 puntos de base de datos. Fuente: MARN. 2020

Debido a la aún limitada cantidad de información en otros lugares del país, se optó de una manera conservadora, por agregar algunos valores de Vs30 calculados a partir de las pendientes del terreno (Wald, 2009), correspondiente al producto del USGS, específicamente para valores superiores e iguales a 640 m/s; siendo cuidadoso en los lugares, como los grábenes, donde se sabe que no existen afloramientos rocosos y que corresponden a zonas de depósitos de productos volcánicos no consolidados.

Lo anterior, se hizo seleccionando espacialmente una muestra aleatoria de 100 puntos, con el único fin de enriquecer la información a ser tratada con las técnicas de aprendizaje supervisado, al momento de llevar a cabo los modelos de interpolación, y así garantizar una distribución homogénea de los datos. En la figura 57 se observa la ubicación de estos 100 puntos, los cuales se localizan principalmente en zonas altas.

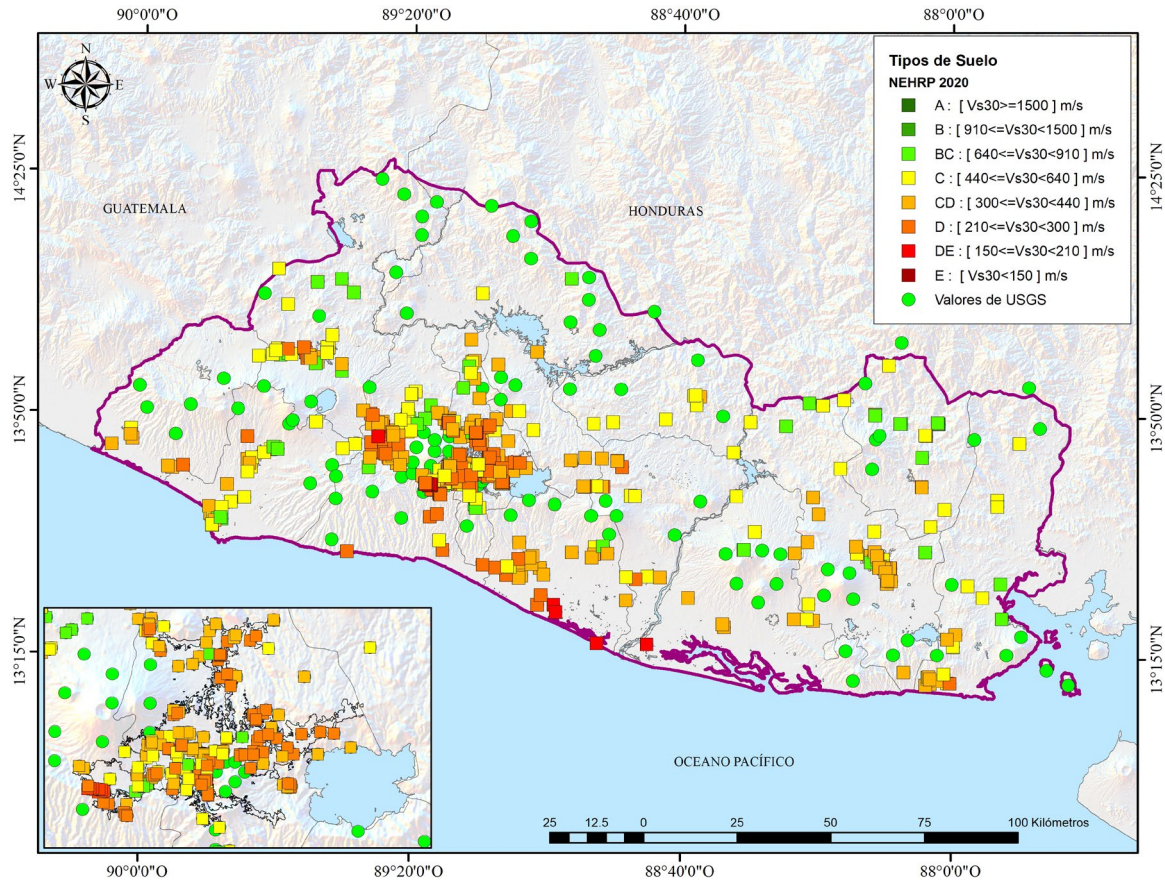


Figura 57. Ubicación geográfica de 619 puntos de base de datos y puntos USGS. Fuente: MARN. 2020

7. Generación de mapa integrado de Vs30

En este apartado, se abordará el tema sobre el procedimiento seguido para la generación de un modelo que represente de manera óptima un mapa, que registre los valores de Vs30 a nivel nacional, el cual será traducido de acuerdo a los rangos de velocidades establecidos en la clasificación de tipos de suelos según la NEHRP 2020.

Esta representación se busca que se realice a través de la creación de modelos confiables que permitan, mediante el entrenamiento de variables conocidas, poder predecir valores en lugares donde no se tenga información, tratando de obtener la más mínima incertidumbre. Estos modelos se basarán en la utilización de herramientas de aprendizaje supervisado o automático (machine learning) manejando como variable objetivo el valor numérico de Vs30 (al tratarse de modelos de regresión), o bien considerando variables categóricas como las clases de tipo de suelo correspondientes a los rangos de velocidades que propone la NEHRP 2020 (al tratarse de modelos de clasificación).

Primeramente, se abordará la teoría de los modelos y como éstos pueden representar adecuadamente los resultados y obtener así un mapa más confiable.

7.1. Modelos de aprendizaje supervisado

Básicamente, los modelos son generalizaciones representativas utilizadas para la predicción. En otras palabras, consisten en la práctica de poder tomar un conjunto de datos y hacer una especie de generalización sobre los mismos, por lo que se puede usar para hacer una predicción en lugares donde no se tiene información sobre la variable de interés.

Un modelo de confianza es aquel que puede generalizar la tendencia que se observa en los datos y no necesariamente que imite o simule demasiado bien los datos con los que estamos entrenando. A esto se le llama que un modelo está sobre ajustado (*overfitting*).

Este sobreajuste es muy frecuente en los modelos de aprendizaje automático y representa uno de los principales problemas al usar estas técnicas. Es decir, que podemos hacer un muy buen trabajo al explicar exactamente los datos que le dimos al modelo, pero cualquier otro conjunto de datos nuevos (asociados a la variable de interés) que le introduzcamos al modelo realizará un trabajo bastante malo en la predicción, lo cual es problemático ya que nuestro objetivo es la predicción. Por lo tanto, no es necesario que hagamos un buen trabajo al predecir los datos que ya tenemos (porque ya los tenemos).

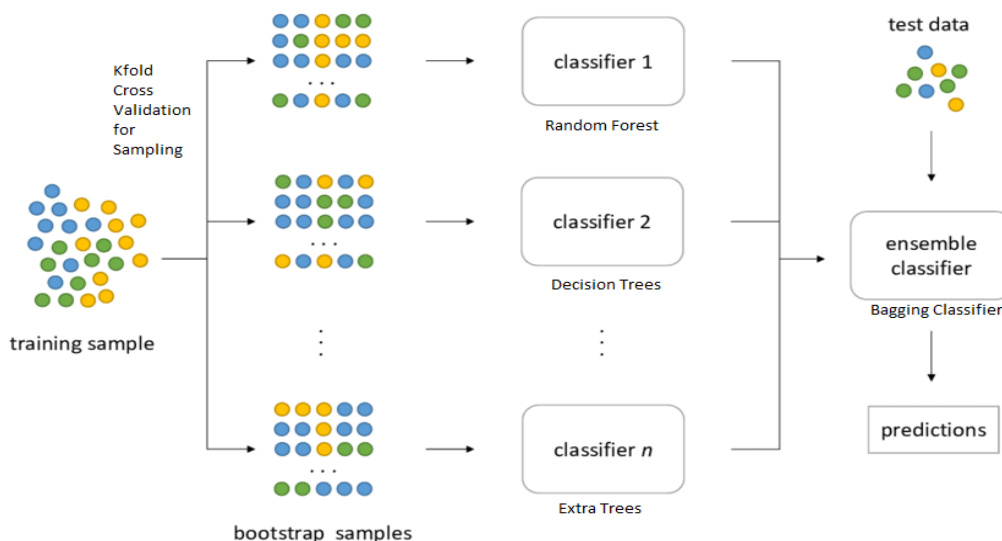
Entonces, si no es una buena generalización del trabajo, no seremos capaces de predecir en lugares en los que aún no tenemos datos y ese, al final, es realmente nuestro objetivo. Entre los modelos más comunes que se utilizan para realizar predicción en ubicaciones geográficas son los siguientes:

- Random Forest con Predicción espacial (RFsp)
- Técnica de Redes de Neuronas Artificiales (RNA)

Ambas técnicas serán desarrolladas en este trabajo, con lo que se tendrá al final diferentes mapas de Vs30, de los cuales, posteriormente, se superpondrán entre sí, con el objetivo de lograr construir un Mapa Integrado de Vs30, capturando las bondades y virtudes de ambos algoritmos; mejorando así la calidad de las zonificaciones representadas, asociándose a lugares con respuesta sísmica homogénea

La idea anterior radica en resolver el siguiente cuestionamiento: ¿se podrá aplicar alguna estrategia para aumentar la precisión de los resultados que se obtengan de los modelos que se definan con las dos técnicas RFsp y RNA? La respuesta es sí, en ambas técnicas, se utilizan métodos denominados *bagging* y *boosting*, ambos son ejemplos de *ensemble methods* (los métodos de conjunto utilizan múltiples algoritmos de aprendizaje para obtener un mejor rendimiento predictivo que el que se podría obtener de cualquiera de los algoritmos de aprendizaje constituyentes), que en resumen es un método para utilizar combinaciones de otros modelos. Ver figura 58.

Para explicar mejor el algoritmo que se ha utilizado para el incremento de la precisión en el modelamiento del Vs30, se presenta la siguiente analogía. Un paciente desea obtener un diagnóstico a una enfermedad basado en sus síntomas. En lugar de consultar con un solo doctor, decide consultar a varios. Si cierto diagnóstico se repite con frecuencia entonces, el paciente podría escogerlo como el mejor. Este es el corazón de la metodología, el diagnóstico final se basa en una mayoría de votación, donde cada doctor tiene la oportunidad de votar equitativamente. Ahora, se puede reemplazar cada doctor por un modelo que estime el Vs30 y se obtiene la esencia del algoritmo empleado. Es importante mencionar que, de forma intuitiva, una votación de mayoría hecha por un grupo grande de doctores será más confiable que una votación hecha por un grupo pequeño.



Bagging Classifier Process Flow

Figura 58. Diagrama de flujo del método de ensamblaje. Fuente: MARN. 2020

A continuación, se detallan los tres modelos que se han tomado en consideración para obtener el mapa integrado de Vs30 de El Salvador.

7.1.1. Random Forest con Predicción Espacial

7.1.1.1. Teoría sobre Random Forest

Random Forest (RF) es un algoritmo de machine learning que surge como evolución de los árboles de decisión. Los árboles de decisión clasifican los datos de un modelo realizando una serie de particiones binarias, permitiendo realizar predicciones futuras. Los RF estarán formados por un gran número de árboles de decisión, creados a partir de la técnica del *bagging* o *bootstrap aggregating*, relacionada a la toma de muestras con reemplazamiento (ver figura 58). Este algoritmo mejora los árboles de decisión, creando un modelo más fiable evitando problemas como el *overfitting* (García Ruiz, M. 2018).

RF es bueno para lidiar con este problema de sobreajuste debido a su naturaleza de aleatoriedad. Una de las razones por las que se llama así "Bosque aleatorio" es porque en todas las partes de su algoritmo se incluye dicha aleatoriedad. Es decir, no se proporcionan aleatoriamente a todos los árboles todos los datos, ni tampoco se brindan aleatoriamente en cada árbol todas las variables, sino que, en cambio, se toman aleatoriamente todo tipo de decisiones a lo largo de ese proceso de análisis. De modo que, al final, se obtienen modelos que generalmente no están sobreajustados, que es en particular uno de los poderes reales de RF, en comparación con muchos otros métodos de *machine learning* que existen.

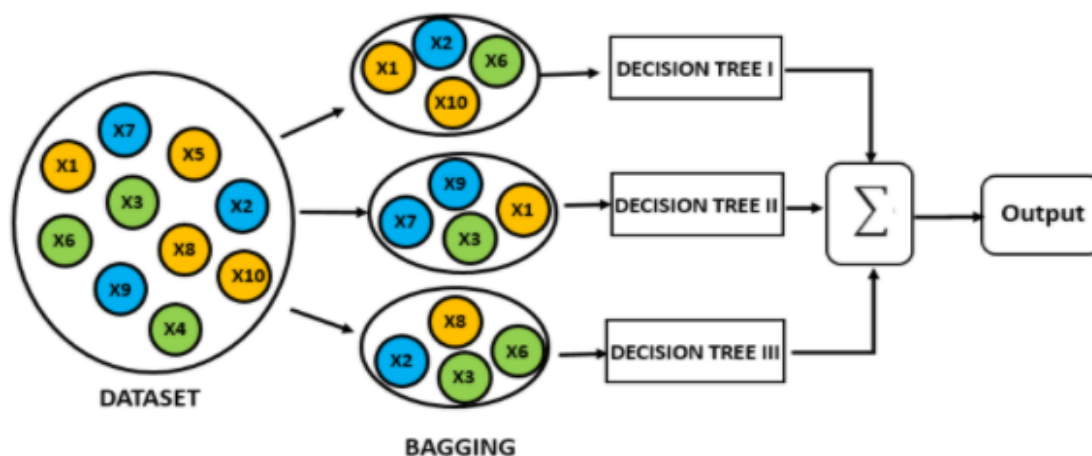


Figura 58. Algoritmo de RF utilizando técnica de Bagging para reducir la varianza de las predicciones.
 Fuente: <https://bookdown.org/content/2031/ensambladores-random-forest-parte-i.html#random-forest>

Prácticamente, con RF se busca en promediar los resultados de un conjunto de modelos individuales (árboles de decisión). Cada modelo se estima partir de un remuestreo de la muestra original. Además, en cada división de los árboles no se consideran todas las variables de entrada sino un subconjunto elegido al azar. La herramienta, es capaz de tratar modelos muy complejos con un gran número de variables y observaciones (García Ruiz, M. 2018).

Un aspecto importante a destacar, es que con estos modelos de aprendizaje supervisado se pueden simular dos tipos de problemas: a) Clasificación y b) Regresión. En RF se puede generar modelos para predecir variables objetivas que procedan de problemas de Clasificación o Regresión.

Con problemas de Clasificación se predice una variable categórica, por ejemplo, puede ser algo como la distribución de una especie, las causas de los incendios forestales, la presencia o ausencia de algo, podría ser un tipo binario, la clasificación de la tierra, etc. Por otra parte, con la Regresión, se estaría prediciendo una variable continua, por ejemplo, gasto en atención médica, tasa de delincuencia, tasa de mortalidad, ganancias de ventas, algún número continuo a lo largo de una recta numérica, etc.

7.1.1.2. Algoritmo de RF (Tomado de García Ruiz, M. 2018)

Cada árbol de decisión se formará de la siguiente manera:

1. Teniendo un conjunto N observaciones diferentes, se elegirá una muestra N aleatoria con **reemplazamiento**. Esta técnica, recibe el nombre de *bootstrapping* y es una técnica utilizada en varios algoritmos de *machine learning*. Además, introduce aleatoriedad al algoritmo, ya que cada árbol se forma de manera ligeramente distinta.
2. Dadas las M variables de entrada, en cada nodo se seleccionarán de forma aleatoria $p < M$ variables. Este número p, será constante en todo el proceso de formación del árbol e introducirá el segundo elemento de aleatoriedad en el algoritmo
3. Se dejará crecer el árbol, sin podar hasta la máxima extensión posible.

Como se ha mencionado anteriormente, en la formación del bosque se introduce una aleatoriedad para reducir la varianza en el modelo. La aleatoriedad, reducirá la correlación entre árboles, ya que en la formación de los mismos cada uno partirá de una muestra ligeramente distinta y en cada nodo la selección de las variables también será diferente.

La formación del bosque, será igual si estamos tratando un problema de clasificación o de regresión. Una vez construido el bosque, se utilizará para realizar la predicción, siendo esta la media entre las predicciones de cada árbol, en el caso de un problema de regresión. En un problema de clasificación la predicción será la clase más votada entre todos los árboles.

Por lo tanto, se puede establecer que los RF dependen de dos parámetros fundamentales:

- Ntree: Número de árboles que forman el bosque. Es un parámetro importante puesto que los experimentos indican que a mayor número de árboles menos problemas de *overfitting* deberían ocurrir. Este parámetro también tiene efecto en la precisión de la predicción. Como es lógico, a mayor número de árboles mejor será la predicción, puesto que el número de datos para hacer el promedio es mayor. Sin embargo, existe un valor para el cual, el error ya no disminuye y se estanca, aumentando solo el tiempo del algoritmo
- Mtree: Número de variables “p” que se seleccionarán en cada partición (o nodo) de cada árbol del bosque. Este importante parámetro introduce una gran aleatoriedad al modelo, y dependiendo del número de variables seleccionadas se obtendrán mejores resultados en las predicciones. Sus valores recomendados son: $p^{1/2}$ para un problema de clasificación y $p/3$ para un problema de regresión.

Se puede establecer que la tasa de error de los RF está relacionada con estos parámetros. Al reducir el número de variables “p”, se reduce la correlación entre los árboles, ya que en cada nodo se tienen menos posibilidades entre las que elegir; y, por lo tanto, disminuirá la varianza y más preciso será el árbol. Sin embargo, al reducir p, también se reduce la precisión del árbol.

7.1.1.3. Evaluación de desempeño de un modelo en RF

A continuación, se presentan 3 conceptos importantes y necesarios de conocer para poder evaluar y emitir un juicio sobre los resultados que se obtienen de un modelo generado con RF. Estos serán determinados posteriormente al momento de la generación del modelo para obtener el mapa de VS30, con los datos de valores de velocidad de ondas de corte.

7.1.1.3.1. OOB Out of Bag Error (OOB)

El *Out of bag error* (OOB) es una medida de error aplicada a modelos que utilizan la técnica del *bootstrapping*, que sirve para evaluar su rendimiento. Esta medida se determina dentro del subconjunto de datos seleccionados para entrenamiento (figura 59). Por lo que se hace la elección aleatoria de los N datos con reemplazamiento en una muestra (*Sample Set 1*), luego utilizando el *bootstrapping*, se selecciona otra muestra (*Sample Set 2*) y así sucesivamente. Dentro de cada muestra, se realiza una fracción muestral, estableciendo un tamaño de datos para ser tratados “dentro de la bolsa” (In Bag) y otros “fuera de la bolsa” (OOB). Es precisamente estos últimos datos en OOB a los cuales se les determina su error.

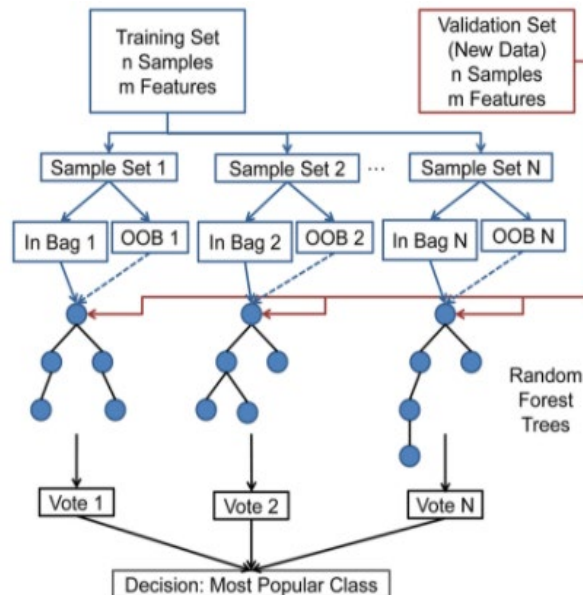


Figura 59. Proceso de selección de datos de entrenamiento con reemplazo para calcular el OOB error.
Fuente: <https://bookdown.org/content/2031/ensambladores-random-forest-parte-i.html#random-forest>

Prácticamente, esta medida se entiende como ¿qué tan bien puede predecir cada árbol las características excluidas? El tamaño de la fracción muestral, puede ser calibrada de acuerdo a la naturaleza de los datos. Comúnmente, cada árbol solo obtiene un subconjunto aleatorio de los datos de entrenamiento (aproximadamente $2/3$). Entonces, después de entrenar el modelo, el $1/3$ restante (que se excluyó) se vuelve a ejecutar a través de ese árbol y luego probamos para ver qué tan bien ese árbol puede predecir ese $1/3$ que se excluyó. Por lo tanto, el OOB error, representa el error de predicción cometido por el bosque cuando se tienen en cuenta este conjunto de variables que han quedado “fuera de la bolsa”.

Un dato importante, es que el OOB disminuye al aumentar el número de árboles, llegándose a un valor asintótico para un elevado número de árboles. No obstante, al llegar a un número determinado, el OOB, se estanca y ya no disminuye más al aumentar el parámetro del número de árboles y puede producir problemas como el *overfitting*.

Cuando se decide por tomar muestras con reemplazo, los dos valores muestrales son independientes. Prácticamente, esto significa que lo que obtenemos en el primero no afecta lo que obtenemos en el segundo. Matemáticamente, esto significa que la covarianza entre los dos es cero. Por otro lado, en el muestreo sin reemplazo, los dos valores de muestra no son independientes. Matemáticamente, esto significa que la covarianza entre los dos no es cero. (<https://web.ma.utexas.edu/users/parker/sampling/repl.htm>).

Sin embargo, en el modelo que se genera para el mapa de Vs30, se tratará de buscar el tamaño de fracción muestral óptimo, mediante un análisis de sensibilidad.

7.1.1.3.2. Validación Cruzada

La validación cruzada, o *Cross validation*, es un método de evaluación en modelos en los que se tratan un gran número de datos para garantizar la independencia de los resultados (predicción) de la partición inicial de datos del training y test (entrenamiento y prueba).

Prácticamente, esta validación significa que ¿qué tan bien puede predecir el bosque las características que no se utilizan en el entrenamiento? Comúnmente, se excluye de todo el conjunto de datos un valor del 10 % con fines de validación. Este porcentaje se selecciona al azar y no se usa en absoluto para entrenar el modelo. Entonces, después que el modelo ha sido entrenado, se ejecuta ese 10% a través de todo el bosque, donde esta vez, en lugar de solo el árbol individual, podemos observar qué tan bien puede predecir el bosque ese 10%.

No es raro ver un R cuadrado realmente alto para los datos de entrenamiento y un R cuadrado más bajo para los datos de validación. Esto es lo que se espera ver. Por lo tanto,

el modelo hace un mejor trabajo al predecir los datos que se usaron para enseñar que los datos que nunca antes se habían visto.

El método *holdout*, es el método de validación cruzada más simple (ver figura 60). Este método se basa en dividir los datos en un set de entrenamiento (training) y en un set de prueba (test). Se utilizará sólo el set de entrenamiento para construir el modelo de predicción. Una vez construido el modelo, se hará una validación con los datos asignados como de prueba.



Figura 60. Método de validación cruzada Holdout.

Fuente: <https://bookdown.org/content/2031/ensambladores-random-forest-parte-i.html#random-forest>

El problema que surge con este método, es la elevada variancia que tiene, puesto que el resultado puede resultar muy diferente dependiendo de la división inicial de los datos.

Para corregir el problema de la dependencia de la división inicial de los datos, existen otro método de *cross validation*, llamado *K-fold*. En este método *K-fold* se dividen los datos en K subconjuntos y se repite el *holdout method* k veces de manera que, en cada iteración, se escoge uno de los K subconjuntos como el set de prueba y los K-1 restantes, se utilizan como set de entrenamiento.

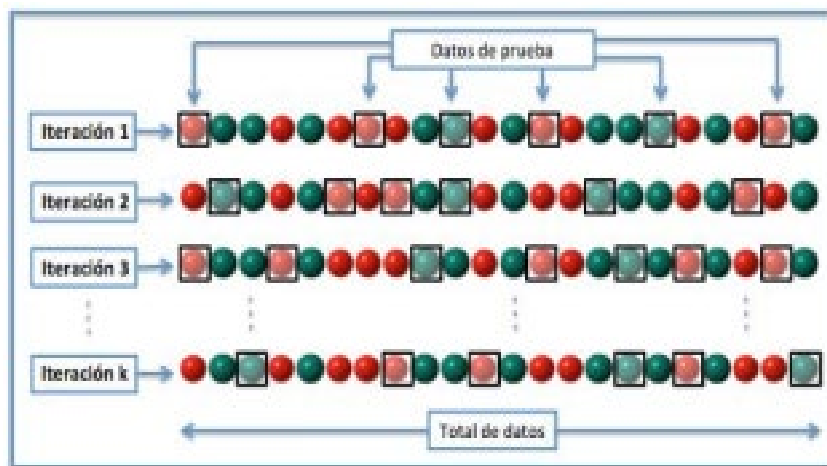


Figura 61. Método de validación cruzada K-fold.

Fuente: <https://bookdown.org/content/2031/ensambladores-random-forest-parte-i.html#random-forest>

La ventaja de este método, es que el resultado depende menos de la división que se hace de los datos en los sets de entrenamiento y prueba. Una vez realizados todos los modelos, se obtendrá el resultado final como la media aritmética de las K iteraciones.

7.1.1.3.3. Importancia de las variables

Este concepto se entiende como ¿Qué tan bueno es el desempeño de cada variable en la división de los árboles? La variable más importante, significa que es un diferenciador realmente importante. A estas variables, se les conoce como “variables explanatorias o explicativas” o simplemente como “covariables”, las cuales pueden ser atributos, características de distancia y/o proximidad. Estas covariables deben de tener un grado de influencia en la determinación de la variable objetivo.

Si bien se habló antes en la medida de error de OOB, de que se hace la selección de muestras con reemplazo de los datos de entrenamiento, también se realiza una selección aleatoria de un subconjunto de variables explicativas. Donde estas últimas, ayudarán a construir la formulación o ecuación de regresión junto a la variable de objetivo.

En el caso de RF, para medir la importancia de las variables, para cada árbol que forman el bosque, se calcula el MSE (error cuadrático medio) con las variables del OOB, que son las que han quedado fuera. Se podría resumir este concepto en ver cómo afecta a la salida del modelo cuando se realizan cambios en las variables de entrada. Las variables de entrada que más variabilidad produzcan en la salida, serán aquellas que más influencia tengan y, por tanto, aquellas que mejor explicarán el modelo y serán más importantes

7.1.1.4. Conceptos adicionales en RF

A continuación, se presenta un resumen de algunos conceptos claves para explicar la creación y procesamiento de un modelo desarrollado en RF:

- En RF, una rama se podría convertir en un predictor débil. Pero cuando se combinan un montón de predictores débiles, se convierten en predictores fuertes y, al final, la mayoría de votos ganará y podrá identificar correctamente la prueba.
- En resumen, cada árbol solo obtiene un subconjunto aleatorio de variables de entrenamiento y un subconjunto aleatorio de variables explicativas; de lo contrario, predeciría perfectamente, pero solo dentro del conjunto de datos de entrenamiento, por lo que no podría predecir cosas de las que no solía aprender. Así que, se introduce la aleatoriedad y luego se combinan todos los predictores débiles en un predictor fuerte y se clasifica con la mayoría de votos ganados.
- Cada vez que se modela, tenemos que evaluar el desempeño de nuestro modelo; o qué tan bien nuestro modelo está haciendo la predicción. Cuando se predice

una regresión, se obtiene un valor R^2 . Esto ayuda a comparar qué tan lejos de los valores observados están nuestros valores predichos. En cambio, cuando se realiza una clasificación, se obtendrá lo que se llama una matriz de confusión.

- En la matriz de confusión, se proporcionan dos medidas diferentes: a) Sensibilidad y b) Precisión.

Para Sensibilidad, se indica la proporción de cuántas predicciones correctas sobre la predicción total que predicen el valor específico del objetivo (verdadero positivo sobre verdadero positivo + falso positivo).

Por precisión, no solo se preocupa por la cantidad de veces que el objetivo fue identificado correctamente, sino que también por la cantidad de veces que NO se identificó correctamente el objetivo, es decir, la proporción de la suma de los verdaderos positivos y falsos negativos, sobre el total.

- La RF es muy mala para la extrapolación.
El modelo solo sabe lo que ha visto, es decir, todo el mundo dentro del cual vive ese modelo, por lo que si entrenamos el modelo con datos que no son representativos de todos los rangos de valores posibles; no podrá predecir algún valor fuera del rango con ese uso para predecir, por lo que realmente debe pensar si RF es apropiado si desea extrapolarlo, probablemente no esté bien.
- En resumen, el flujo de trabajo del modelado sería el siguiente:
Paso 0: prepara tus datos. Paso 1: entrenar un modelo. Paso 2: Evaluar el desempeño del modelo (Verifique los errores OOB y R^2 o la matriz de confusión). Paso 3: volver a entrenar con diferentes parámetros. Paso 4: comparar modelos. Paso 5: Repetir una y otra vez hasta que encuentre un modelo de predicción realmente bueno. Paso 6: utilice el mejor modelo para predecir valores desconocidos.
- En teoría, cuantos más datos haya visto el modelo implementado, mejor será la generalización. Entonces, si se entrenó el modelo con el conjunto completo de datos u observaciones que se tiene disponible, debería generalizarse mejor que un modelo que solo vio, por ejemplo, el subconjunto de entrenamiento o de validación (por ejemplo, ~ 90%) del conjunto de datos completo.

El problema con esto (y la razón por la que se divide los datos en subconjuntos de entrenamiento/validación/prueba en primer lugar), es que se desea poder hacer afirmaciones estadísticas en cuanto a la precisión de los datos no vistos. Tan pronto como se vuelva a entrenar un modelo con todos los datos, ya no es posible hacer tales afirmaciones.

7.1.1.5. Consideración de Auto-correlación espacial

El bosque aleatorio y técnicas similares de aprendizaje supervisado ya se utilizan para generar predicciones espaciales, pero la ubicación espacial de puntos (geografía) a menudo se ignora en el proceso de modelado. La autocorrelación espacial que pueda existir en los datos, indica que, si no se toma en cuenta en el análisis, las predicciones pueden estar sesgadas, y esto podría resultar en la construcción de modelo no óptimo. (Hengl, T. et al, 2018)

Modelar la relación con las covariables y la autocorrelación espacial de forma conjunta utilizando técnicas de aprendizaje supervisado es relativamente novedoso y no está del todo resuelto. Un aspecto importante, es que al utilizar únicamente las coordenadas geográficas como covariables en un modelo de RF, puede no ayudar al proceso de predicción, ya que conduce a límites lineales en el mapa resultante que están directamente relacionados con la configuración del plan de muestreo. Por lo que se necesita un uso más sensato y robusto del espacio geográfico (Hengl, T. et al, 2018).

Es acá donde surge una adición al algoritmo de RF, específicamente en la consideración de covariables que involucren no solo las coordenadas de las observaciones, sino que las distancias de influencia (buffer) entre los puntos de observación. A este nuevo aspecto se le conoce como RFsp (*Random Forest spatial prediction*).

Estas distancias buffer se utilizan como variables explicativas (covariables), incorporando así efectos de proximidad geográfica en el proceso de predicción, similar al concepto basado en interpolaciones kriging.

Por tanto, la clave para hacer que la RF sea aplicable a los problemas de las estadísticas espaciales reside también en la preparación de medidas geográficas de proximidad y conectividad entre observaciones, de modo que se tenga en cuenta la Autocorrelación espacial (Hengl, T. et al, 2018).

7.1.1.6. Aplicación de RFsp para generar mapa de Vs30

Una vez entendido la metodología y conocido las virtudes del algoritmo de RF, así como su incursión en las predicciones espaciales, asumiendo que los datos presentan autocorrelación espacial; se procederá a aplicar el procedimiento para generar un modelo de Vs30 (a nivel nacional) utilizando como base los datos recabados y calculados de velocidades de ondas de corte en los 30 m superiores, en los puntos de la base de datos, mostrados anteriormente.

Se abordará este proceso como un problema de Regresión, donde se utilizará como variable objetivo (continua) el valor calculado anteriormente de Vs30 para cada uno de los 619 puntos, mostrados en la figura 57.

7.1.1.6.1. Asignación de pesos

Como se mencionó en un inicio, dentro de la información recabada, fue evidente la calidad de la misma, la cual se tomó en consideración al momento de ser utilizada. Esto se desarrolló a través de la asignación de pesos, asociada a la confiabilidad de su determinación.

Básicamente, la consideración de datos con mayor peso que otros, funciona de la siguiente manera: las observaciones con pesos más grandes se seleccionarán con mayor probabilidad en las muestras *bootstrap* para los árboles. De igual manera, se realizaron estimaciones considerando una igual asignación de pesos, a efectos de comparación.

Estos pesos han sido asignados cualitativamente dada la precisión de cada técnica y, por tanto, traducidos de acuerdo al porcentaje de error en su medición. Por ejemplo, los puntos donde se ha hecho ensayos geofísicos Downhole, presentan un peso mayor, a comparación de aquellos lugares que se han extraído de fórmulas empíricas o a partir de correlaciones respecto a la pendiente, tal es el caso de los puntos extraídos del modelo propuesto por el Servicio Geológico de Estados Unidos (USGS, sus siglas en inglés). En la figura 62, se muestran los 619 puntos recabados y su técnica asociada.

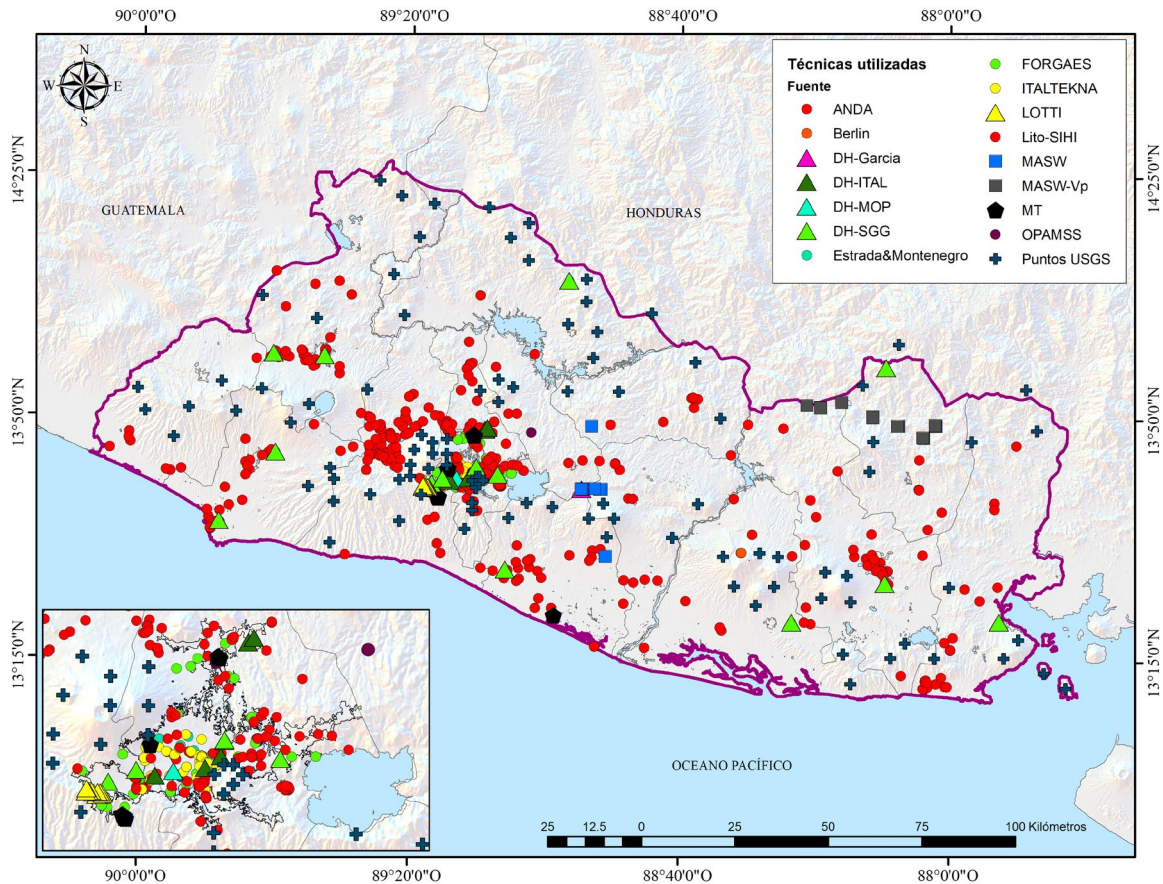


Figura 62. Ubicación de 619 puntos de base de datos y puntos USGS con sus técnicas. Fuente: MARN. 2020

Las técnicas son las siguientes: Down-hole (en símbolo de triángulo), MASW (en símbolo de cuadrado) (Análisis multicanal espectral de las ondas sísmicas de superficie usando fuentes activas, estimando ondas primarias Vp y ondas secundarias de corte Vp), MASW – Vp (procesándose la Vp, y con una aproximación del módulo de Poisson dinámico se estimó la Vs), Microtremores (en símbolo de pentágono, medición de ruido ambiental con fuente pasiva) y los demás puntos se derivan de perfiles geológicos de pozos (en círculos).

Por lo tanto, estos valores observados con sus respectivos pesos sirvieron como puntos de entrenamiento al modelo. En la tabla 16, se muestra la jerarquía de los pesos asignados a cada una de las técnicas relacionada a su calidad de determinación en el valor de Vs30; para luego ser utilizados en el procedimiento de RFsp.

Los pesos se derivan a partir de valores de calidad de la técnica utilizada. Esta calidad ha sido discutida entre el grupo investigador, considerando que los pesos se debiesen calcular utilizando la varianza de medida inversa ($1/\text{Calidad}^2$). Entre más cercano a 1 el peso, los valores de esos puntos se seleccionarán con mayor probabilidad al momento de hacer la fracción muestral con los datos de entrenamiento.

Tabla 16. Asignación de pesos a técnicas utilizadas para determinar el valor de Vs30 en 619 puntos de base de datos

Técnica	Calidad	Pesos Varianza de medida inversa ($1/\text{Calidad}^2$)
Down-hole	1	1
MASW	1.15	0.756
MT	1.15	0.756
MASW-Vp	1.25	0.640
Perfiles geológicos	1.30	0.592
USGS	2.50	0.160

7.1.1.6.2. Selección de covariables

En el proceso de construcción del modelo, se deben asignar covariables las cuales ayudaran a crear una ecuación de regresión para ayudar al modelo a adaptarse a valores más equilibrados que influyan o estén relacionadas directamente con el valor de la velocidad de las ondas de corte de las unidades geológicas.

Para ello, se siguen algunos aspectos expresados en el trabajo de Matsuoka et al. 2006, donde desarrolla regresiones numéricas para la determinación de fórmulas empíricas de Vs30 (empleando 2000 perfiles de pozos en Japón), en función de unidades geomorfológicas y características geográficas, como variables explanatorias o covariables.

Estas covariables son:

- **La elevación del terreno (en metros)**, donde se asumen que tiene relación con los cambios de granulometría de los sedimentos en los ríos, los cuales cambian de grano grueso en el río arriba a granos finos en el río abajo.
- **La distancia más cercana a ríos principales (en metros)**; asociada a la geomorfología formada por inundaciones, donde los depósitos de sedimentos (de consistencia blanda) cerca de un río se vuelven más gruesos y más delgados a medida que aumenta la distancia al río.
- **El ángulo de la pendiente del terreno (en grados y multiplicada por 1000)**; donde se considera que pendientes elevadas indicarían estratos firmes o rocosos (formados por avalanchas de escombros) y valles planos indicarían suelo o sedimentos, y la transición a pendientes intermedias.

Lo anterior, indicaría que los materiales más densos, con mayores velocidades de onda de corte asociadas, son capaces de mantener pendientes pronunciadas, mientras que sedimentos y grandes depósitos de flujos piroclásticos en general, poco consolidados y con velocidades de onda muy bajas, son depositados en valles relativamente planos. (D. J. Wald and Allen 2007).

- **La distancia más cercana a línea de costa (en metros)**; relacionada al espesor de las capas sedimentarias en los deltas y tierras bajas costeras.
- **La distancia más cercana a una montaña de la era del Terciario (en metros)**; relacionada también al espesor de las capas sedimentarias que influyen en el Vs30 de las tierras bajas en valles y terrazas a lo largo de los ríos en regiones montañosas.

De igual manera, se asocia a los efectos cuenca relacionadas a la generación de ondas superficiales en sus bordes y el consecuente efecto de sitio. Asimismo, esta característica presenta una correlación con la profundidad del lecho rocoso debajo de la capa sedimentaria de la unidad geomorfológica formada durante el Cuaternario. Por lo que, una montaña o cerro formado durante el período más antiguo (Pre-Terciario o Terciario) es la referencia que se seleccionará para medir las distancias más cercanas a las observaciones de la variable objetivo.

- **La geología superficial a partir del mapa geológico** de la Misión Geológica Alemana (Bosse et al. 1978), la cual es extraída de manera cualitativa, a partir de una reclasificación hecha por García, I.; 2006; y por lo tanto tratada como una variable categórica convertida en una simple clasificación numérica.

7.1.1.6.3. Algoritmo de RFsp para generar predicciones

A continuación, se irá explicando el código del algoritmo de RFsp que se ha construido para poder crear el modelo de predicción de valores de Vs30 a nivel de todo el país, utilizando los 619 registros de Vs30 consolidados en la base de datos, la asignación de pesos a los datos específicos y la preparación de las covariables.

El código se ha desarrollado con el lenguaje de programación contenido en la plataforma Rstudio del paquete estadístico R, versión 3.6.4. Este forma parte del sistema GNU Project, por lo que forma parte de un proyecto colaborativo. Además, está formado por un amplio conjunto de paquetes que permiten utilizar una amplia gama de herramientas estadísticas como modelos lineales, no lineales, análisis de series temporales, etc.

Otra gran herramienta es la posibilidad de realizar gráficas muy completas. Todas las librerías se obtienen de la página web oficial de *CRAN projects*, sin embargo, se ha utilizado Rstudio que es una plataforma IDE (*Integrated Development Environment*) que facilita el manejo del entorno de R.

Primero se prepara la ruta del espacio de trabajo y las diferentes librerías a ser utilizadas en el código. Acá figuran las librerías `randomForest()`, `ranger()` y `tuneRanger()`, las cuales ayudan a realizar la ejecución del algoritmo de RF.

```
setwd("C:/Users/lmixco/Documents/2020/Febrero/Interpolacion_Vs/Random_Forest")
list.of.packages <- c("plyr", "parallel", "randomForest", "quantregForest",
"plotKML", "GSIF", "RCurl", "raster", "rgdal", "geoR", "gstat", "scales",
"gdistance", "entropy", "lattice", "gridExtra", "intamap", "maxlike",
"spatstat", "DescTools", "gdistance")
new.packages <- list.of.packages[!(list.of.packages %in%
installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages, dependencies = TRUE)

library(rsample); library(sp); library(raster); library(dplyr); library(h2o);
library(rgdal); library(maptools); library(mapview); library(ggplot2);
library(dismo); library(rgeos); library(gstat); library(GSIF); library(rgdal);
library(plyr); library(randomForest); library(plotKML); library(scales);
library(RCurl); library(parallel); library(geoR); library(lattice);
library(gridExtra); library(maxlike); library(spatstat); library(entropy);
library(gdistance); library(DescTools); library(ranger); library(tuneRanger)
```

Luego, se ingresan las observaciones de los 619 puntos a través de un archivo *.csv, el cual posee 4 columnas con las coordenadas geográficas, el valor obtenido de Vs30 y el valor del peso asignado (ver tabla 17).

Las coordenadas geográficas son convertidas a coordenadas planimétricas UTM (*Universal Transverse Mercator*) utilizando el mismo *Datum* con un huso horario de la zona 16 Norte, correspondiente a El Salvador. Lo anterior, se hace para manejar siempre las mismas unidades métricas las cuales serán utilizadas también en la generación de las covariables.

Tabla 17. Extracto de archivo csv de datos observados de Vs30

X	Y	vs30	Weight.sd
167394.9967	1522813.755	344.828	1
172452.8603	1525369.72	418.03	1
172629.8351	1526940.895	638.806	0.592
172968.5603	1524295.724	358.566	1
174937.7242	1538475.348	723.068	0.592
176809.9257	1532471.531	713.759	0.756
182435.4597	1516771.192	364.344	0.592

De igual manera, se ingresa un archivo *.csv con las covariables a utilizar, el cual posee 8 columnas, con las coordenadas UTM, la pendiente, la geología, la elevación, la distancia al río más cercano, la distancia a la costa y la distancia a las montañas (ver tabla 18). Cada una de las covariables fue calculada manualmente utilizando herramientas de geoprocésamiento, análisis espacial y de proximidad, en un programa de SIG.

Tabla 18. Extracto de archivo csv de datos observados de Vs30 con las covariables

X	Y	SLOPE_1000	GEO_SUELO	ELEV_MTS	RIO_DIS_MTS	COSTA_DIS_MTS	MONTA_DIS_MTS
167395.00	1522813.76	0.00	2	4.64	1829.23	3538.14	12814.30
172452.86	1525369.72	90.14	2	36.83	7279.93	7945.87	7677.48
172629.84	1526940.90	218.66	3	45.69	8020.79	9459.51	6265.73
172968.56	1524295.72	113.19	3	45.35	7514.89	7193.76	8401.25
174937.72	1538475.35	509.90	3	522.10	3048.25	20921.12	1518.59
176809.93	1532471.53	552.27	3	510.79	7791.26	16218.99	673.24
182435.46	1516771.19	55.90	3	34.30	15009.47	4656.72	10385.02

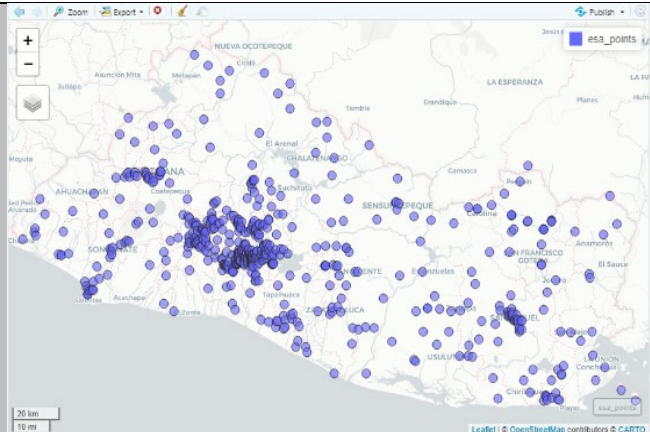
Luego, se genera a partir de los puntos observados un archivo tipo "DataFrame", para poder ser ingresado en el algoritmo de RF, al cual se le asignará la proyección UTM (epsg:32616). También se utilizará un polígono de todo el país, así como un buffer de 15km, con el objetivo de poder generar una grilla equiespaciada que funja como espacio de trabajo que comprenda dentro los 619 observados. A continuación, se muestra el código para este proceso.

```

# Create a data frame
esa_points <- data.frame(esa)
# Convert to SpatialPointsDataFrame
coordinates(esa_points) <- ~X+Y
# Set the projection
proj4string(esa_points) <-
CRS("+init=epsg:32616")

# View the station location using
the mapview function
mapview(esa_points)

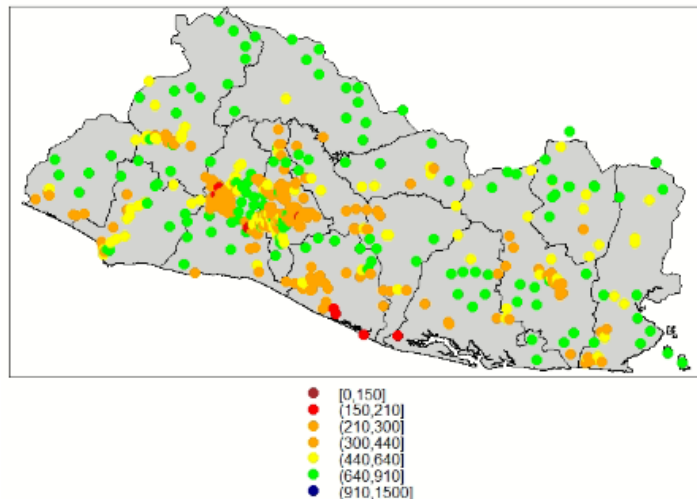
```




```

shapeESA <-
readShapePoly("C:/Users/lmixco/Documents/2020/Febrero/Interpolacion_Vs/Random_Forest/Csv/7_covariables/RANDOM_FOREST/SLV_adml.shp",proj4string =
CRS("+init=epsg:4326"))
shapeESA_buffer15km <-
readShapePoly("C:/Users/lmixco/Documents/2020/Febrero/Interpolacion_Vs/Random_Forest/Csv/7_covariables/RANDOM_FOREST/SLV_adml_buffer15km.shp",proj4string =
CRS("+init=epsg:4326"))
TA <- CRS("+init=epsg:32616")
shape_ESA_UTM <- spTransform(shapeESA, TA)
shape_ESA_buffer15km_UTM <- spTransform(shapeESA_buffer15km, TA)
# define groups for mapping
cuts <- c(0,150,210,300,440,640,910,1500)
# set up a palette of interpolated colors
blues <- colorRampPalette(c('brown', 'red', 'orange', 'orange', 'yellow',
'green', 'dark blue'))
pols <- list("sp.polygons", shape_ESA_UTM, fill = "lightgray")
spplot(es_a_points, 'vs30', cuts=cuts, col.regions=blues(7), sp.layout=pols,
pch=20, cex=2)

```

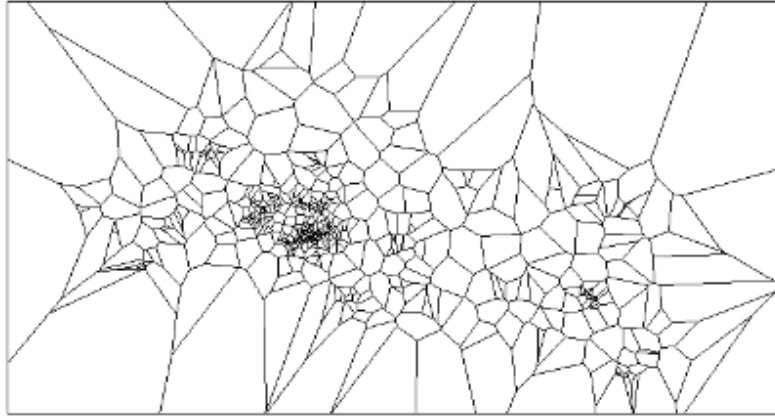


Seguidamente, se procederá a realizar un procedimiento para generar una grilla que se extiende a lo largo de todo el polígono de El Salvador (que es donde se quiere realizar la predicción). Lo anterior, tiene como objetivo generar un mapa de puntos (a partir de dicha grilla) y que sirva como base para construir 619 covariables adicionales basadas en el cálculo de las distancias buffer, considerando la autocorrelación espacial existente entre los puntos de observación. Estas nuevas covariables se sumarán a las 6 covariables mencionadas anteriormente y serán ingresadas dentro de la formulación del modelo.

El procedimiento del mallado se basará en la creación de una interpolación de los datos con la técnica *"Inverse Distance Weighted"*, así como también en la creación de un mapa de polígonos de Thiessen, para destacar la influencia espacial de cada punto observado. Cabe mencionar, que esta interpolación es solamente un paso necesario para generar el archivo de malla de un tipo *"SpatialPixelsDataFrame"* que ayudará a calcular las distancias buffer entre los puntos observados. A continuación, se muestra el detalle del código:

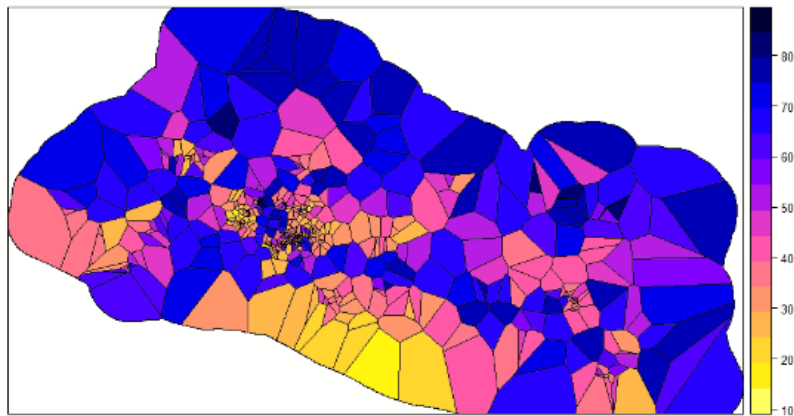
```

#Creacion de poligonos de Thiesen
v <- voronoi(esa_points)
plot(v)
  
```



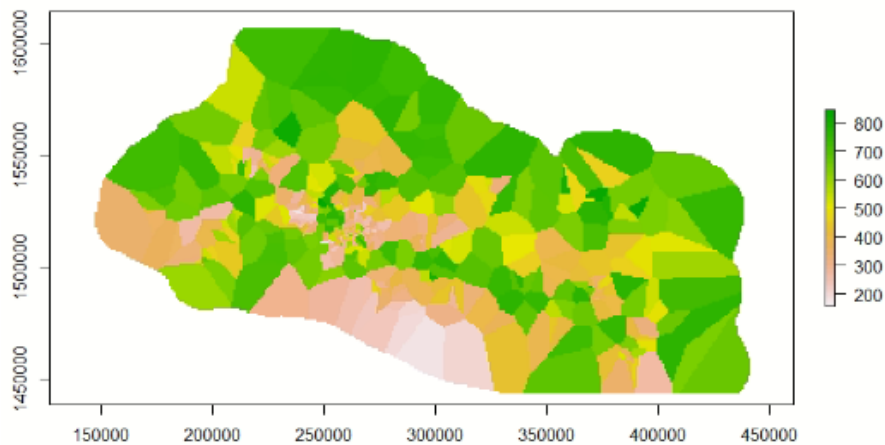
```

ca <- aggregate(shape_ESA_buffer15km_UTM)
#Asignacion de colores de acuerdo a variable de interes
vca <- intersect(v, ca)
spplot(vca, 'vs30', col.regions=rev(get_col_regions()))
  
```

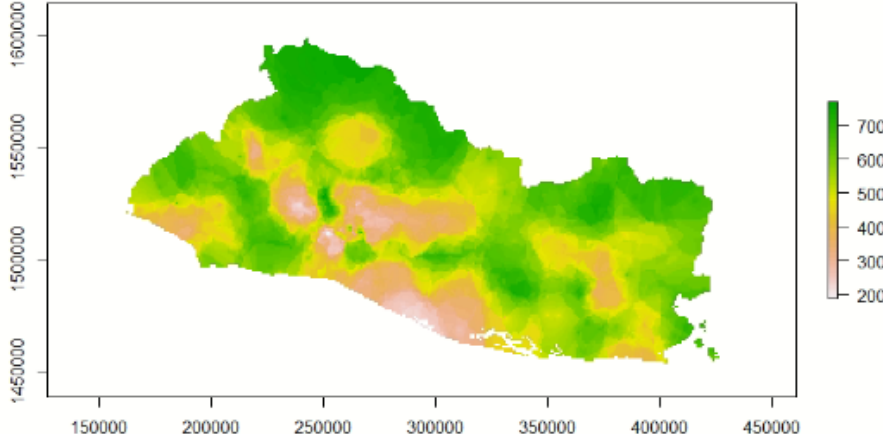


```

#Rasterizado de polígono con los puntos
r <- raster(shape_ESA_buffer15km_UTM, res=1000)
vr <- rasterize(vca, r, 'vs30')
plot(vr)
  
```



```
#Interpolacion de puntos mediante la tecnica Inverse distance Weighted
#con el objetivo de generar un archivo grid y data de tipo
"SpatialPixelsDataFrame"
gs <- gstat(id = "vs30", formula=vs30~1, locations = esa_points, nmax=7,
set=list(idp = .5))
gs
nn <- interpolate(r, gs)
## [inverse distance weighted interpolation]
nnmsk <- mask(nn, shape_ESA_UTM)
plot(nnmsk)
```



```
#Creacion de mallado del tipo "SpatialPixelsDataFrame", que servira para
calculo de buffer distances
grid2 <- as(nnmsk, "SpatialPixelsDataFrame")
```

Una vez obtenida la malla con el formato especificado y con un número de puntos de 32,528 que cubre todo El Salvador, se procede a determinar las covariables de distancias buffer. A continuación, se muestra el código utilizado para esta tarea haciendo uso de la librería GSIF. Este proceso tarda ya que genera 619 mapas en cuadrícula. El valor de la variable objetivo “vs30” ahora se puede modelar como función de distancias buffer.

```
#Spatial Autocorrelation
#Buffer Distances
grid.distP <- GSIF::buffer.dist(esa_points["vs30"], grid2[1],
as.factor(1:nrow(esa_points)))
dnP <- paste(names(grid.distP), collapse="+")
```

Luego se procede a la formulación completa de las covariables, contabilizando 627.

```
covariables <- paste(names(esalkmcovs), collapse = "+")
fm.esa <- as.formula(paste("vs30 ~ ", covariables, " + ", dnP))
fm.esa
vs30 ~ SLOPE 1000 + GEO SUELO + ELEV MTS + RIO DIS MTS + COSTA DIS MTS + MONTA DIS MTS +
layer.1 + layer.2 + layer.3 + layer.4 + layer.5 + layer.6 + layer.7 + layer.8 + layer.9 +
layer.10 + layer.11 + layer.12 + layer.13 + layer.14 + layer.15 + layer.16 + layer.17
+ layer.18 + layer.19 + layer.20 + layer.21 + layer.22 + layer.23 + . . . + layer.617 +
layer.618 + layer.619
```

El análisis posterior es similar a cualquier análisis de regresión. Lo que se hace primero es superponer los puntos y las cuadrículas, para así crear una matriz de regresión; es decir conjugar todas las covariables.

```
ov.esa <- over(esa_points["vs30"], grid.distP)
ov.esa1 <- over(esa_points["vs30"], esalkmcovs)
rm.esa1 <- do.call(cbind, list(esa_points@data["vs30"], ov.esa, ov.esa1))
```

Una vez teniendo toda esta información se procede a ejecutar el algoritmo de RF. En esta ocasión, se utilizará la librería randomForest() la cual no utiliza el factor de asignación de pesos y se ejecutará para parámetro de comparación con la otra librería ranger().

Primeramente, haremos una ejecución con todos los datos sin discriminar en entrenamiento o de prueba (test); con el objetivo de visualizar como se encuentra el desempeño máximo del OOB error. Luego, se procede a dividir la data en subconjuntos de datos de entrenamiento y de test, estableciendo una fracción de 2/3 para datos de entrenamiento y 1/3 restante en datos para test. Primero se utiliza una semilla para garantizar que los resultados, cifras, etc. sean reproducibles. El valor del parámetro “Mtry” se calcula preliminarmente como $(p/3)$, es decir, el número de variables entre 3

```
set.seed(101)
vs30_split <- initial_split(rm.esa1, prop = 2/3)
vs30_train <- training(vs30_split)
vs30_test <- testing(vs30_split)

# RF model con la libreria randomForest, utilizando toda la data
m1 <- randomForest(formula = fm.esa, data = rm.esa1, ntree=1000, mtry = 207)
m1
Call:
 randomForest(formula = fm.esa, data = rm.esa1, ntree = 1000,          mtry
 = 208)

      Type of random forest: regression
      Number of trees: 1000
No. of variables tried at each split: 208

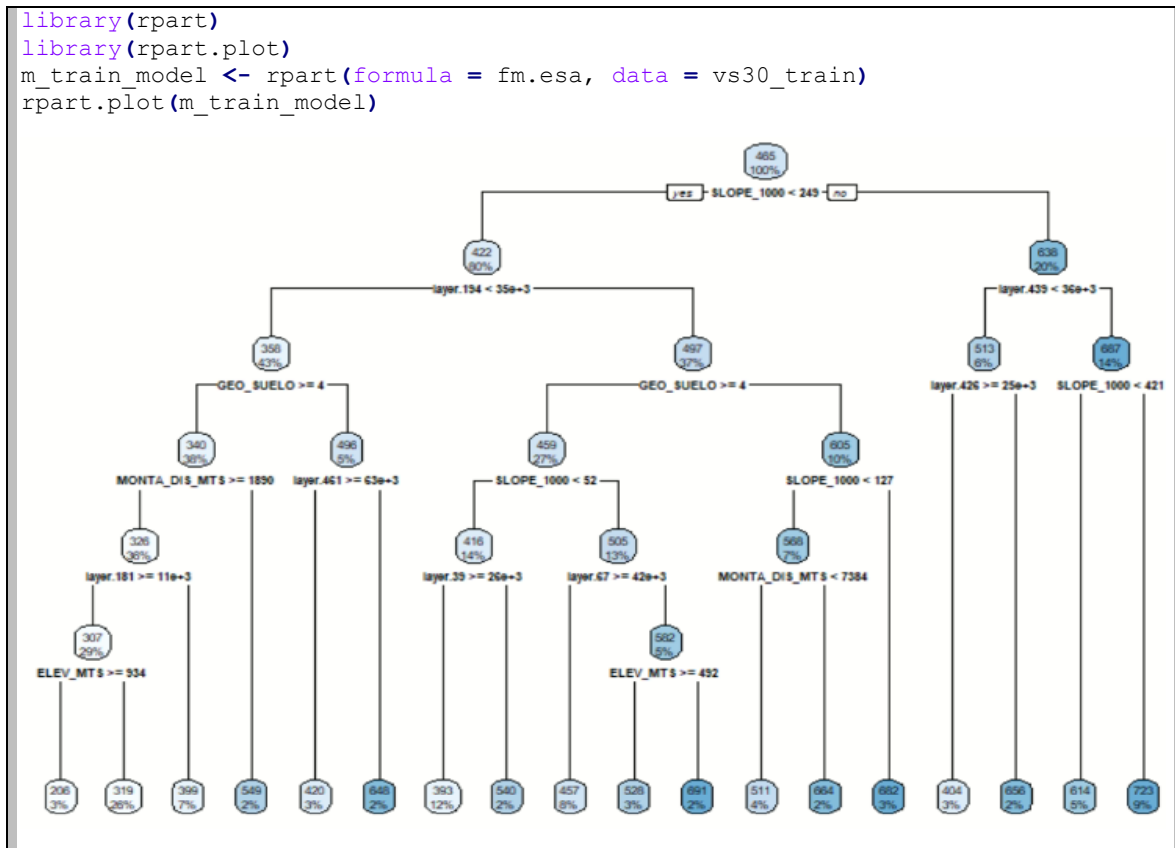
      Mean of squared residuals: 11145.26
      % Var explained: 62.69

# Solo entrenamiento
m_train <- randomForest(formula = fm.esa, data = vs30_train, ntree=1000, mtry =
207)
m_train
Call:
 randomForest(formula = fm.esa, data = vs30_train, ntree = 1000,
mtry = 208)

      Type of random forest: regression
      Number of trees: 1000
No. of variables tried at each split: 208

      Mean of squared residuals: 13257.41
      % Var explained: 54.07
```


También existe la opción de observar un extracto de los árboles de decisión, sus nodos, ramas y hojas, al hacer un gráfico entre las covariables y la data de entrenamiento. Esto se hace con la librería `rpart()` y `rpart.plot`. A continuación, el detalle:



Seguidamente, se hace una predicción preliminar para observar como nuestro modelo se desempeña con los datos de prueba (test) que no han sido entrenados. Se dice preliminar, debido a que no se ha realizado todavía un proceso de sensibilidad o tuneo de los parámetros (llamados también hiperparámetros) del modelo; donde se buscarán valores óptimos para poder lograr la minimización del error en la mayoría de lo posible.

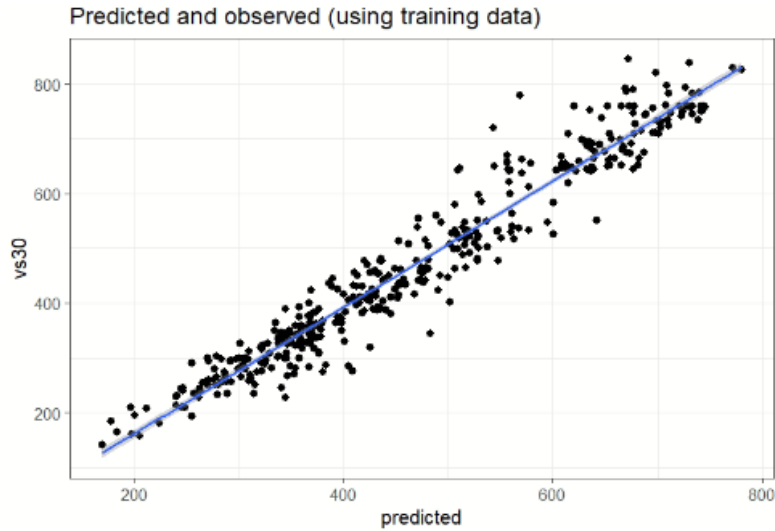
Continuando con las predicciones, se plotea en un gráfico comparativo los valores predichos con los valores reales, tanto para los datos de entrenamiento como los valores de prueba. Es lógico que con los datos de entrenamiento (utilizados para entrenar el modelo) obtengamos menos dispersión; no es el caso, para los datos de prueba (test) los cuales el modelo nunca ha visto. A continuación, se presenta el detalle del código.

```
#Hacer una predicción sin realizar un tuneo de los hiperparámetros
pred_randomForest_mtrain <- predict(m_train, vs30_test)

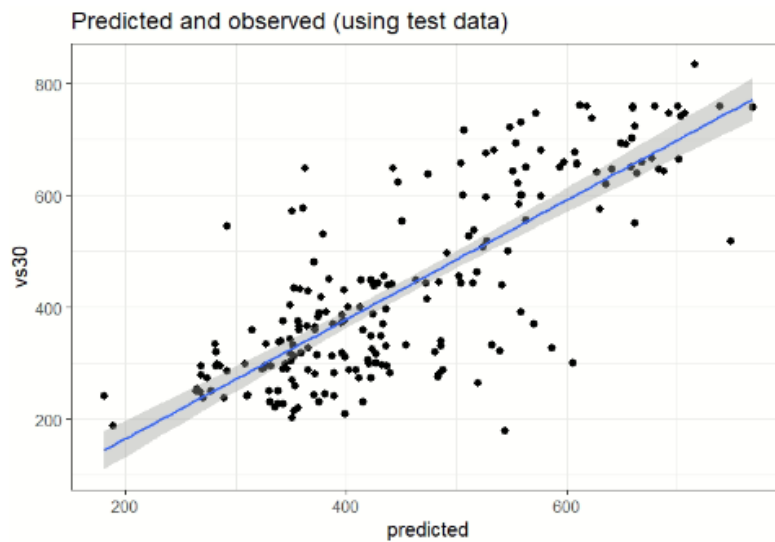
rmse <- function(x) sqrt(sum((x - vs30_test$vs30)^2))
rmse(pred_randomForest_mtrain)
[1] 1506.279
```

```
head(pred_randomForest_mtrain)
      2      3      5      7      9     12
377.3087 474.1962 661.9172 371.4879 371.3941 683.6427

#Grafica comparativa de predicciones preliminares
vs30_train %>%
  mutate(predicted = predict(m_train, vs30_train)) %>%
  ggplot(aes(predicted, vs30)) +
  geom_point(colour = "#ff6767", alpha = 1) +
  stat_summary(fun.data=mean_cl_normal) +
  geom_smooth(method='lm', formula= y~x, level=0.95) +
  labs(title = "Predicted and observed") + theme_bw(15)
```



```
vs30_test %>%
  mutate(predicted = predict(m_train, vs30_test)) %>%
  ggplot(aes(predicted, vs30)) +
  geom_point(colour = "#ff6767", alpha = 1) +
  stat_summary(fun.data=mean_cl_normal) +
  geom_smooth(method='lm', formula= y~x, level=0.95) +
  labs(title = "Predicted and observed") + theme_bw(15)
```



Como se esperaba, en las pasadas figuras se pudo apreciar el desempeño de nuestro modelo utilizando los diferentes tipos de datos. Cabe desatacar, que se trabajó con un intervalo de nivel de confianza del 95% para predicciones de un modelo lineal. Básicamente, indica el "rango" en el que estarían nuestras predicciones si repitiéramos el experimento (muestreo) una y otra vez. Un muestreo conduce a una sola línea recta de predicciones; teniendo en cuenta la variabilidad de los datos, las zonas grises de las figuras indican un rango de posibles líneas rectas. Al establecer el nivel en .95, indicaría que "si repitiéramos el muestreo una y otra vez, el 95% de las líneas de regresión estarían dentro de esa zona gris".

Ahora ha llegado el momento de hacer un ajuste (tuneo) o análisis de sensibilidad de los hiperparámetros que dependen del algoritmo de RF. Esto radica en el argumento que, al realizar las predicciones preliminares, hemos introducido los parámetros al modelo sin saber si podrían ser los óptimos (asociado a reducir el error), asimismo, se piensa que este proceso se comporte en cierta manera como una caja negra al que se le introducen unos valores y de la que obtenemos una respuesta.

Los hiperparámetros a tunear son los siguientes: a) mtry, b) min.node.size y c) sample fraction. Como se infirió anteriormente, los hiperparámetros mtry, min.node.size y sample.fraction determinan el grado de aleatoriedad y deben ajustarse.

El parámetro sample.fraction especifica la fracción de observaciones que se utilizarán en cada árbol. Las fracciones más pequeñas conducen a una mayor diversidad y, por lo tanto, a árboles menos correlacionados, lo que a menudo es deseable. El valor del número de árboles (ntree) no se revisará ya que es conocido que a mayor número de árboles los errores tienden a ser menores, hasta cierto límite, donde la respuesta ya se vuelve constante. A continuación, se muestra el código.

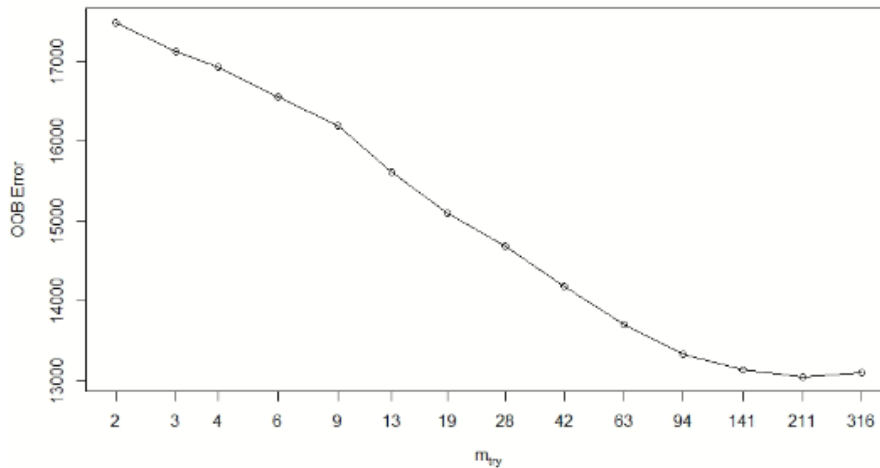
```
#Tuning inicial
#Si se desea ajustar inicialmente el parámetro mtry se puede usar el comando
randomForest::tuneRF
#para una evaluación rápida y sencilla.
#La función tuneRF necesita indicar de manera separada los datos predictores y
#la variable objetivo (respuesta).

# Se genera una matrix o dataframe predictor donde se excluya el valor de la
variable objetivo "vs30" y solo queden las covariables
features <- setdiff(names(vs30_train), "vs30")
set.seed(123)

m_train_tune <- tuneRF(
  x      = vs30_train[features],
  y      = vs30_train$vs30,
  ntreeTry = 1000, # número de árboles
  mtryStart = 2,   # valor inicial
  stepFactor = 1.5, # factor de paso para incremento de mtry
  improve   = 0.001, # OBB error para si ha tenido una mejora del 1%
  trace     = TRUE, plot=TRUE, doBest=FALSE #
)
```

```

mtry = 2 OOB error = 17486.94
Searching left ...
Searching right ...
mtry = 3 OOB error = 17124.08
0.02074983 0.001
mtry = 4 OOB error = 16927.58
0.01147547 0.001
mtry = 6 OOB error = 16560.04
0.02171252 0.001
mtry = 9 OOB error = 16195.72
0.02199974 0.001
mtry = 13 OOB error = 15618.98
0.03561095 0.001
mtry = 19 OOB error = 15097.81
0.03336715 0.001
mtry = 28 OOB error = 14683.58
0.027437 0.001
mtry = 42 OOB error = 14184.9
0.03396132 0.001
mtry = 63 OOB error = 13712.39
0.03331083 0.001
mtry = 94 OOB error = 13327.79
0.02804746 0.001
mtry = 141 OOB error = 13136.98
0.01431711 0.001
mtry = 211 OOB error = 13052.32
0.006444559 0.001
mtry = 316 OOB error = 13097.39
-0.003453584 0.001
  
```



De los resultados anteriores se puede descifrar que un valor óptimo de “mtry” se encuentra en el rango de 141 y 316 (ya que se obtiene un menor valor de OOB), encontrándose dentro el valor de 208 ($p/3$) utilizado anteriormente en la predicción preliminar. Nuestra tarea será identificar qué valor de “mtry” es el más adecuado para nuestros datos. Para ello, se construirá una búsqueda más precisa de este parámetro, incluyendo los dos restantes. En esta ocasión, se hará uso de la librería ranger(), el cual es el mismo algoritmo de RF, pero con la obtención de resultados más rápidos. Por el momento, no se colocarán la asignación de pesos. A continuación, el detalle en el código.


```
#Para generar un "grid search", vamos a construir nuestra grilla de hyper-
parámetros.

#Vamos a buscar a través de modelos diferentes (aproximadamente 220) con
variaciones en mtry, mínimo tamaño de muestras en el nodo (entre 1 y 10), y
tamaño de la fracción muestral (3 pruebas: 70%, 75% y 80%, para ser considerado
como los datos In Bag, y el resto como OOB)

# hyperparameter grid search
hyper_grid <- expand.grid(
  mtry      = seq(141, 316, by = 10),
  node_size = seq(1, 10, by = 1),
  sampe_size = c(.70, .75, .80),
  OOB_RMSE  = 0
)

# total number of combinations
nrow(hyper_grid)
[1] 220

#Ahora iteramos a través de las combinaciones y aplicamos 1000 árboles

for(i in 1:nrow(hyper_grid)) {
  m_train_model_search <- ranger(
    formula      = fm.esa,
    data         = vs30_train,
    num.trees    = 1000,
    mtry         = hyper_grid$mtry[i],
    min.node.size = hyper_grid$node_size[i],
    sample.fraction = hyper_grid$sampe_size[i],
    seed        = 123 # Notese el seteo de la semilla
  )

  # add OOB error to grid
  hyper_grid$OOB_RMSE[i] <- sqrt(m_train_model_search$prediction.error)
}

hyper_grid %>%
  dplyr::arrange(OOB_RMSE) %>%
  head(10)

  mtry node_size sampe_size OOB_RMSE
1   190         5         0.80 113.7655
2   170         5         0.75 113.7697
3   200         5         0.75 113.7744
4   200         3         0.80 113.7863
5   290         3         0.80 113.8170
6   290         4         0.80 113.8302
7   260         5         0.80 113.8521
8   290         2         0.80 113.8559
9   220         5         0.80 113.8610
10  240         4         0.75 113.8747
```

En la lista mostrada en el resultado, nos indica que para nuestros datos de entrenamiento la combinación de mtry, min.node.size y sample fraction, es 190, 5 y 0.80, respectivamente.; ya que se obtiene el menor valor de OOB_RMSE de 113.7655.

Por lo que se procede a realizar las predicciones con dichos parámetros. Antes de realizar las predicciones, veremos un histograma de como varia el error de OOB y finalmente conocer cuáles son las variables con mayor influencia en el proceso de predicción de la variable objetivo Vs30.

```
set.seed(123)

optimal_ranger <- ranger(formula = fm.esa,
  data = vs30_train,
  num.trees = 30000,
  mtry = 190,
  min.node.size = 5,
  sample.fraction = 0.80,
  importance = 'impurity') #para evaluar importancia de variables

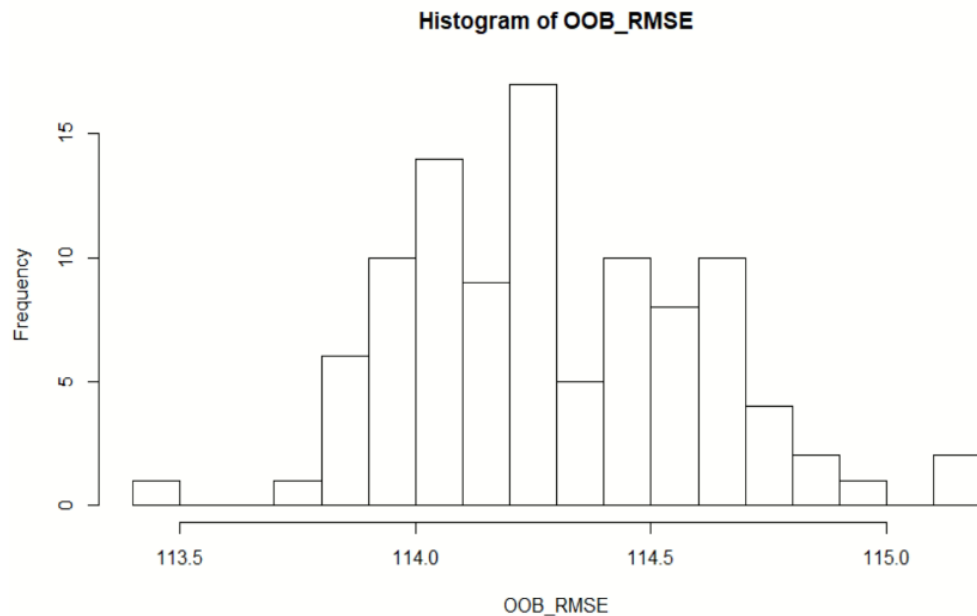
optimal_ranger

Ranger result

Call:
  ranger(formula = fm.esa, data = vs30_train, num.trees = 30000,      mtry =
190, min.node.size = 5, sample.fraction = 0.8, importance = "impurity")

Type:                Regression
Number of trees:     30000
Sample size:         413
Number of independent variables: 625
Mtry:                190
Target node size:    5
Variable importance mode: impurity
Splitrule:           variance
OOB prediction error (MSE): 13037.13
R squared (OOB):     0.5662154

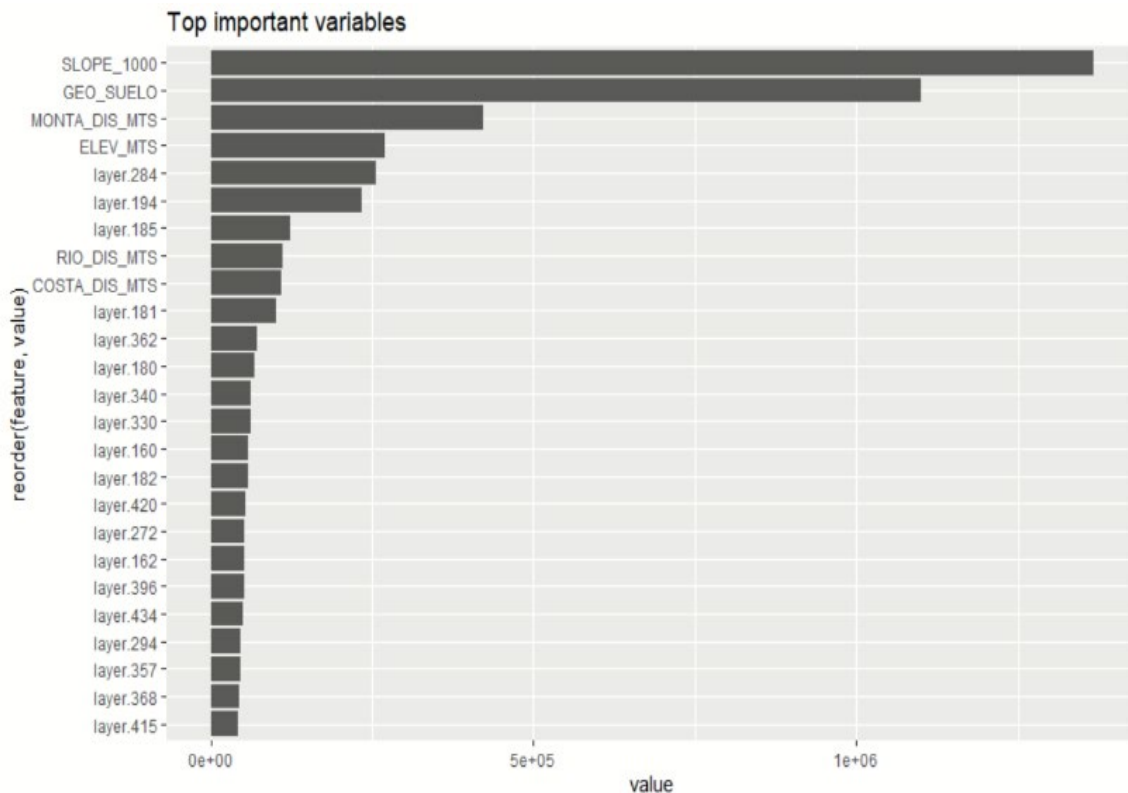
hist(OOB_RMSE, breaks = 20)
```



```
#Inspecciones los resultados de importancia de atributos.

optimal_ranger$variable.importance
list_features=data.frame("feature"=
names(optimal_ranger$variable.importance),"value"=unname(optimal_ranger$variabl
e.importance))

list_features %>%
  dplyr::arrange(desc(value)) %>%
  dplyr::top_n(25) %>%
  ggplot(aes(reorder(feature, value), value)) +
  geom_col() +
  coord_flip() +
  ggtitle("Top important variables")
```



De los resultados, podemos visualizar que la variable más importante, es la pendiente del terreno, seguida de la geología y la distancia a las montañas de la era geológica del Terciario. Esto presenta buena correspondencia con la tesis del Wald del USGS, que las pendientes son una característica diferenciadora para la estimación del valor de las velocidades de ondas de corte; la cual debe estar acompañada con el criterio geológico.

Seguidamente, se procederá a revisar el desempeño del algoritmo de RF considerando la asignación de los pesos. Asimismo, se revisará de nuevo el desempeño considerando similar calidad de los datos (igual peso). Se utilizarán un total de 30 mil árboles para disminuir el error lo más posible. A continuación, se presenta el código y los resultados.

```
#Revision de desempeño de algoritmos

optimal_ranger <- ranger(
  formula      = fm.esa,
  data         = vs30_train,
  num.trees    = 30000,
  mtry         = 190,
  min.node.size = 5,
  sample.fraction = .80,
  importance   = 'impurity') #para evaluar importancia de variables
Growing trees.. Progress: 34%. Estimated remaining time: 1 minute, 0 seconds.
Growing trees.. Progress: 68%. Estimated remaining time: 29 seconds.
optimal_ranger

Ranger result

Call:
  ranger(formula = fm.esa, data = vs30_train, num.trees = 30000,      mtry =
190, min.node.size = 5, sample.fraction = 0.8, importance = "impurity")

Type:                Regression
Number of trees:     30000
Sample size:         413
Number of independent variables: 625
Mtry:                190
Target node size:    5
Variable importance mode:  impurity
Splitrule:           variance
OOB prediction error (MSE): 13094.26
R squared (OOB):     0.5671374

optimal_ranger_pesos <- ranger(
  formula      = fm.esa,
  data         = vs30_train,
  num.trees    = 30000,
  mtry         = 190,
  min.node.size = 5,
  case.weights = ((rm.esa_pesos$Weight.sd)^2),
  sample.fraction = .80,
  importance)   = 'impurity' #para evaluar importancia de variables
Growing trees.. Progress: 34%. Estimated remaining time: 59 seconds.
Growing trees.. Progress: 68%. Estimated remaining time: 28 seconds.
optimal_ranger_pesos

Ranger result

Call:
  ranger(formula = fm.esa, data = vs30_train, num.trees = 30000,      mtry =
190, min.node.size = 5, case.weights = ((rm.esa_pesos$Weight.sd)^2),
sample.fraction = 0.8, importance = "impurity")

Type:                Regression
Number of trees:     30000
Sample size:         413
Number of independent variables: 625
Mtry:                190
Target node size:    5
Variable importance mode:  impurity
Splitrule:           variance
OOB prediction error (MSE): 15311.21
R squared (OOB):     0.512146
```


De los resultados anteriores, se observa un valor del MSE (Predicción de error del OOB de 13094.26. El valor de este número no es tan representativo, dado que hay que entender que un valor de MSE ideal no es 0, ya que entonces se tendría un modelo que predice perfectamente los datos de entrenamiento, pero que es muy poco probable que prediga perfectamente cualquier otro dato. Lo que desea es un equilibrio entre sobreajuste (overfit), es decir un MSE muy bajo para los datos de entrenamiento y subajuste (subfit), es decir, un MSE muy alto para pruebas o datos de test, los cuales son datos no vistos por el modelo.

Por otro lado, a partir de los resultados se puede estimar que el modelo junto a la influencia de todas las covariables explica aproximadamente el 57 % de la variabilidad la variable objetivo “vs30”, para el caso donde no se consideran los pesos. Se obtiene un menor de R-cuadrado cuando se considera la calidad de las técnicas.

El valor obtenido (R-squared OOB), se entiende como la proporción de la varianza total de la variable explicada por la regresión. También se le conoce como coeficiente de determinación (solo regresión), de la variable objetivo; la cual se calcula únicamente con los datos fuera de la bolsa. En otras palabras, este coeficiente de determinación, refleja la bondad del ajuste de un modelo a la variable que pretender explicar.

Es importante saber que el resultado del coeficiente de determinación oscila entre 0 y 1. **Cuanto más cerca de 1 se sitúe su valor, mayor será el ajuste del modelo a la variable que estamos intentando explicar.** De forma inversa, cuanto más cerca de cero, menos ajustado estará el modelo y, por tanto, menos fiable será. Por lo tanto, **se podría decir que nuestro modelo, si bien se explica la variabilidad del vs30 en más de la mitad de los datos fuera de la bolsa, pero no se logra considerar como un modelo confiable.**

Sin embargo, podría ayudar a estimar o predecir valores de vs30 que al final se encuentren dentro de un rango, correspondiente a las clases de tipos de suelo propuesto por la NEHRP, pudiendo incrementar la precisión en la determinación de dichas clases. Este ejercicio se realizará con la otra técnica de *machine learning* (Redes Neuronales)

7.1.1.6.3.1. Predicciones con el modelo final

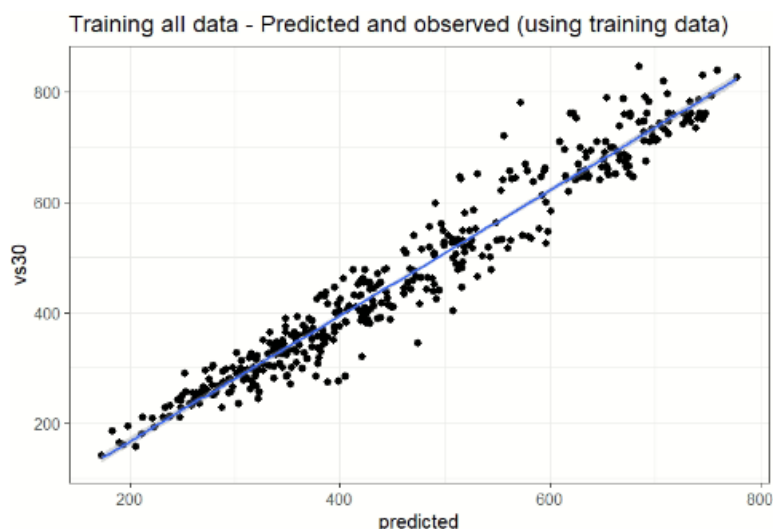
En teoría, cuantos más datos haya visto el modelo implementado, mejor será la generalización. Es por esto, que, si se pensara en entrenar el modelo con el conjunto completo de datos que se tiene disponible, debería generalizarse mejor que un modelo que solo vio, por ejemplo, el 80% del conjunto de datos completos, como nuestro modelo. El problema con esto (y la razón por la que dividimos los datos en subconjuntos de entrenamiento y de prueba (test) en primer lugar), es que se desea poder hacer afirmaciones estadísticas en cuanto a la precisión de los datos no vistos. **Tan pronto, como volvamos a entrenar un modelo con todos los datos, ya no es posible hacer tales afirmaciones.**

La cuestión de si el reentrenamiento con el conjunto de datos completo mejorará el rendimiento en datos futuros no visibles, **no es estrictamente algo que se pueda probar**. Una verificación sensata sería **probar el modelo reentrenado final nuevamente con el conjunto de prueba original**; esperando que obtenga una puntuación más alta que nunca cuando el modelo solo vio el conjunto de entrenamiento, ya que, en esta ocasión, el modelo en realidad ha visto el conjunto de prueba durante el reentrenamiento.

Lo anterior, no indicaría un 100% de confiabilidad de que este nuevo modelo final es superior a todos los otros casos, **pero al menos es tan bueno como puede ser, con los datos proporcionados**. Por lo que se procede a realizar este análisis y entrenar todo el conjunto de datos y ver la mejoría de las predicciones en el conjunto de datos de prueba. A continuación, se muestra el código.

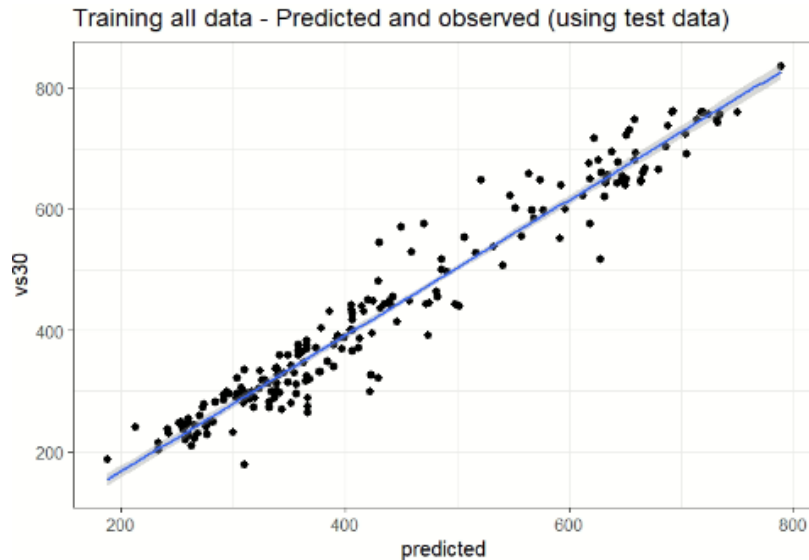
```
#set.seed(101)
# Todos los datos
m_total <- randomForest(formula = fm.esa, data = rm.esal, ntree=30000, mtry =
190)
m_total

#Hacer una prediccion sin realizar un tuneo de los hiperparámetros
pred_randomForest_mtotal <- predict(m_total, vs30_test)
rmse <- function(x) sqrt(sum((x - vs30_test$vs30)^2))
rmse(pred_randomForest_mtotal)
[1] 582.1858
head(pred_randomForest_mtotal)
      2      3      5      7      9     12
406.1230 592.9851 698.1627 363.5112 361.0033 663.5475
vs30_train %>%
  mutate(predicted = predict(m_total, vs30_train)) %>%
  ggplot(aes(predicted, vs30)) +
  geom_point(colour = "#ff6767", alpha = 1) +
  stat_summary(fun.data=mean_cl_normal) +
  geom_smooth(method='lm', formula=y~x, level=0.95) +
  labs(title = "Training all data - Predicted and observed (using training
data)") + theme_bw(15)
```

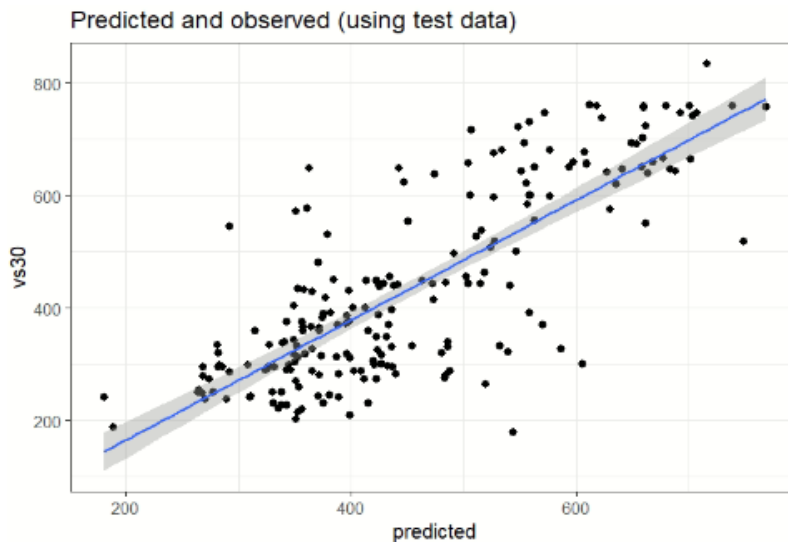


```

vs30_test %>%
  mutate(predicted = predict(m_total, vs30_test)) %>%
  ggplot(aes(predicted, vs30)) +
  geom_point(colour = "#ff6767", alpha = 1) +
  stat_summary(fun.data=mean_cl_normal) +
  geom_smooth(method='lm', formula= y~x, level=0.95) +
  labs(title = "Training all data - Predicted and observed (using test data)")
+ theme_bw(15)
  
```



Comparando con la gráfica que no uso todos los datos:



Al observar las pasadas figuras, es evidente que existe una mejoría en la precisión del modelo al utilizar toda la data disponible. Si bien se mencionó anteriormente que al ocupar toda la data como entrenamiento, no valdría la pena determinar el R-cuadrado, se procederá a calcularlo, tanto para el caso de consideración de los pesos (según la calidad de las técnicas) como para los con similar calidad. A continuación, el detalle del código.

```
#Revisión de desempeño de algoritmos usando la data total como entrenamiento

optimal_ranger_total <- ranger(
  formula      = fm.esa,
  data         = rm.esa1,
  num.trees    = 30000,
  mtry         = 190,
  min.node.size = 5,
  sample.fraction = .80,
  importance   = 'impurity' #para evaluar importancia de variables
)
Growing trees.. Progress: 12%. Estimated remaining time: 3 minutes, 53 seconds.
Growing trees.. Progress: 23%. Estimated remaining time: 3 minutes, 22 seconds.

optimal_ranger_total
Ranger result

Call:
  ranger(formula = fm.esa, data = rm.esa1, num.trees = 30000, mtry = 190,
min.node.size = 5, sample.fraction = 0.8, importance = "impurity")

Type:                Regression
Number of trees:     30000
Sample size:         619
Number of independent variables: 625
Mtry:                190
Target node size:    5
Variable importance mode:  impurity
Splitrule:           variance
OOB prediction error (MSE): 11191.78
R squared (OOB):     0.6259202

optimal_ranger_total_pesos <- ranger(
  formula      = fm.esa,
  data         = rm.esa1,
  num.trees    = 30000,
  mtry         = 190,
  min.node.size = 5,
  case.weights  = ((rm.esa_pesos$Weight.sd)^2),
  sample.fraction = .80,
  importance   = 'impurity' #para evaluar importancia de variables
)
Growing trees.. Progress: 12%. Estimated remaining time: 3 minutes, 53 seconds.
Growing trees.. Progress: 23%. Estimated remaining time: 3 minutes, 22 seconds.

optimal_ranger_total_pesos
Ranger result

Call:
  ranger(formula = fm.esa, data = rm.esa1, num.trees = 30000, mtry = 190,
min.node.size = 5, case.weights = ((rm.esa_pesos$Weight.sd)^2),
sample.fraction = 0.8, importance = "impurity")

Type:                Regression
Number of trees:     30000
Sample size:         619
Number of independent variables: 625
Mtry:                190
Target node size:    5
Variable importance mode:  impurity
Splitrule:           variance
OOB prediction error (MSE): 14738.43
R squared (OOB):     0.5073754
```


Una vez visualizado la mejoría en el modelo utilizando todos los datos entrenados, se ejecutará una predicción general basado en un mallado específico, que concentre a todo El Salvador. Esto se hará utilizando tanto para el modelo donde se considera la calidad en las técnicas, mediante la asignación de pesos, como considerándolas todos los datos con la misma calidad.

Para ello es necesario preparar otro archivo *.csv con dicho mallado, con sus respectivos cálculos de los valores de sus covariables, los cuales son necesarios para poder predecir el valor de Vs30 en la ubicación de cada punto de la grilla. A continuación, se muestra el detalle del código.

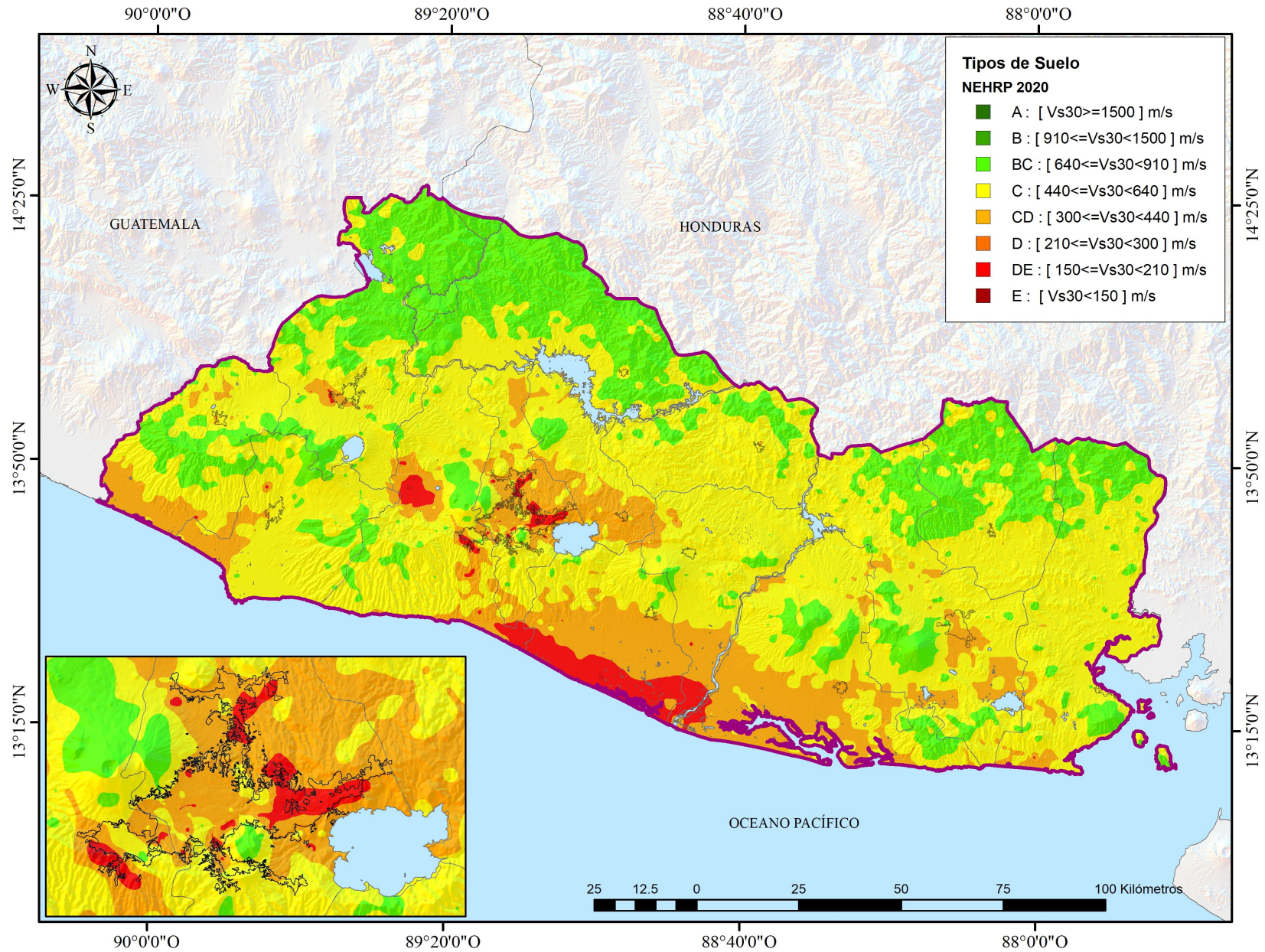
```
#####  
# PREDECIR EN UNA MALLA DE EL SALVADOR  
#####  
datlui2 <- read.csv("Csv/7_covariables/RANDOM_FOREST/PREDECIR_OK_TOTAL_OK.csv",  
header = TRUE)  
saveRDS(datlui2, "Csv/7_covariables/RANDOM_FOREST/MISCELANEA/predecir_oks.rds")  
esalkmcovspredecir_ok2 <-  
readRDS("Csv/7_covariables/RANDOM_FOREST/MISCELANEA/predecir_oks.rds")  
coordinates(esalkmcovspredecir_ok2) <- ~X+Y  
proj4string(esalkmcovspredecir_ok2) <- CRS("+init=epsg:32616")  
esal.rfd <- predict(optimal_ranger_total,  
esalkmcovspredecir_ok2@data)$predictions  
esalkmcovspredecir_ok2$esal_rfd = esal.rfd  
#Guardar en CSV la prediccion  
write.csv(esalkmcovspredecir_ok2,  
"Csv/7_covariables/RANDOM_FOREST/RESULTADOS/RESULT.csv")  
#####
```

Finalmente, se logra exportar los resultados en un archivo *.csv llamado "RESULT.csv", el cual contendrá las predicciones hechas con el modelo para Vs30 determinado anteriormente; con un R-cuadrado de aproximadamente 63% al utilizar toda la data como entrenamiento.

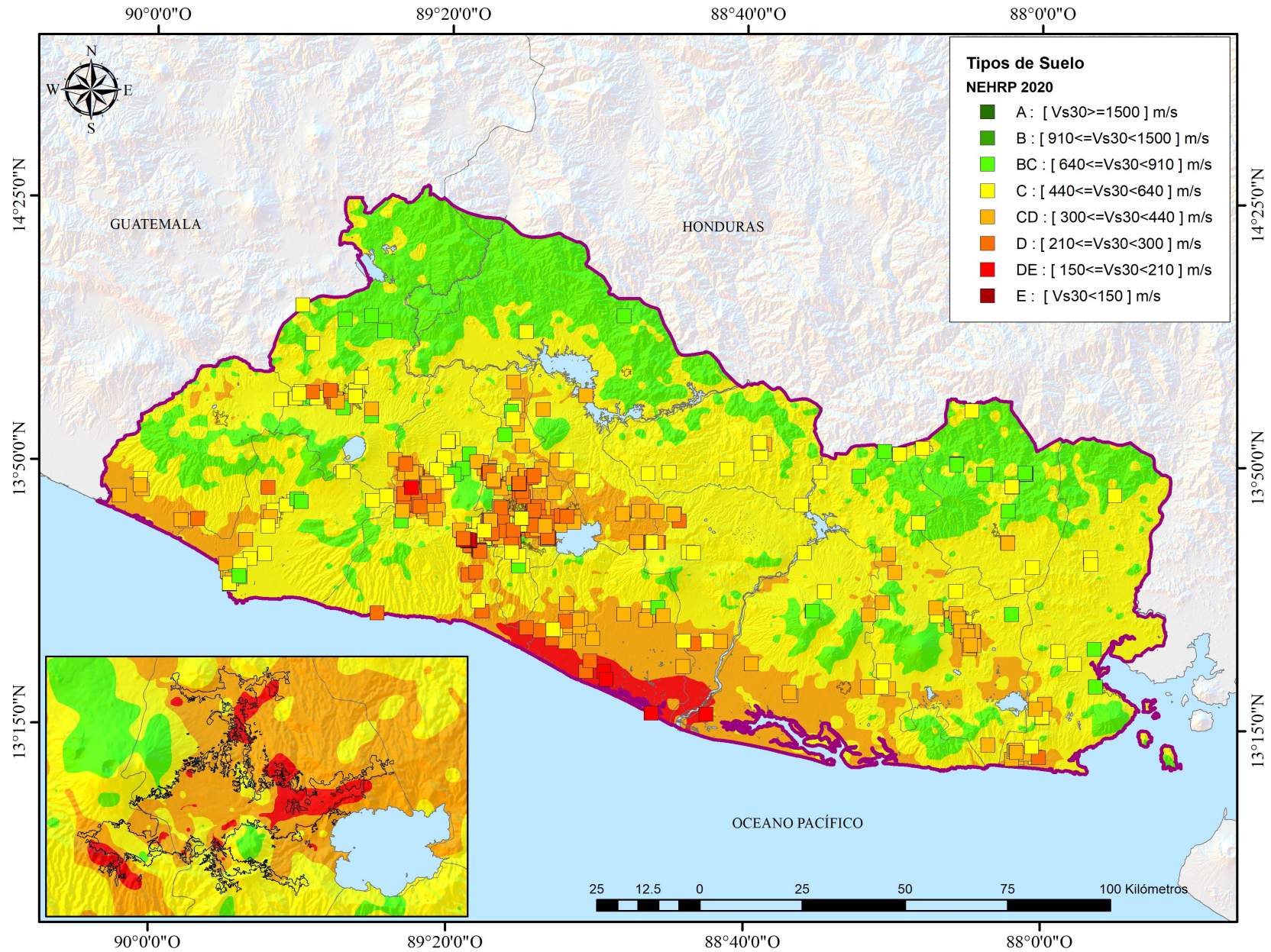
Luego de generar el archivo se procede a exportar los resultados de Vs30 en la malla sobre un Sistema de Información Geográfica y se plotean los puntos. Estos puntos son posteriormente interpolados para crear una superficie. La técnica de interpolación a utilizar será "Vecino Natural" para garantizar lo mayor posible el dato obtenido de los resultados. La resolución de la superficie será de 100 metros, tanto en X como en Y.

En las figuras 63 y 64, se logra apreciar las superficies obtenidas, la primera sin la ubicación de los 619 puntos y la segunda superponiendo dichos puntos.

Generación de mapa de velocidades promedio a 30 metros de profundidad (Vs30) de El Salvador



Generación de mapa de velocidades promedio a 30 metros de profundidad (Vs30) de El Salvador



7.1.2. Técnica de redes de neuronas artificiales (RNA)

7.1.2.1. RNA problema como variable categórica

Para este modelo nuevamente se han tomado en consideración las variables de entrada: latitud, longitud, pendiente, geología, elevación, distancia al río más cercano, distancia a la línea de costa y distancia a las montañas. La diferencia radica en que la variable de salida es categórica, es decir, el modelaje se ha dividido por tipo de suelo. Generando un conjunto de modelos por cada clase NEHRP 2020.

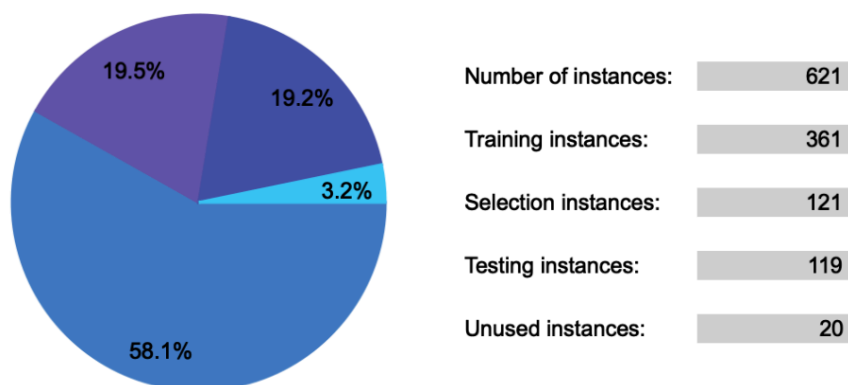
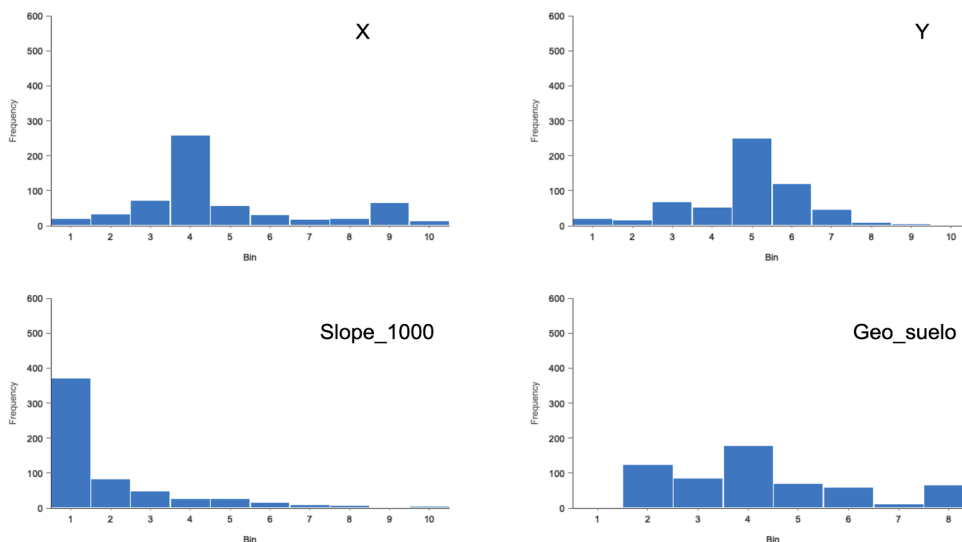


Figura 63. Distribución del conjunto de datos. Fuente: MARN. 2020

El conjunto de datos de entrada se distribuye como muestra la Figura 63, siendo conformado por un total de 619 registros cada uno con 8 variables de entrada y una variable de salida, siendo esta última la probabilidad de que las características particulares de un suelo pertenezcan o no a una clase NEHRP en específico. Cada variable de entrada fue depurada para evitar valores atípicos que puedan dificultar la clasificación, como puede apreciarse en la Figura 64.



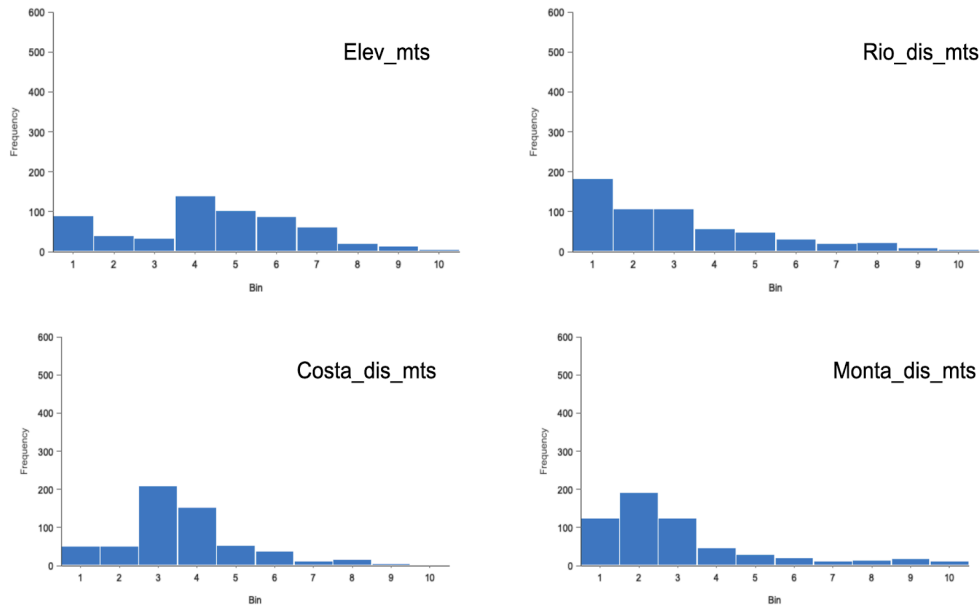
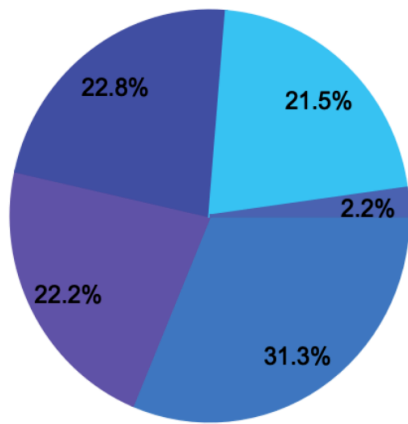


Figura 64. Distribución de datos de cada variable de entrada. Fuente: MARN. 2020

La distribución de la variable de salida se concentra en los tipos de suelo: BC, C, CD y D. Aunque existen registros del tipo de suelo DE (solamente 13 para todo el país), se decidió no incluirlo en el modelado ya que su frecuencia es muy inferior comparada con las demás clasificaciones. Ver figura 65.



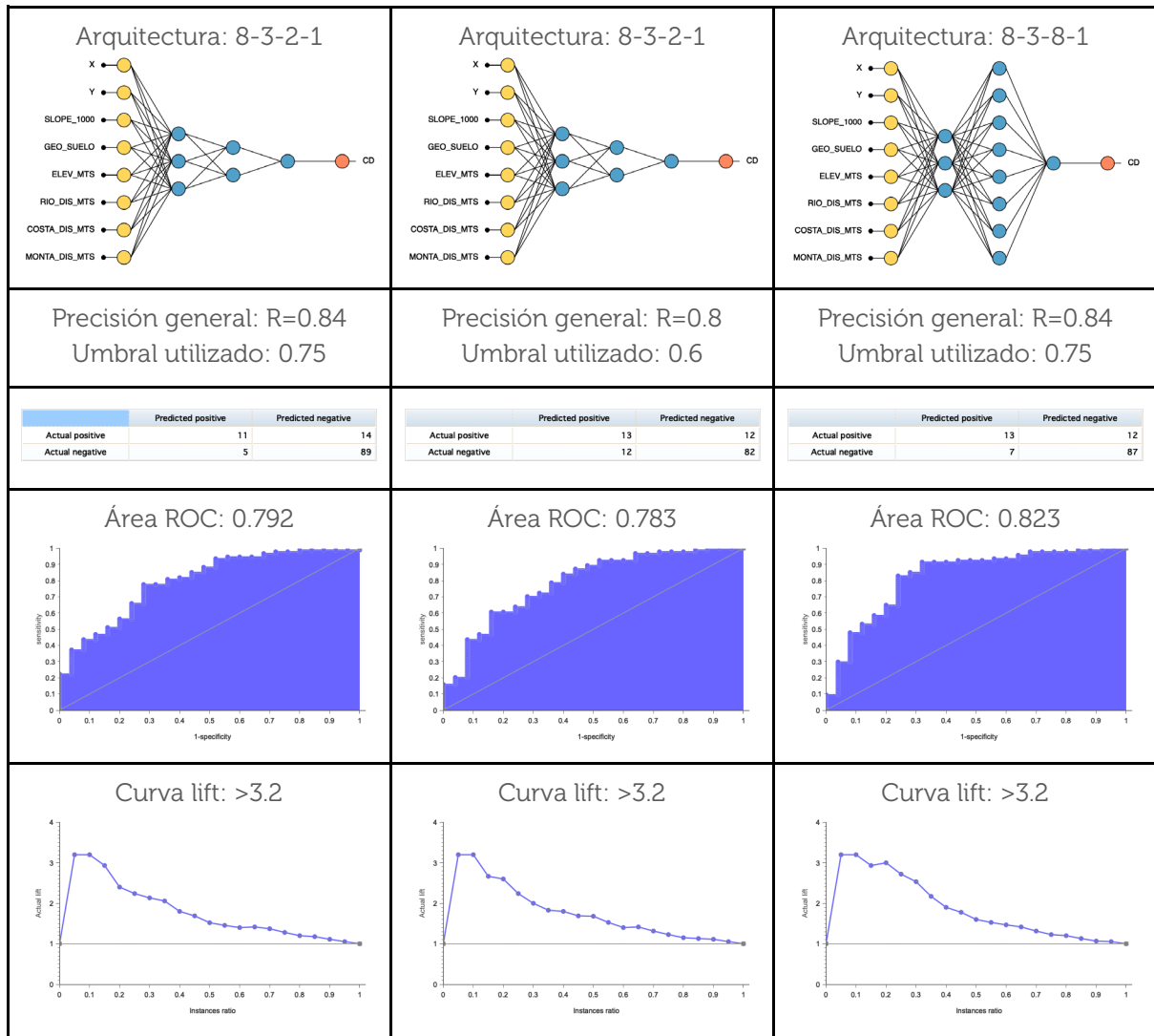
Tipo de suelo	Instancias
CD	188
C	133
BC	137
D	129
DE	13

Figura 65. Distribución de número de instancias por tipo de suelo. Fuente: MARN. 2020

7.1.2.1.1. Modelo 1: Tipo de suelo CD

En la siguiente tabla se resumen los aspectos más importantes de las tres mejores redes de neuronas artificiales que se calcularon para este tipo de suelo en particular.

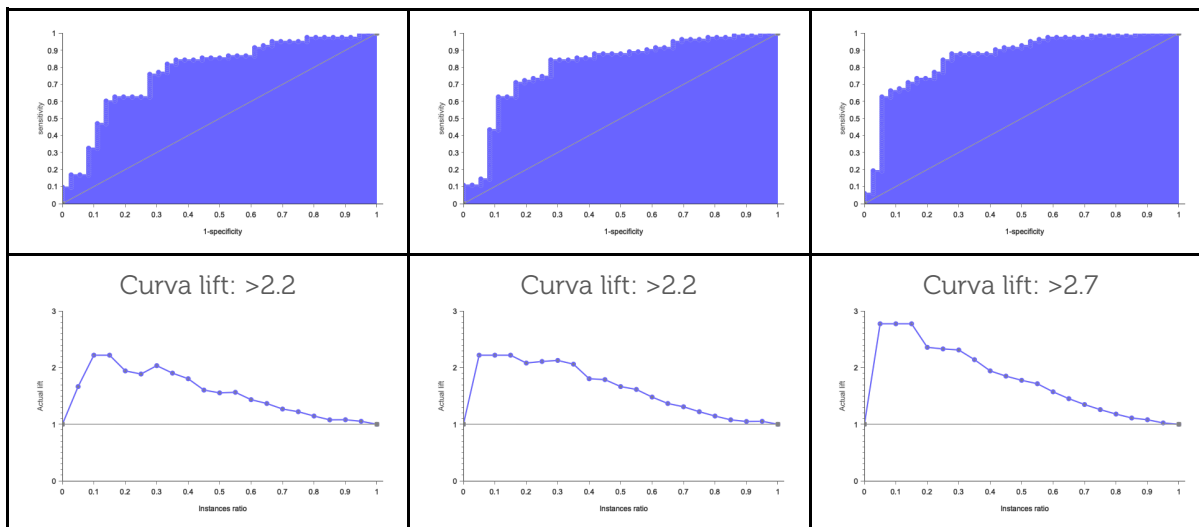
Archivo: CD_075_084.py	Archivo: CD_06_08.py	Archivo: CD_075_084_2.py
------------------------	----------------------	--------------------------



7.1.2.1.2. Modelo 2: Tipo de suelo BC

En la siguiente tabla se resumen los aspectos más importantes de las tres mejores redes de neuronas artificiales que se calcularon para este tipo de suelo en particular.

<p>Archivo: BC_065_078.py</p>	<p>Archivo: BC_06_08.py</p>	<p>Archivo: BC_06_081.py</p>																											
<p>Arquitectura: 8-3-1-1</p>	<p>Arquitectura: 8-3-1-1</p>	<p>Arquitectura: 8-3-5-1</p>																											
<p>Precisión general: R=0.78 Umbral utilizado: 0.65</p>	<p>Precisión general: R=0.8 Umbral utilizado: 0.6</p>	<p>Precisión general: R=0.81 Umbral utilizado: 0.6</p>																											
<table border="1"> <thead> <tr> <th></th> <th>Predicted positive</th> <th>Predicted negative</th> </tr> </thead> <tbody> <tr> <td>Actual positive</td> <td>23</td> <td>13</td> </tr> <tr> <td>Actual negative</td> <td>13</td> <td>70</td> </tr> </tbody> </table>		Predicted positive	Predicted negative	Actual positive	23	13	Actual negative	13	70	<table border="1"> <thead> <tr> <th></th> <th>Predicted positive</th> <th>Predicted negative</th> </tr> </thead> <tbody> <tr> <td>Actual positive</td> <td>26</td> <td>10</td> </tr> <tr> <td>Actual negative</td> <td>14</td> <td>69</td> </tr> </tbody> </table>		Predicted positive	Predicted negative	Actual positive	26	10	Actual negative	14	69	<table border="1"> <thead> <tr> <th></th> <th>Predicted positive</th> <th>Predicted negative</th> </tr> </thead> <tbody> <tr> <td>Actual positive</td> <td>20</td> <td>16</td> </tr> <tr> <td>Actual negative</td> <td>7</td> <td>76</td> </tr> </tbody> </table>		Predicted positive	Predicted negative	Actual positive	20	16	Actual negative	7	76
	Predicted positive	Predicted negative																											
Actual positive	23	13																											
Actual negative	13	70																											
	Predicted positive	Predicted negative																											
Actual positive	26	10																											
Actual negative	14	69																											
	Predicted positive	Predicted negative																											
Actual positive	20	16																											
Actual negative	7	76																											
<p>Área ROC: 0.775</p>	<p>Área ROC: 0.805</p>	<p>Área ROC: 0.856</p>																											



7.1.2.1.3. Modelo 3: Tipo de suelo D

En la siguiente tabla se resumen los aspectos más importantes de las tres mejores redes de neuronas artificiales que se calcularon para este tipo de suelo en particular.

Archivo: D_07_079.py	Archivo: D_07_081.py	Archivo: D_06_082.py																											
Arquitectura: 8-7-1	Arquitectura: 8-7-1	Arquitectura: 8-7-1-1																											
Precisión general: R=0.79 Umbral utilizado: 0.7	Precisión general: R=0.81 Umbral utilizado: 0.7	Precisión general: R=0.82 Umbral utilizado: 0.6																											
<table border="1"> <thead> <tr> <th></th> <th>Predicted positive</th> <th>Predicted negative</th> </tr> </thead> <tbody> <tr> <td>Actual positive</td> <td>12</td> <td>12</td> </tr> <tr> <td>Actual negative</td> <td>13</td> <td>82</td> </tr> </tbody> </table>		Predicted positive	Predicted negative	Actual positive	12	12	Actual negative	13	82	<table border="1"> <thead> <tr> <th></th> <th>Predicted positive</th> <th>Predicted negative</th> </tr> </thead> <tbody> <tr> <td>Actual positive</td> <td>15</td> <td>9</td> </tr> <tr> <td>Actual negative</td> <td>14</td> <td>81</td> </tr> </tbody> </table>		Predicted positive	Predicted negative	Actual positive	15	9	Actual negative	14	81	<table border="1"> <thead> <tr> <th></th> <th>Predicted positive</th> <th>Predicted negative</th> </tr> </thead> <tbody> <tr> <td>Actual positive</td> <td>18</td> <td>6</td> </tr> <tr> <td>Actual negative</td> <td>16</td> <td>79</td> </tr> </tbody> </table>		Predicted positive	Predicted negative	Actual positive	18	6	Actual negative	16	79
	Predicted positive	Predicted negative																											
Actual positive	12	12																											
Actual negative	13	82																											
	Predicted positive	Predicted negative																											
Actual positive	15	9																											
Actual negative	14	81																											
	Predicted positive	Predicted negative																											
Actual positive	18	6																											
Actual negative	16	79																											
Área ROC: 0.782	Área ROC: 0.804	Área ROC: 0.818																											
Curva lift: >2.3	Curva lift: >2.5	Curva lift: >2.5																											

7.1.2.1.4. Modelo 4: Tipo de suelo C

Este tipo de suelo en específico presentó particular dificultad en su modelaje, tal como puede evidenciarse en un modelo que se realizó en un intento por englobar las cuatro clasificaciones en una sola red de neuronas artificiales, su estructura y correspondiente matriz de confusión fueron las siguientes:

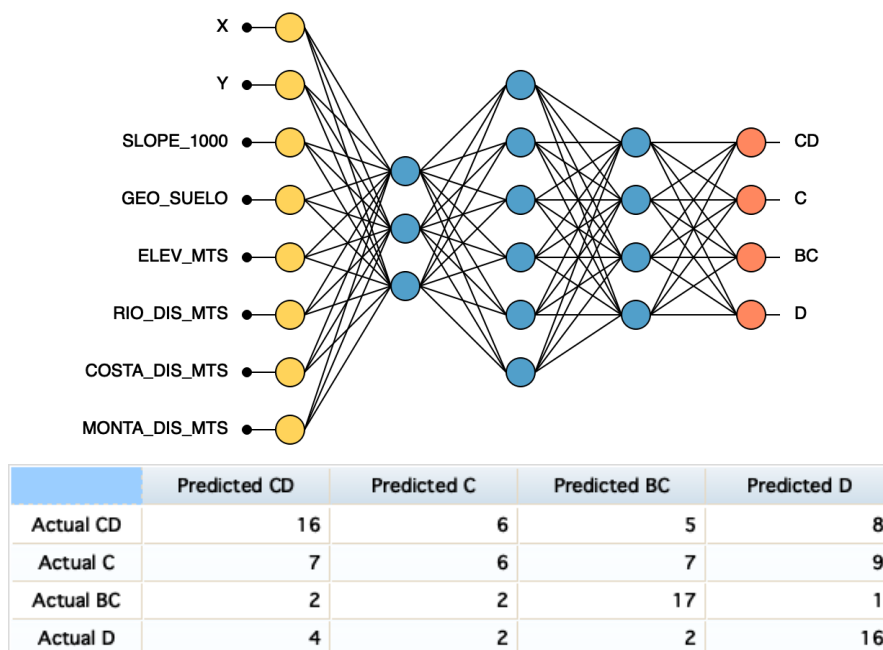


Figura 66. Red de Neuronas Artificiales para todos los tipos de suelo analizados. Fuente: MARN. 2020

Se evidencia en la figura 66 que, aunque los valores en la diagonal principal son favorables para los tipos de suelo CD, BC y D; no lo fueron para C. Además, se intentó generar algún modelo que identificara exclusivamente los suelos tipo C, pero sin mayor fruto, pues el mejor valor de área bajo la curva que se obtuvo fue de 0.603.

Por tanto, se decidió considerar alguna alternativa distinta, clasificar este tipo de suelo en particular por medio de la “eliminación” o el “descarte” es decir que al no indicar otro modelo claros indicios de que la probabilidad es alta, entonces considerar que se es de tipo C; el argumento anterior se respalda al examinar la distribución geográfica de este tipo de suelo, ya que puede verse que se encuentra “disperso”, en otras palabras podría considerarse como una especie de tipo de suelo por defecto.

7.1.2.1.5. Algoritmo para selección de umbral

Especial mención merece la forma en que se seleccionaron los umbrales de los submodelos de los apartados anteriores. Se graficaron las variaciones entre la exactitud, la sensibilidad, la especificidad y la precisión tanto positiva como negativa; para los umbrales: 0.4, 0.5, 0.6, 0.7, 0.75 y 0.8 (ver figura 67). El objetivo era encontrar el umbral

que maximice las cinco variables mencionadas, con énfasis en la sensibilidad y la precisión positiva (probabilidad de que sí pertenezca al tipo de suelo en cuestión). Además, se incluyó el umbral sugerido por la curva ROC, el cual a veces aportó muy buenos resultados, pero otras veces fue superado por otros encontrados previamente.

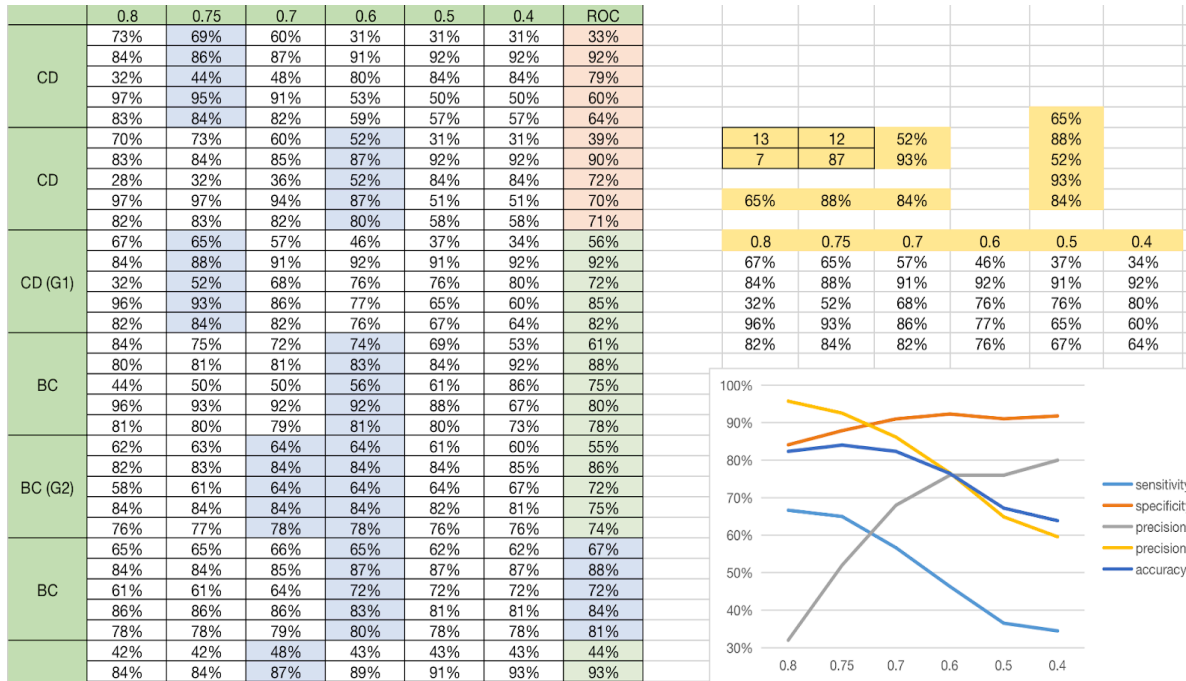


Figura 67. Resultados de RNA para encontrar umbral que maximice las variables. Fuente: MARN. 2020

7.1.2.1.6. Ensamble de sub-modelos de RNA

Para finalizar este modelo de RNA, se utilizan los umbrales obtenidos anteriormente para cada tipo de suelo. Para ello, se implementó un código *Jupyter Notebook*, donde se utilizó un método de ensamblado para encontrar la combinación que arroje los mejores resultados. Se terminó utilizando las siguientes reglas basadas en un árbol de decisión que se generó utilizando una librería del paquete scikit-learn disponible para Python 3.

```

for resultado in listaResultados:
    prediccion = "C"
    BC1, BC2, CD1, CD2, D1, D2 = resultado[:-1]
    max_value = max(resultado[:-1])
    max_index = resultado[:-1].index(max_value)

    if(D1>0.7 or D2>0.6): prediccion = 'D'
    elif(max_index==2 or max_index==3): prediccion = 'CD'
    elif(BC1>0.6 or BC2>0.6): prediccion = 'BC'

    if(BC1+BC2+CD1+CD2+D1+D2<1): prediccion = 'C'

    if(prediccion == 'D' and (D1<0.5 or D2<0.5)): prediccion = 'CD'
    if(prediccion == 'CD' and CD2<0.575): prediccion = 'C'

```

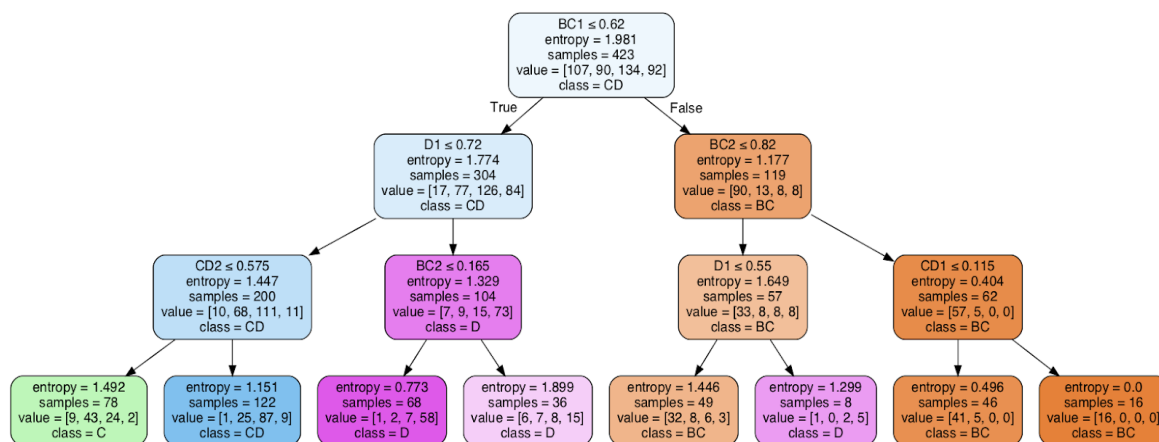


Figura 68. Método de ensamblaje de los sub-modelos que conforman el modelo RNA. Fuente: MARN 2020

Utilizando las reglas que se detallaron, se logró conseguir un valor de exactitud general de 68.23% (se logró 69.9% después de ciertos cambios adicionales) incluyendo todos los tipos de suelo, para todo el territorio nacional. La matriz de confusión es la siguiente:

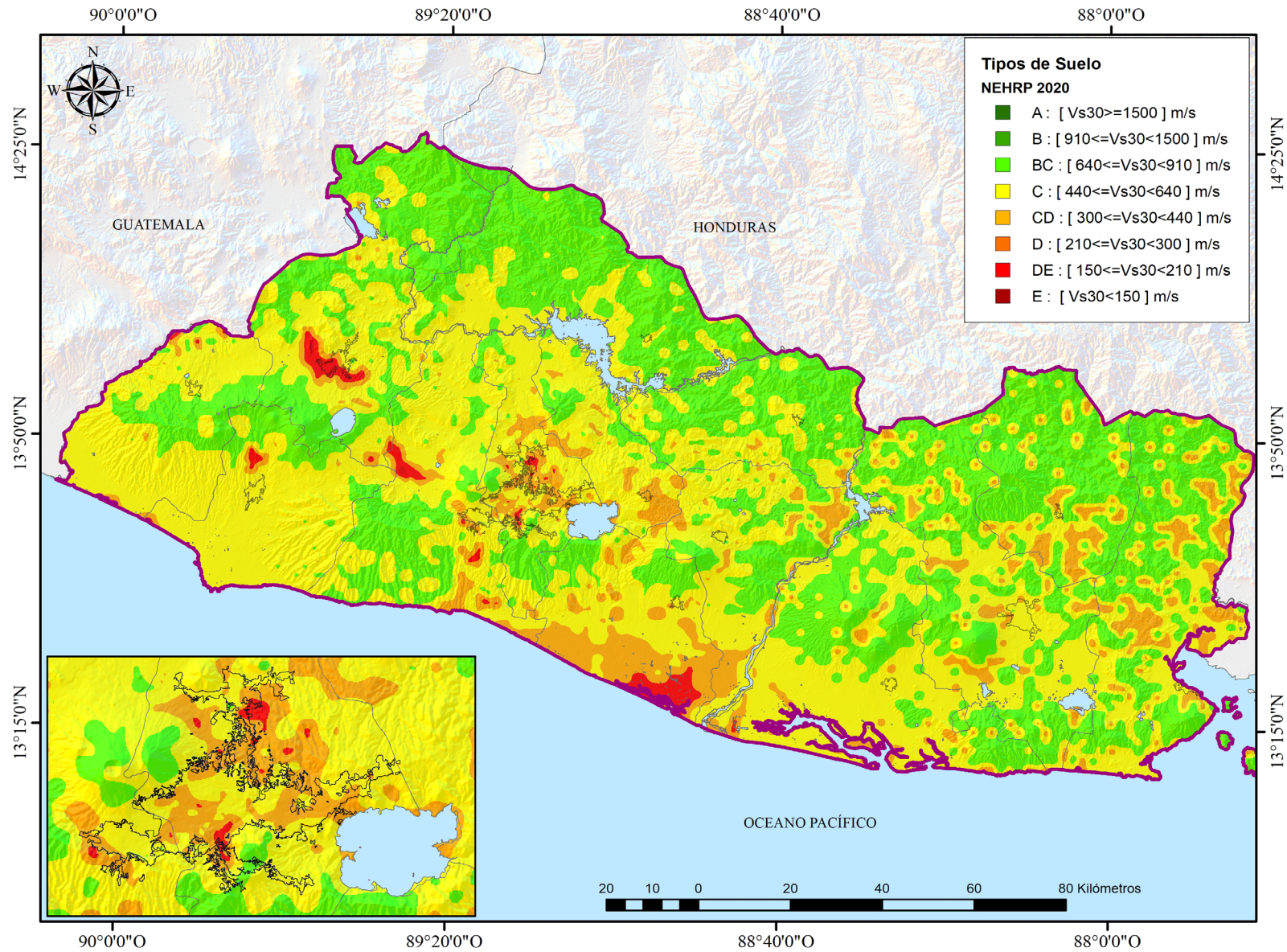
	Pred BC	Pred C	Pred CD	Pred D	Precisión
Real BC	121	20	1	7	81%
Real C	27	66	34	8	49%
Real CD	9	35	129	18	68%
Real D	3	7	13	107	82%
Sensibilidad	76%	52%	73%	76%	69.9%

Figura 69. Matriz de confusión obtenida para los tipos de suelo analizados. Fuente: MARN. 2020

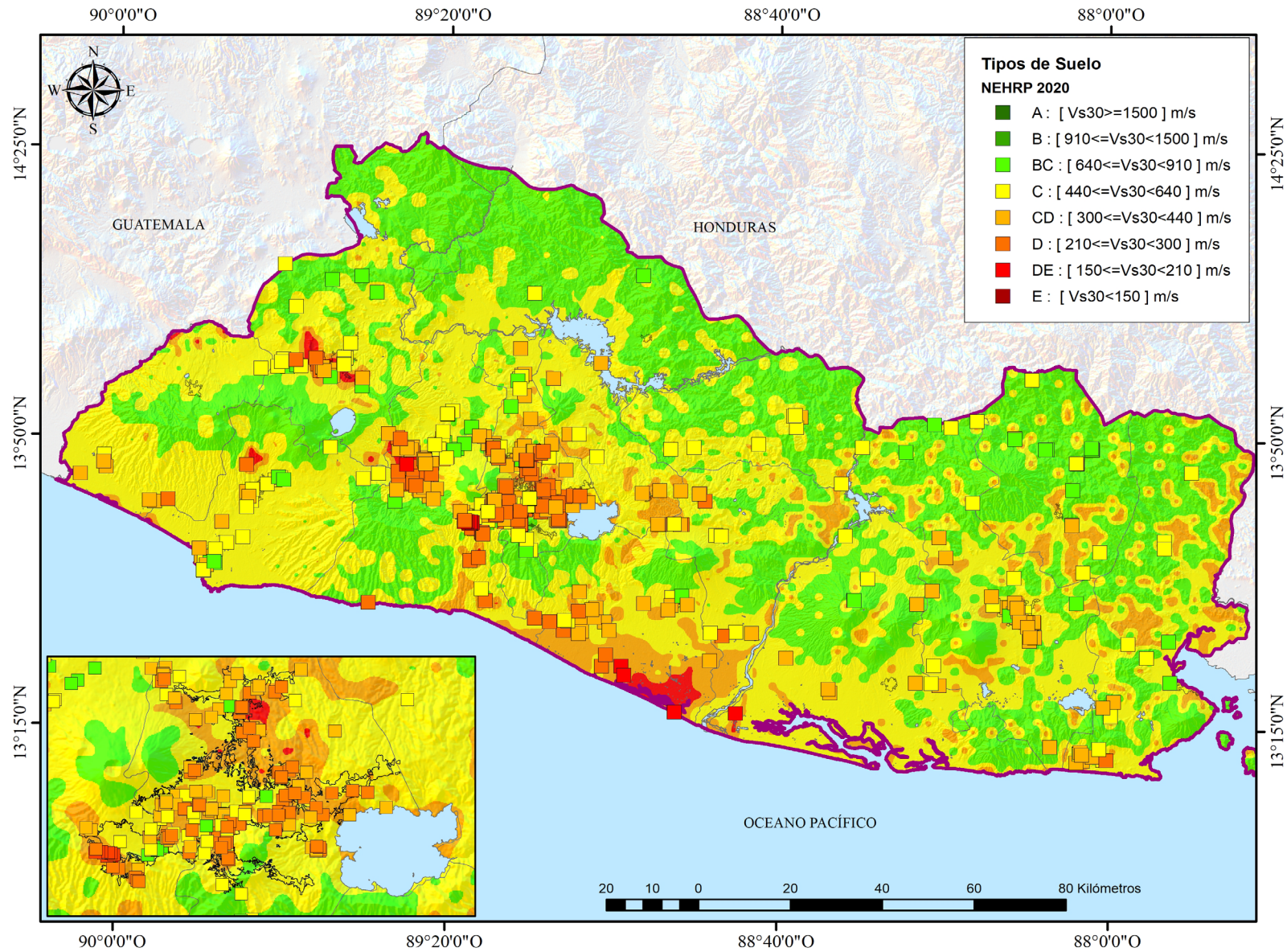
Los valores de la diagonal principal son bastante prometedores para un modelo que unifica los cuatro tipos de suelo, rozando el 70%. Esto significa siete puntos porcentuales de mejor rendimiento que el modelo correspondiente a *RandomForest Sp*. Como comentario final, puede apreciarse que el modelaje de la clase C siempre obtiene los resultados más bajos, mientras que las clases BC y D ofrecen valores considerablemente más altos. Se espera que los puntos fuertes de cada modelo se fusionen para formar un mapa que estime de la mejor manera posible los tipos de suelo según la NEHRP.

En resumen, para este modelo de RNA, primero se evalúan los resultados de las 6 mejores RNA's encontradas (2 para cada tipo de suelo: BC1, BC2, CD1, CD2, D1 Y D2, excepto C que se tomó por defecto), luego se efectúa un árbol de decisión (mostrado en la figura 68), para finalmente devolver la clase predicha junto con el centroide de dicha clase en la variable correspondiente al vs30. Luego, se procede a exportar los resultados obtenidos y se genera una superficie de igual manera que se hizo con la técnica de RFsp. En las figuras 70 y 71, se muestran las superficies obtenidas, la primera sin la ubicación de los 619 puntos y la segunda superponiendo dichos puntos.

Generación de mapa de velocidades promedio a 30 metros de profundidad (Vs30) de El Salvador



Generación de mapa de velocidades promedio a 30 metros de profundidad (Vs30) de El Salvador



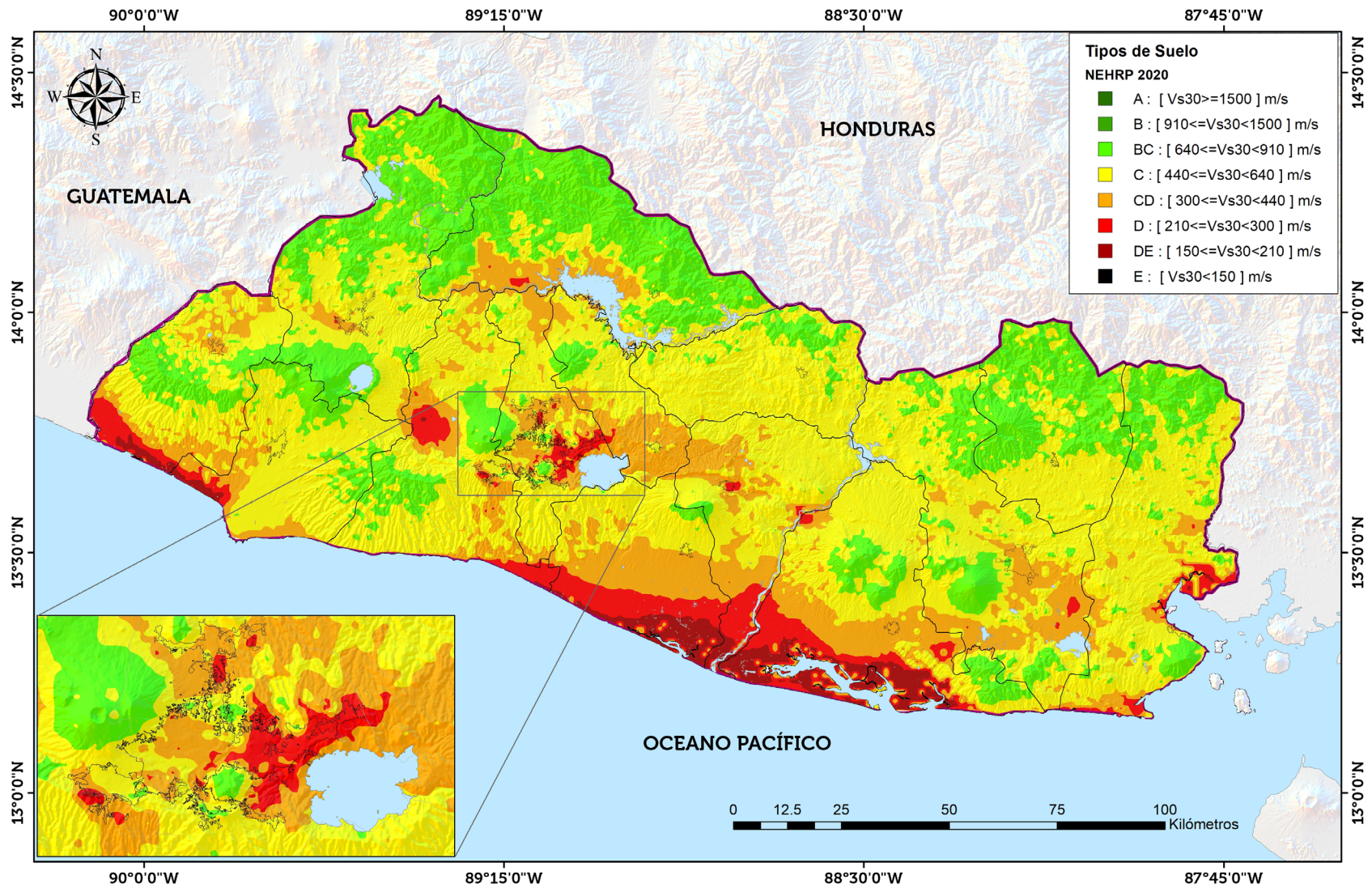
7.1.2.1.7. Ensamblaje de resultados de Modelos (RandomForest SP & RNA)

Una vez generado los mapas de las dos técnicas consideradas, se procede a ejecutar un álgebra de mapas para generar el ensamblaje de los dos modelos. Tal como se explicó antes, el modelo RNA con resultados discretos (clases de tipo de suelo) fue traducido a variables continuas; para luego combinarlos con la asignación de pesos iguales. En otras palabras, se obtuvo el promedio de los resultados, para considera la contribución de ambos modelos.

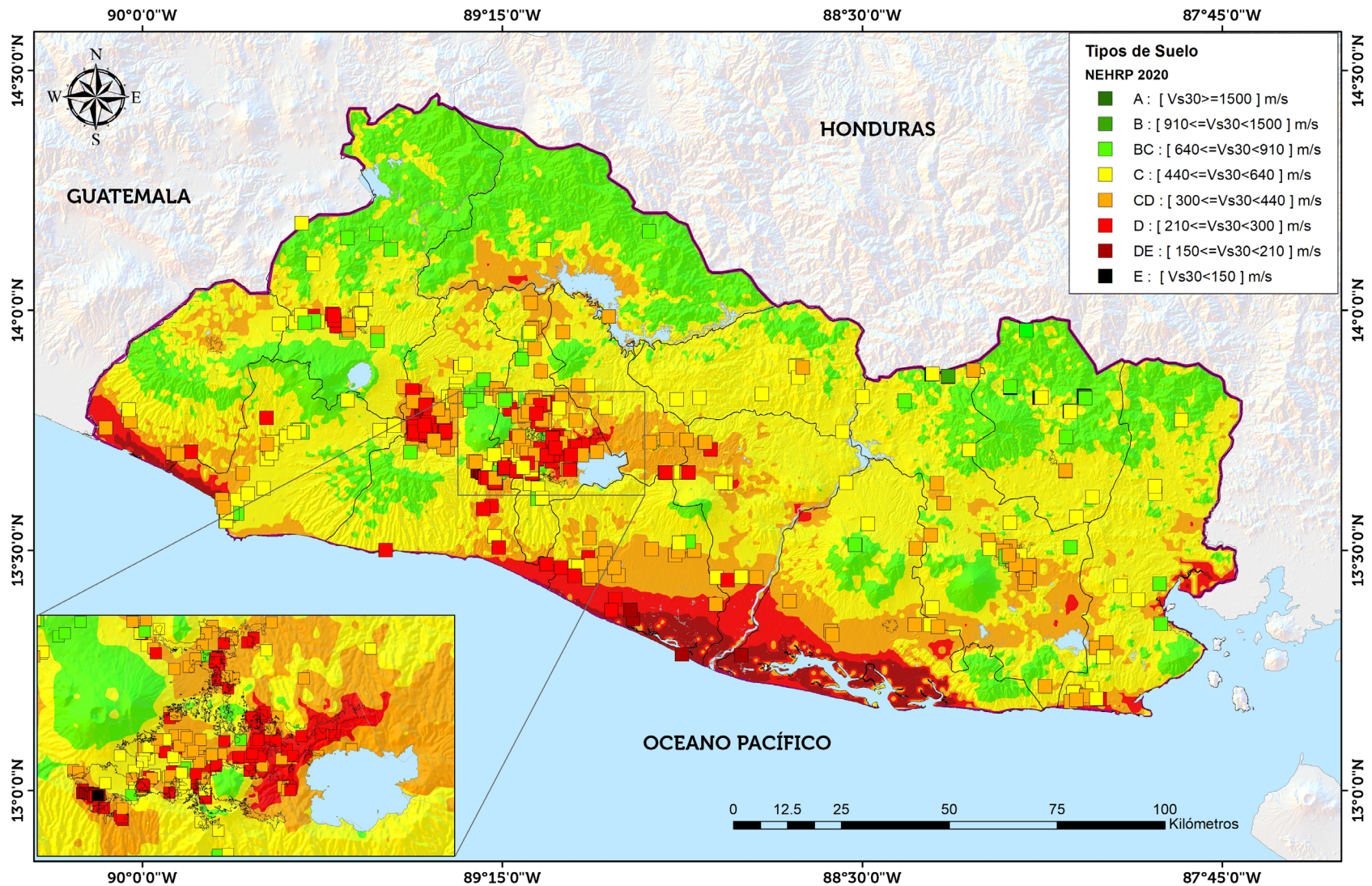
Finalmente, se obtiene el promedio de los datos y se procede a realizar la interpolación con la técnica de “Vecino Natural” para crear la superficie final. La resolución de la superficie se mantendrá en 100 metros, tanto en X como en Y.

En las figuras 72 y 73, se logra apreciar las superficies obtenidas, la primera sin la ubicación de los 619 puntos y la segunda superponiendo dichos puntos. Este mapa correspondería al mapa integrado de Vs30 para El Salvador.

Generación de mapa de velocidades promedio a 30 metros de profundidad (Vs30) de El Salvador



Generación de mapa de velocidades promedio a 30 metros de profundidad (Vs30) de El Salvador



8. Conclusiones, recomendaciones y comentarios finales

- Para la conformación del mapa de zonas de respuesta sísmica homogénea de El Salvador (en términos de Vs30), se recopilieron un total 519 puntos a nivel nacional, con información de la estructura de velocidad de ondas de corte en profundidad, determinada tanto de forma directa como inferida geológicamente. Aproximadamente un 44% se concentra principalmente en el Área Metropolitana de San Salvador (AMSS).
- Debido a la aún limitada cantidad de información en otros lugares del país, se optó por agregar algunos valores de VS30 calculados a partir de las pendientes del terreno, correspondiente al producto del USGS, específicamente para valores superiores e iguales a 640 m/s; siendo cuidadoso en los lugares, como los grábenes, donde se sabe que no existen afloramientos rocosos.
- Se utilizaron las técnicas de aprendizaje supervisado (*machine learning*), basadas en los algoritmos de *Random Forest* con Predicción espacial (Rfsp) y la Técnica de redes de neuronas artificiales; para la generación de un modelo estadístico espacial que represente una óptima interpolación de valores de Vs30 a nivel nacional.
- La variabilidad litológica de El Salvador se traduce en una amplia gama de comportamientos geotécnicos que van desde rocas muy duras y competentes que no amplificarán la señal sísmica (o muy poco), hasta terrenos muy blandos que pueden amplificarla de forma relevante.
- El territorio nacional ha sido dividido en cinco tipos de suelo de acuerdo a la clasificación propuesta por NEHRP en el año 2020, cuyas divisiones se encuentran en función de la variación de la velocidad de ondas de corte de los primeros 30 superiores; y a las cuales se les ha integrado las diferentes unidades geológicas encontradas en los pozos perforados, desde un punto de vista de edades geológicas, orígenes eruptivos, descripciones físicas y geomecánicas. Los cinco tipos de suelo son: BC (Vs30: 640-910 m/s), C (Vs30: 440-640 m/s), CD (Vs30: 300-440 m/s), D (Vs30: 210-300 m/s) y DE (Vs30: 150-210 m/s).
- La práctica de asignación de velocidades de ondas de corte a partir de perfiles geológicos, se considera muy subjetiva y a su vez con cierto grado de sesgo, por parte de la opinión técnica del profesional de geología; por lo que debe de tomarse como preliminar para su uso en la caracterización de la respuesta de un sitio en específico. Sin embargo, dado que la información era abundante y abarcaba una buena distribución espacial a nivel nacional, se trató de aprovechar en lo mayor posible la información; y considerarla con un menor peso al momento de efectuar las interpolaciones.

- Cada cambio (o escalón) en el perfil de velocidades se le denomina estrato sísmico, el cual proviene de un promedio de velocidades donde se estima que no ha existido un cambio considerable en la pendiente de los tiempos de arribo de las ondas secundarias respecto a la profundidad. Por lo que el analista del ensayo, estima que podría calcularse un promedio de velocidades dentro de un estrato sísmico. Esto indica que no precisamente exista un estrato sísmico correspondiente a un estrato o unidad geológica; lo cual sería el deber ser; sin embargo, se podría inferir que, si bien podría existir cambios, pero no son tan abruptos
- Con el objetivo de poder utilizar la información suministrada en la descripción litológica de cada uno de los perfiles estratigráficos, así como extraer muestras para realizar ensayos geotécnicos, se realizó un proceso de reclasificación sobre la base de criterios asociados a los parámetros que más influyen en las propiedades dinámicas del subsuelo.
- La reclasificación de las unidades geológicas se basó en los siguientes criterios:
 - a) Perfil de espesores de velocidad de corte (Vs) obtenidos por ensayos Down-hole (en primera instancia, y las condiciones lo permitan) y ensayos de Microtremores.
 - b) Edad y formaciones (miembros) geológicas de acuerdo al mapa geológico superficial 1:100,000, desarrollado por la Misión Geológica Alemana entre los años 1967-1971.
 - c) Clasificaciones realizadas por cronología eruptiva; en profundidad.
 - d) Descripciones físicas de las propiedades de los materiales que componen las unidades litológicas, tales como: Dureza, Densidad y Permeabilidad.
 - e) Descripción visual de la granulometría de cada uno de los estratos identificados.
- Para llevar a cabo el proceso de generación de un modelo estadístico espacial de los valores de Vs30, a partir de la aplicación de técnicas de aprendizaje, se tuvo que desarrollar regresiones numéricas en función de variables exploratorias (conocidas como covariables) correspondientes a unidades geomorfológicas y características geográficas para cada localización donde existía un valor de Vs30.
- Las covariables elegidas presentan un cierto grado de influencia en la estimación del efecto de sitio, sin embargo, no se consideran tan representativas en comparación con el grado que existe con la estructura de velocidades de ondas de corte.
- Las covariables seleccionadas fueron las siguientes:
 - a) **elevación del terreno (en metros)**, donde se asumen que tiene relación con los cambios de granulometría de los sedimentos en los ríos, los cuales cambian de grano grueso en el río arriba a granos finos en el río abajo.

- b) **La distancia más cercana a ríos principales (en metros);** asociada a la geomorfología formada por inundaciones, donde los depósitos de sedimentos (de consistencia blanda) cerca de un río se vuelven más gruesos y más delgados a medida que aumenta la distancia al río.
- c) **El ángulo de la pendiente del terreno (en grados y multiplicada por 1000);** donde se considera que pendientes elevadas indicarían estratos firmes o rocosos (formados por avalanchas de escombros) y valles planos indicarían suelo o sedimentos, y la transición a pendientes intermedias.
- d) **La distancia más cercana a línea de costa (en metros);** relacionada al espesor de las capas sedimentarias en los deltas y tierras bajas costeras.
- e) **La distancia más cercana a una montaña de la era del Terciario (en metros);** relacionada también al espesor de las capas sedimentarias que influyen en el Vs30 de las tierras bajas en valles y terrazas a lo largo de los ríos en regiones montañosas. De igual manera, se asocia a los efectos cuenca relacionadas a la generación de ondas superficiales en sus bordes y el consecuente efecto de sitio. Asimismo, esta característica presenta una correlación con la profundidad del lecho rocoso debajo de la capa sedimentaria de la unidad geomorfológica formada durante el Cuaternario. Por lo que, una montaña o cerro formado durante el período más antiguo (Pre-Terciario o Terciario) es la referencia que se seleccionará para medir las distancias más cercanas a las observaciones de la variable objetivo.
- f) **La geología superficial a partir del mapa geológico** de la Misión Geológica Alemana (Bosse et al. 1978), la cual es extraída de manera cualitativa, a partir de una reclasificación hecha por García, I.; 2006; y por lo tanto tratada como una variable categórica convertida en una simple clasificación numérica.

9. Referencias

- American Society of Civil Engineers (ASCE 7-16), 2016, “Minimum Design Load for Building and Other Structures,” ASCE/SEI 7-16.10, Reston, Virginia.
- Bosse, H.R., Lorenz, W., Merino, A., Mihm, A., Rode, K., Schmidt-Thome, M., Wiesemann, G. y Weber, H.S.1978. Geological Map of El Salvador Republic 1:100,000. Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, Germany
- BSSC (2020) NEHRP Recommended Provisions for seismic Regulations for New buildings and other Structures, Part1: Provisions, FEMA 368, Federal Emergency Management Agency, Washington, D.C.
- CCARA Project, 2017, “Evaluación de Amenaza y Riesgo Sísmico en El Caribe y América Central” (CCARA por sus siglas en inglés), a través del Global Earthquake Model (GEM).
- DTO-AMENAZA2019: Documento generado en el MARN en el año 2019, llamado: “Reevaluación probabilística de amenaza sísmica de El Salvador”, desarrollado por Luis Mixco.
- DTO-EMPIRICOS2019: Documento generado en el MARN en el año 2019, llamado: “Evaluación de efectos de sitio en el Salvador a través de métodos empíricos”, desarrollado por Luis Mixco.
- Eurocode 8 (EN 1998-1): Design of structures for earthquake resistance.
- Faccioli, E., Battistella, C., Alemani, P. y Tibaldi, a. (1988). Seismic microzoning investigations in the metropolitan area of San Salvador, El Salvador, following the destructive earthquake of October 10, 1986.
- Ferres López, María Dolors. (2014). “Estratigrafía, geología y evaluación de peligros volcánicos del complejo volcánico de San Salvador, el Salvador”. (Tesis de Doctorado). Universidad Nacional Autónoma de México, México. Recuperado de <https://repositorio.unam.mx/contenidos/71400>.
- García Florez, Ignacio, (2016). Factores geológicos-geotécnicos que controlan los deslizamientos inducidos por terremotos en zonas de alta y media actividad sísmica: caso de El Salvador. Tesis doctoral para optar al grado de doctor en Universidad Complutense de Madrid, España.
- Jiawei Han, Micheline Kamber. Data mining: concepts and techniques.

- Marroquín, G., and Benito, M.B. (2009). “RESIS II Evaluación de la Amenaza Sísmica en El Salvador.”
- Multi-Period Response Spectra and Design Parameters (of the 2020 NEHRP Provisions and ASCE 7-22). Fuente: Proposed Revision of Table 20.3-1 to Include Three New Site Classes (BC, CD and DE) at Current Site Class Boundaries.
- SIHI-MARN (2019). Plataforma para visualización de sistema de información Hídrica de El Salvador.
- Panjamani Anbazhagan, Anjali Uday, Sayed S R Moustafa, Nassir S N Al-Arifi, Correlation of densities with shear wave velocities and SPT *N* values, *Journal of Geophysics and Engineering*, Volume 13, Issue 3, June 2016, Pages 320–341, <https://doi.org/10.1088/1742-2132/13/3/320>.
- Pedro Isasi Viñuela, Inés M. Galván León. Redes de neuronas artificiales: un enfoque práctico.
- RADES-MARN (2020). Plataforma de registros acelerográficos de El Salvador.
- Salazar, Walter & Brown, Lyndon & Hernández, Walter & Guerra, José. (2013). An Earthquake Catalogue for El Salvador and Neighbouring Central American Countries and its Implication in the Seismic Hazard Assessment. *Journal of Civil Engineering and Architecture*. 7. 1018-1045. 10.17265/1934-7359/2013.08.011.
- Schnabel, P. B., Lysmer, J., and Seed, H. B. (1972). "SHAKE: A Computer Program for Earthquake Ground Response Analysis for Horizontally Layered Sites." Report No: EERC 72- 12, Earthquake Engineering Research Center, University of California, Berkeley, California.
- Schmidt Díaz, V., (2010). Avances para estudios del riesgo sísmico a escala regional y local: aplicación a América Central y a la Bahía de Cádiz (Sur de España). Tesis Doctoral, Universidad Politécnica de Cataluña, Barcelona, España.
- Smith, V. C., Costa, A., Aguirre-Díaz, G., Pedrazzi, D., Scifo, A., Plunkett, G., ... & McConnell, J. R. (2020). The magnitude and impact of the 431 CE Tierra Blanca Joven eruption of Ilopango, El Salvador. *Proceedings of the National Academy of Sciences of the United States of America-PNAS*, 117(42), 26061-26068.



MINISTERIO DE
MEDIO AMBIENTE
Y RECURSOS
NATURALES

www.marn.gob.sv | medioambiente@marn.gob.sv

