



ESPE

**UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA**

DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

**CARRERA DE INGENIERÍA EN ELECTRÓNICA, Y
TELECOMUNICACIONES**

**TRABAJO DE TITULACIÓN, PREVIO A LA OBTENCIÓN DEL
TÍTULO DE INGENIERO EN ELECTRÓNICA, Y
TELECOMUNICACIONES**

**TEMA: GENERACIÓN DE MODELOS MULTIVARIANTES NO
LINEALES DE UNA PLANTILLA DE PRESIÓN PLANTAR A
TRAVES DE TÉCNICAS DE MACHINE LEARNING**

**AUTOR:
SR. PÉREZ SORIA, CRISTIAN GABRIEL**

DIRECTOR: ING. GUAMAN NOVILLO, ANA VERÓNICA PHD.

SANGOLQUÍ

2017



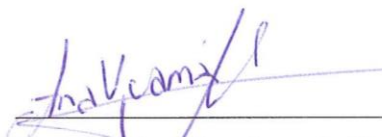
DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

**CARRERA DE INGENIERÍA EN ELECTRÓNICA,
Y TELECOMUNICACIONES**

CERTIFICACIÓN

Certifico que el trabajo de titulación: “*GENERACIÓN DE MODELOS MULTIVARIANTES NO LINEALES DE UNA PLANTILLA DE PRESIÓN PLANTAR A TRAVES DE TÉCNICAS DE MACHINE LEARNING*”, realizado por el señor *PEREZ SORIA CRISTIAN GABRIEL*, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar al señor *PEREZ SORIA CRISTIAN GABRIEL* para que lo sustenten públicamente.

Sangolquí, Septiembre del 2017



Ing. Ana Verónica Guamán Novillo PhD.

DIRECTOR



DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA EN ELECTRÓNICA,
Y TELECOMUNICACIONES

AUTORÍA DE RESPONSABILIDAD

Yo, **PEREZ SORIA CRISTIAN GABRIEL**, con cédula de identidad N°: 1722793195, declaro que este trabajo de titulación **“GENERACIÓN DE MODELOS MULTIVARIANTES NO LINEALES DE UNA PLANTILLA DE PRESIÓN PLANTAR A TRAVES DE TÉCNICAS DE MACHINE LEARNING”** ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaro que este trabajo es de mi autoría, en virtud de ello me declaro responsable del contenido, veracidad y alcance de la investigación mencionada.

Sangolquí, Julio del 2017

Sr. Pérez Soria Cristian Gabriel

C.C. 1722793195



DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

**CARRERA DE INGENIERÍA EN ELECTRÓNICA,
Y TELECOMUNICACIONES**

AUTORIZACIÓN

Yo, ***PEREZ SORIA CRISTIAN GABRIEL***, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca virtual de la institución el presente trabajo de titulación “***GENERACIÓN DE MODELOS MULTIVARIANTES NO LINEALES DE UNA PLANTILLA DE PRESIÓN PLANTAR A TRAVES DE TÉCNICAS DE MACHINE LEARNING***”, cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

Sangolquí, Julio del 2017

Sr. Pérez Soria Cristian Gabriel

C.C. 1722793195

DEDICATORIA

Quiero dedicar este trabajo de investigación a mi madre Irene, que incondicionalmente ha estado presente durante toda mi vida en logros y fracasos, con el constante apoyo aún más en el transcurso de mi formación profesional ya que gracias a ella es posible la elaboración de este proyecto, quiero recalcar además que ella me ha demostrado que con mucho esfuerzo y perseverancia se logran las metas propuestas aunque el camino hacia esas metas se torne difícil o parezca imposible.

A mi abuelo Gilberto, que ha sido como un padre para mí, aunque ya no se encuentre junto a nosotros hubiera querido entregarle a él mi título pero ya que Dios lo quiso así no fue posible, de todas maneras sé que en donde él se encuentre estará orgulloso de mi, ya que este trabajo de investigación también es fruto de sus enseñanzas y valores que con tanto cariño y sabiduría ha plasmado.

A mi abuela Beatriz, por ser mi segunda madre y estar pendiente de que vaya por buena senda a nivel personal y profesional, enseñando siempre a conservar la humildad y el respeto hacia las demás personas.

A mi padre Marco, por apoyarme cuando lo necesito con sus mejores intenciones y consejos, a pesar que por situaciones de la vida no se encuentre siempre a mi lado, estará presente no solo como un padre sino también como un amigo.

Dedico este trabajo de investigación a ustedes, porque ustedes han formado un pilar fundamental en mi vida para poder concluir la carrera universitaria con la realización de este trabajo.

Cristian Gabriel Pérez Soria

AGRADECIMIENTO

Agradezco a Dios en primera instancia por darme la madre que tengo, además de la sabiduría para poder desenvolverme dentro de la carrera universitaria y así poder concluir con el presente proyecto de titulación.

A mi madre Irene, por nunca dejar que me rindiera, aportando siempre con apoyo moral durante las incansables veladas de estudio fundamentales para poder aprobar cada uno de los niveles de la carrera universitaria, sin olvidar de los concejos que harían que mi vida vaya por buen camino, siendo una excelente guía llena de valores y cariño, Gracias madre sin ti no sería posible esto.

A mi novia Jessica, por ser una persona tan dulce y llena de buenos sentimientos e intenciones, brindándome todo su apoyo y cariño durante más de media carrera universitaria siendo un apoyo fundamental no solo en el transcurso de la ingeniería sino también por el cambio que ha incentivado en mi vida ayudándome a ser cada día que pasa mejor persona, demostrándome todo lo bueno que las personas pueden llegar a tener a pesar de que en una primera perspectiva no lo parezca; a pesar de todos los altos y bajos que hemos vivido juntos ha estado siempre junto a mí, por eso y muchas cosas que no puedo expresar en estas cortas líneas, Gracias Jessica y si Dios lo permite sigas formando parte de mi vida.

A mi familia por brindarme el entendimiento de estar ausente en varias ocasiones en donde la unión de toda la familia era importante, pero gracias a esos sacrificios eh podido concluir con mi carrera universitaria, adicional quiero agradecerles por formar parte en mi vida y lo que cada uno apporto para que pudiera finiquitar con este proyecto de titulación.

A todos mis amigos y compañeros que formaron parte de esta etapa de mi vida quiero agradecerles por brindarme su amistad y muchas veces su apoyo como equipo de estudio, viviendo experiencias inolvidables en el transcurso de mi formación académica,

muchas gracias a todas esas personas que saben que han formado parte de mi vida y que si ustedes lo permiten sigamos siendo buenos amigos.

A la Dra. Ana Guamán, por ser una docente tan inteligente y capaz, que sin dudar me ha brindado las herramientas y conocimientos necesarios para poder realizar el presente proyecto de investigación, y en general a todos los excelentes profesionales con los que cuenta la Universidad de las Fuerzas Armadas – ESPE, que han sido unos muy buenos tutores durante el proceso académico como docentes y como amigos, Gracias.

Cristian Gabriel Pérez Soria

ÍNDICE DE CONTENIDO

CERTIFICACIÓN.....	ii
AUTORÍA DE RESPONSABILIDAD	iii
AUTORIZACIÓN	iv
DEDICATORIA.....	v
AGRADECIMIENTO	vi
ÍNDICE DE CONTENIDO	viii
ÍNDICE DE TABLAS	xi
ÍNDICE DE FIGURAS	xii
RESUMEN.....	xv
ABSTRACT	xvi
CAPÍTULO I	1
INTRODUCCIÓN	1
1.1. Antecedentes	1
1.2. Justificación e importancia.....	4
1.3. Objetivos	6
1.4.1. Objetivo general	6
1.4.2. Objetivos específicos.....	6
CAPÍTULO II	7
MARCO TEÓRICO	7
2.1. Introducción.....	7
2.2. Regresión univariante	8
2.2.1. Regresión lineal univariante.....	8
2.2.1.1. Estimación o predicción	9
2.2.2. Regresión no lineal univariante	10

2.3. Regresión multivariable	10
2.4. Support Vector Regression.....	13
2.4.1. Validación de parámetros.....	17
CAPÍTULO III	20
ANÁLISIS DE DATOS PRECEDENTES	20
3.1. Introducción.....	20
3.2. Plantilla de presión plantar precedente	20
3.3. Respuesta no lineal de sensores piezoresistivos	23
3.4. Correlación entre sensores adyacentes.....	24
3.5. Análisis de resultados con calibración univariante.....	25
3.6. Análisis de resultados con calibración multivariable	27
3.4.1. Construcción de matrices para entrenar y predecir el modelo	29
3.4.2. Validación cruzada del modelo multivariante.....	31
CAPÍTULO IV	35
MODELOS MULTIVARIABLES DE ÁREAS CON MAYOR PRESIÓN PLANTAR	35
4.1. Introducción.....	35
4.2. Delimitación de áreas con mayor presión plantar	35
4.3. Diseño de Protocolo de Experimentos.....	37
4.4. Generación de Modelos de Regresión Multivariantes	44
4.5. Entrenamiento de Modelos de Regresión Multivariantes	45
4.6. Optimización de Parámetros con Cross Validation	46
CAPÍTULO V	61
VALIDACIÓN DE MODELO MULTIVARIANTE	61
5.1. Introducción.....	61
5.2. Predicción y pruebas de modelo multivariantes	61
5.3. Interfaz de comunicación PC-plantilla de presión plantar	67

5.4. Pruebas con pisada normal.....	70
CAPÍTULO VI	77
CONCLUSIONES Y RECOMENDACIONES	77
6.1. Conclusiones	77
6.2. Recomendaciones	78
BIBLIOGRAFÍA	80

ÍNDICE DE TABLAS

Tabla 1 Base de datos organizada	12
Tabla 2 Materiales de elaboración de plantilla precedente	21
Tabla 3 Matriz XCF1	29
Tabla 4 Matriz YCF1	30
Tabla 5 Matriz XC4 para entrenamiento	30
Tabla 6 Matriz YC4 para entrenamiento	30
Tabla 7 Matriz XTF1 para validación	31
Tabla 8 Matriz YTF1 para validación	31
Tabla 9 Optimización de parámetros con datos precedentes	32
Tabla 10 Número de sensores distribuidos en cada área o modelo de regresión	45
Tabla 11 Optimización con k-fold cross validation para modelo hallux	47
Tabla 12 Optimización con k-fold cross validation para modelo dedos del pie	48
Tabla 13 Optimización con k-fold cross validation para modelo CM1	49
Tabla 14 Optimización con k-fold cross validation para modelo CM2.....	50
Tabla 15 Optimización con k-fold cross validation para modelo CM3.....	51
Tabla 16 Optimización con k-fold cross validation para modelo CM4.....	52
Tabla 17 Optimización con k-fold cross validation para modelo CM5.....	53
Tabla 18 Optimización con k-fold cross validation para modelo midfoot	54
Tabla 19 Optimización con k-fold cross validation para modelo heel	55
Tabla 20 Resultados de RMSE Y R2 de Modelos Multivariantes.....	66
Tabla 21 Resultados finales de fuerzas predichas con pruebas de pisada normal ..	76

ÍNDICE DE FIGURAS

Figura 1 División del pie en 10 áreas	2
Figura 2 Plantilla de presión plantar.....	4
Figura 3 Regresión lineal Univariante	8
Figura 4 Regresión no lineal univariante	10
Figura 5 Diagrama de dispersión en 3 ejes	11
Figura 6 SVR lineal	14
Figura 7 SVR No lineal	15
Figura 8 Separación lineal por sobredimensión	16
Figura 9 Espacio de funciones inducido por el kernel.....	16
Figura 10 Método de k-fold cross validation	18
Figura 11 Método de random cross validation.....	19
Figura 12 Método de leave one out cross validation	19
Figura 13 Materiales plantilla de presión.....	21
Figura 14 Distribución de capas en la construcción del sensor	21
Figura 15 Estructura de un FSR	22
Figura 16 Prototipo de plantilla para medir la fuerza de la pisada.....	23
Figura 17 Respuesta del sensor fuerza vs voltaje.....	23
Figura 18 Correlación entre sensores adyacentes.....	24
Figura 19 Curva de comportamiento del sensor de la columna 4 de fila 1 (x_4).....	25
Figura 20 Regresión univariante - sensor C4.....	26
Figura 21 Modelo de calibración – sensor C4	26
Figura 22 Sensores distribuidos en la plantilla de presión plantar	27
Figura 23 Modelo multivariable fila 1 fuerza predicha vs fuerza real con datos de entrenamiento	33
Figura 24 Modelo multivariable fila 1 fuerza predicha versus fuerza real con datos de validación.....	34
Figura 25 Plantilla con 9 áreas de presión plantar.....	37
Figura 26 Bloque hueso hallux.....	38
Figura 27 Bloque huesos dedos de pie menores	38
Figura 28 Bloque 1 cabeza metatarsiana 1.....	38
Figura 29 Bloque 2 cabeza metatarsiana 2.....	39
Figura 30 Bloque 3 cabeza metatarsiana 3.....	39

Figura 31 Bloque 4 cabeza metatarsiana 4.....	39
Figura 32 Bloque 5 cabeza metatarsiana 5.....	40
Figura 33 Bloque hueso pie medio	40
Figura 34 Bloque hueso heel.....	40
Figura 35 Nueve Prototipos impresos en 3D	41
Figura 36 Pruebas finales con máquina de presión puntual.....	41
Figura 37 Diseño de Hardware para interpretación de datos de sensores	42
Figura 38 Hardware terminado para interpretación de datos de sensores	42
Figura 39 Recepción de datos de PCB por puerto de comunicación COM 7	43
Figura 40 Respuesta de sensores por comunicación serial alámbrica	43
Figura 41 Sensores distribuidos en cada área con mayor presión plantar	44
Figura 42 K-fold cross validation con k=10	46
Figura 43 Fuerza predicha vs. fuerza real aplicada en área hallux con datos de entrenamiento.....	56
Figura 44 Fuerza predicha vs. fuerza real aplicada en área dedos menores del pie con datos de entrenamiento.....	57
Figura 45 Fuerza predicha vs. Fuerza real aplicada en área CM1 con datos de entrenamiento.....	57
Figura 46 Fuerza predicha vs. fuerza real aplicada en área CM2 con datos de entrenamiento.....	58
Figura 47 Fuerza predicha vs. fuerza real aplicada en área CM3 con datos de entrenamiento.....	58
Figura 48 Fuerza predicha vs. fuerza real aplicada en área CM4 con datos de entrenamiento.....	59
Figura 49 Fuerza predicha vs. fuerza real aplicada en área CM5 con datos de entrenamiento.....	59
Figura 50 Fuerza predicha vs. fuerza real aplicada en área midfoot con datos de entrenamiento.....	60
Figura 51 Fuerza predicha vs. fuerza real aplicada en área heel con datos de entrenamiento.....	60
Figura 52 Fuerza predicha vs. fuerza real aplicada en área hallux con datos de validación.....	62
Figura 53 Fuerza predicha vs. fuerza real aplicada en área dedos menores del pie con datos de validación	62

Figura 54 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 1 con datos de validación	63
Figura 55 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 2 con datos de validación	63
Figura 56 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 3 con datos de validación	64
Figura 57 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 4 con datos de validación	64
Figura 58 Fuerza predicha vs. fuerza real aplicada en área cabeza	65
Figura 59 Fuerza predicha vs. fuerza real aplicada en área midfoot con datos de validación.....	65
Figura 60 Fuerza predicha vs. fuerza real aplicada en área heel con datos de validación.....	66
Figura 61 Módulo Bluetooth HC-06	68
Figura 62 Plantilla de presión plantar con módulo bluetooth	68
Figura 63 Respuesta de sensores por comunicación serial inalámbrica	69
Figura 64 Variables Independientes de pisada normal del área Hallux.....	70
Figura 65 Fuerza predicha en área hallux	71
Figura 66 Fuerza predicha en área dedos menores del pie	71
Figura 67 Fuerza predicha en área cabeza metatarsiana 1	72
Figura 68 Fuerza predicha en área cabeza metatarsiana 2	72
Figura 69 Fuerza predicha en área cabeza metatarsiana 3	73
Figura 70 Fuerza predicha en área cabeza metatarsiana 4.....	73
Figura 71 Fuerza predicha en área cabeza metatarsiana 5	74
Figura 72 Fuerza predicha en área midfoot	74
Figura 73 Fuerza predicha en área heel	75

RESUMEN

En el presente proyecto se propone generar modelos de calibración multivariantes de plantillas de presión para ayudar al estudio y prevención de afectaciones en nueve áreas del pie con mayor presión plantar. Para ello es necesario entender la dinámica existente en los sensores de la plantilla y estudiar la correlación que existe entre la fuerza aplicada y el voltaje generado por cada uno de los sensores piezoresistivos. Luego se creará una nueva base de datos aplicando fuerzas controladas en cada área del pie usando máquinas de presión del laboratorio de Resistencia de los Materiales del Departamento de Ciencias de la Energía y Mecánica de la Universidad de las Fuerzas Armadas-ESPE. Con la información obtenida se generarán modelos multivariantes como Support Vector Regression. Se realizarán pruebas de validación de los modelos obtenidos para finalmente implementar un modelo por cada área de presión y monitorizar la fuerza de la plantilla en un programa ejecutado en una PC a través de una interfaz de comunicación. Esta plantilla de presión otorgará una herramienta más para el diagnóstico de posibles anomalías o afectaciones en las áreas con mayor presión plantar del pie, pretendiendo prevenir consecuencias más graves y costosas. Se obtienen modelos multivariantes de una plantilla de presión plantar calibrada lo más exacto posible, delimitada por las nueve áreas con mayor presión plantar.

PALABRAS CLAVE:

- **APRENDIZAJE DE MÁQUINA**
- **VECTORES DE SOPORTE DE REGRESION (SVR)**
- **CALIBRACIÓN**

ABSTRACT

In the current project, the use of multivariate calibration models was proposed for calibration of pressure insoles as solution in the study and prevention of affectations in nine areas of the foot with biggest plantar pressure. In order to achieve the main objective of this project, it was necessary to understand the sensors dynamics in the insole and study the correlation between the applied force and the voltage generated by each of the piezoresistive sensors. A new database was created where controlled forces to each area was applied of the foot using pressure machine from Materials Resistance Laboratory of the Department of Energy and Mechanics of the University of the Armed Forces (ESPE). Once the information was obtained, multivariate models were generated as support vector regression. Then, the final models were realized to finally implement a model of each area of pressure and monitor the strength of the template in a program executed in a PC through a communication interface. This pressure insole will provide an additional tool for the diagnosis of possible anomalies or affectations in the areas with biggest pressure of the foot, in order to prevent more serious and costly consequences. Multivariate models of a calibrated plantar pressure insole are generated as accurate as possible, delineated by the nine areas with biggest plantar pressure.

KEY WORDS:

- **MACHINE LEARNING**
- **SUPPORT VECTOR REGRESSION (SVR)**
- **CALIBRATION**

CAPÍTULO I

INTRODUCCIÓN

1.1. Antecedentes

Las enfermedades que se pueden producir por una mala distribución de presión plantar al momento de realizar una pisada en contacto con el suelo pueden ser desapercibidas, silenciosas y muy peligrosas. Las presiones plantares son el reflejo de la forma de apoyo durante la posición estática “de pie” o durante la marcha; de hecho, el modo de apoyo depende de las características estructurales y funcionales del pie (Morag, E; Cavanagh, P R, 1999), y de las otras articulaciones de la extremidad (Wrobel, Connolly, & Beach, 2004).

Así mismo, la postura, la disimetría en extremidades y la patología podológica introducen alteraciones en la intensidad y forma en que se distribuyen las presiones plantares en el área de la pisada (Martínez Assucena, y otros, 2003). Es así que los estudios han demostrado que las complicaciones en la movilidad de las articulaciones, la formación de callos, y varias alteraciones en el pie pueden ser causados por las elevadas presiones plantares en posición estática y durante la marcha (Brand, 1979). (Pecoraro, Rehntr, & Burgess, 1990). (San Tsung, Zhang, Tat Mak, & Nar Wong, 2004).

Existen dos métodos para medir la distribución de presión bajo el pie durante la marcha. El primer método utiliza un sistema de placa de fuerza, por lo general las placas de fuerza están introducidas en el suelo de una plataforma, por lo tanto la persona debe golpear sus pies dentro de los límites de la superficie de medición (Bauman & Brand, 1963). Sin embargo, este sistema resulta ineficaz al tratar de medir la presión de los pasos en una marcha en condiciones normales, ya que la persona se encuentra en pie descalzo y no se tiene delimitada el área del pie en donde exactamente realiza la presión. Otro método para medir la presión que ejerce el pie con el suelo es colocar sensores en la plantilla de un zapato, este método puede medir las presiones plantares durante la actividad normal de la persona con su calzado nominal de uso diario (Bauman & Brand, 1963).

Se han desarrollado varias plantillas de presión de distintos materiales para identificar y corregir las áreas con mayor presión plantar durante la marcha, así como también se han desarrollado plantillas para reducir la presión plantar en áreas específicas como: las cinco cabeceras metatarsianas, el pie medio, el talón medial y lateral; y en personas con amputación de primer rayo (primer metatarsiano y hueso Hallux) causados por la diabetes (El-Hilaly, Elshazly, & Amer, 2013).

Existe un caso puntual de plantilla ya elaborada, la plantilla Biofoot (Pérez García, 2004), que consta de 64 sensores piezoeléctricos, divididos en 10 áreas que corresponden al talón posterior (Heel), pie medio, todo el antepié, las cinco cabezas metatarsianas, hueso halux y dedos de los pies menores, como se observa en la figura 1 (Pérez García, 2004). La prueba de la plantilla se lo realizó en marcha y en caminata normal (Bauman & Brand, 1963), además existen trabajos de investigación orientados al análisis de la precisión de las plantillas de presión plantar en un escenario de carrera rápida (Kati, Forrester, & Fleming, 2010).

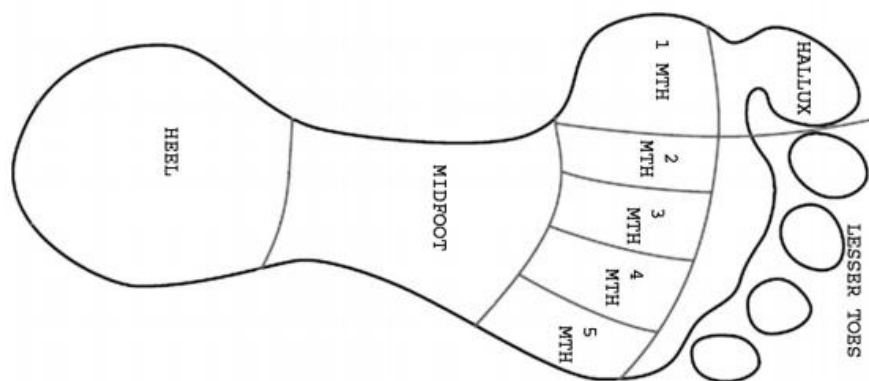


Figura 1 División del pie en 10 áreas

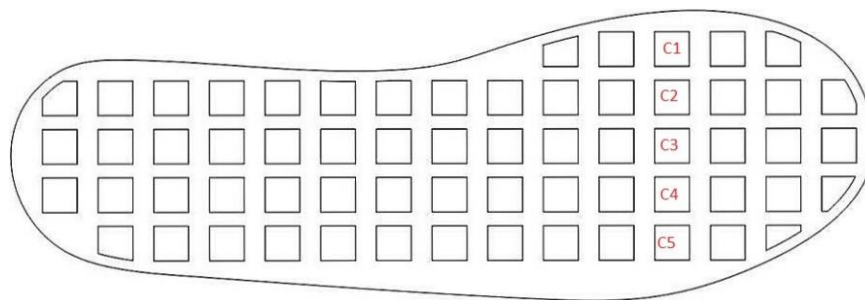
Fuente: (Pérez García, 2004)

En varias investigaciones el análisis de datos que brindan los sensores se lo realiza a través de dos métodos de transmisión: obtención de datos por medios cableados, y obtención de datos por medios inalámbricos (Hadi, Razak, Zayegh, Begg, & Wahab, 2012). Para la obtención de datos por medios de transmisión por cable se han usado un tipo de conexión serial mientras que para la obtención de datos

por medios de transmisión inalámbrica se usa “Smart Wireless DAQ and control ” (Hadi, Razak, Zayegh, Begg, & Wahab, 2012), permitiendo mejorar los diseños y movilidad. Los datos de sincronización del tiempo, cálculos de fuerzas, correlaciones, errores cuadráticos medios (RMSEs) y los datos de medición brutos son calibrados usando Matlab (SR) R2013b (Math Works Inc., Natick, EE.UU.) e IBM (SPSS) Estadísticas 22 (IBM Corporation, NewYork, EE.UU.). Finalmente Koch y Lunde (2016) usaron para calcular los valores de fuerza con la sumatoria de las presiones de todos los sensores individuales de la plantilla.

La Universidad de las Fuerzas Armadas ESPE ha elaborado un primer prototipo de plantilla de presión (Bermeo & Guamán, 2017). La plantilla se basa en 63 sensores piezoresistivos distribuidos en toda el área de presión plantar tal como se observa en figura 2(a) e implementada tal como se observa en la figura 2(b). En una etapa inicial se realizó un conjunto de pruebas para entender el comportamiento de los sensores de forma individual. Las pruebas se realizaron ejerciendo una fuerza variable creciente en cada uno de los 63 sensores y se almaceno todos los datos en una matriz de 15166 filas por 6 columnas, donde las filas representan el número de datos obtenidos en el experimento y las columnas representan los cinco sensores por cada fila más la sexta columna que es la fuerza aplicada.

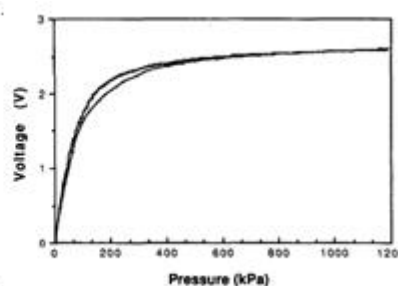
En las pruebas obtenidas la respuesta de los sensores presentan una respuesta no lineal tal como se observa en la figura 2(c). Además, se observó que la información de los sensores está correlacionadas entre ellas impidiendo obtener modelos de calibración fiables por cada sensor o por secciones de sensores. Por consiguiente, para resolver estos problemas se requiere un abordaje de calibración multivariante utilizando técnicas actuales de machine learning que permitan establecer modelos no lineales y eliminar la correlación existente entre los sensores.



(a) Plantilla elaborada con 63 sensores



(b) Plantilla con Sensores Piezoeléctricos



(c) Respuesta no lineal de los sensores

Figura 2 Plantilla de presión plantar

Fuente: (Bermeo & Guamán, 2017).

1.2. Justificación e importancia

Con la elaboración de una plantilla de presión plantar de la Universidad de las Fuerzas Armadas ESPE se dispone de un prototipo formado por 63 sensores piezoresistivos distribuidos en toda el área plantar siendo cada sensor capaz de percibir una fuerza de presión de 0N hasta 1200N con una exactitud de hasta el 70% en un periodo de tiempo determinado por las pruebas. Los datos de los sensores se enviaron por transmisión serial en tiempo real hacia una PC y fueron procesados con Matlab (SR) donde se almacenaron los datos en una matriz de 15366x6 según Bermeo y Guamán (2016). Este prototipo de plantilla de presión plantar sería un aporte al diagnóstico para una posterior solución de los problemas que se generan en los pies, por los valores pico de las presiones plantares en ciertas áreas con mayor afectación, ya sea por mala postura o por mala marcha al caminar.

En el prototipo existente cada sensor es independiente en la toma de datos en función de la presión ejercida puntualmente en cada uno. Una prueba preliminar fue realizada en el prototipo existente con el fin de calibrar los sensores y ver la relación que podría existir con la fuerza aplicada durante la marcha. Se encontró que los 63 sensores no son réplicas exactas (Bermeo & Guamán, 2017) entre ellas, y además se observó que hay una alta correlación en la respuesta de cada uno de los sensores, de hecho el aplicar una fuerza puntal en un sensor determinado, los sensores adyacentes tienden a cambiar su resistividad de forma no lineal. Como metodología de calibración se aplicó la técnica de regresión univariante ajustando polinomios de tercer orden. Sin embargo los resultados fueron pobres en términos de calibración obteniendo sensibilidad de ajuste (R^2) entre 0.5 al 0.7.

Tomando en cuenta que esta plantilla debe ser correctamente calibrada para poder predecir los valores de fuerza aplicada por futuros sujetos de prueba, es necesario garantizar modelos de calibración con errores inferiores a los ya obtenidos con modelos univariantes. Este problema podría ser solucionado utilizando métodos de calibración multivariantes para rechazar la correlación existente entre cada uno de los sensores y modelar las no linealidades. Este modelo multivariable servirá para poder predecir la fuerza ejercida durante la marcha.

1.3. Objetivos

1.4.1. Objetivo general

- Generar modelos de calibración multivariantes no lineales de la plantilla piezoresistiva.

1.4.2. Objetivos específicos

- Definir el estado del arte de modelos multivariantes, y las áreas con presión plantar elevadas.
- Establecer los sensores de la plantilla de presión plantar para cada área definida con mayor presión plantar.
- Generar modelos multivariantes por cada área delimitada.
- Elaborar un algoritmo computacional que permita calibrar la plantilla en sus diferentes áreas de presión.
- Diseñar una interfaz de comunicación serial alámbrica o inalámbrica para el monitoreo de la fuerza aplicada en la plantilla.
- Determinar el protocolo de pruebas de los modelos multivariantes.

CAPÍTULO II

MARCO TEÓRICO

2.1. Introducción

Ante la necesidad de calibrar correctamente la respuesta de cada sensor y eliminar las correlaciones existentes entre ellos, es necesario introducir herramientas de regresión multivariante. En el presente capítulo se empezará abordando información correspondiente a regresión univariante lineal y no lineal, necesarios para poder entender posteriormente regresión multivariable.

En general, la regresión establece relaciones o dependencias entre dos variables x e y , donde estas relaciones pueden ser funcionales (determinísticas), o aleatorias (no determinísticas). Funcionales o determinísticas es cuando existe una función matemática exacta que relacione ambas variables, y aleatorias o no determinísticas cuando no necesariamente existe una relación exacta pero se puede observar una cierta tendencia de relación ante los comportamientos de ambas variables. Para poder observar la tendencia de relación entre ambas variables es necesario un diagrama de dispersión en donde representa los datos de ambas variables en el plano como una nube de puntos, entonces la regresión consiste en ajustar la nube de puntos para ambas variables x e y , de manera aproximada a una función conocida (Gorgas García, Cardiel López, & Zamorano Calvo, 2011).

En el caso de una regresión lineal en donde la nube de puntos se distribuye a lo largo del plano tomando una tendencia de una función lineal, al ajustar la nube de puntos se formaría entonces usando una recta de regresión; o ya sea el caso de regresión no lineal donde la nube de puntos se distribuya tomando la forma de una función no lineal que al ajustar tomaría la forma de una curva de regresión (Gorgas García, Cardiel López, & Zamorano Calvo, 2011).

En regresión, se toma en consideración que la variable y puede estar en función de una variable x o varias variables $x_1, x_2, x_3, \dots, x_n$; en donde el caso de única variable x toma lugar para el análisis de regresión univariante, mientras que para el caso de varias variables se considera el análisis de regresión multivariable. Para el

estudio de regresión multivariable se hará énfasis en la técnica de Support Vector Regression y los respectivos parámetros de validación y optimización (García, 2008).

2.2. Regresión univariante

La regresión univariante es utilizada para ajustar la relación o dependencia entre una variable dependiente y , con una variable independiente x (García, 2008), donde se obtendrá un modelo explicativo de y en función de única variable x . El modelo puede distribuirse de tal forma que se pueda asociar a una relación lineal según la dispersión de la nube de puntos, o bien la dispersión puede distribuirse de tal forma que se la asocie a una relación no lineal. (Gorgas García, Cardiel López, & Zamorano Calvo, 2011).

2.2.1. Regresión lineal univariante

Si los valores a ser ajustados gráficamente muestran una correlación que tiende a ser lineal, entonces se obtiene una recta al ajustar los valores, como se observa en la figura 3.

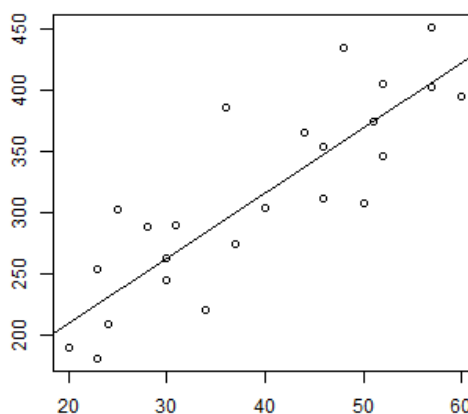


Figura 3 Regresión lineal Univariante

De esta manera el modelo de Regresión lineal está dado por la ecuación de la recta (1):

$$y = ax + b + \epsilon \quad (1)$$

Donde, el coeficiente a es la covarianza de las dos variables x e y , dividida por la varianza de la variable x , además establece la pendiente o inclinación de la recta, el coeficiente b es el valor de ajuste de error para que la variable independiente x se relacione más aproximado con la variable dependiente y , además establece el valor donde la variable dependiente y cruza por el eje vertical cuando la variable independiente x es igual a cero (García, 2008).

El coeficiente ϵ es una variable aleatoria distribuida con $E(\epsilon) = 0$ y $Var(\epsilon) = \sigma^2$, donde la cantidad σ^2 normalmente se la denomina varianza del error o varianza residual, es decir que ϵ vendría a ser el error del modelo con distribución normal y media igual a cero, y varianza constante σ^2 . Este error es añadido debido a que se considera que la relación entre x e y no es perfecta (Walpole, Myers, Myers, & Ye, 2007).

2.2.1.1. Estimación o predicción

Para la estimación o predicción en regresión lineal el objetivo es poder predecir los valores de la variable y en función de la variable x . De este modo se genera la ecuación de regresión y' en base a cada x como se observa en la ecuación (2) (Salinas, 2005).

$$y' = ax + b \quad (2)$$

La variable x corresponde a cada valor de predicción de y' , que además también corresponde a cada valor de y en la ecuación (1) pero con el error residual ϵ que corresponde a cada valor de y , por lo que el error residual estará presente en la ecuación de predicción y' representado por la ecuación 3. La ecuación (3) es verificable también si se reemplaza la ecuación (2) en la ecuación (1).

$$y = y' + \epsilon \quad (3)$$

Despejando y' de la ecuación (3) se obtendrá entonces la estimación o predicción en la ecuación (4).

$$y' = y - \epsilon \quad (4)$$

2.2.2. Regresión no lineal univariante

En ciencias e ingeniería existen muchos fenómenos de naturaleza no lineal y, cuando se conoce la estructura real, se debería intentar ajustar al modelo real. Los modelos no lineales no determinísticos o aleatorios se refieren a condiciones no ideales, en las cuales se asegura de que la respuesta y , y por lo tanto, la respuesta del error del modelo ϵ , no tiene distribución normal, sino que más bien siguen una distribución binomial, o de Bernoulli, o de Poisson como es de esperar en problemas reales (Walpole, Myers, Myers, & Ye, 2007).

En Regresión no lineal los valores a ser ajustados gráficamente muestran una correlación que tiende a ser no lineal, entonces se pretende obtener una curva al ajustar los valores tal como se observa en la figura 4.

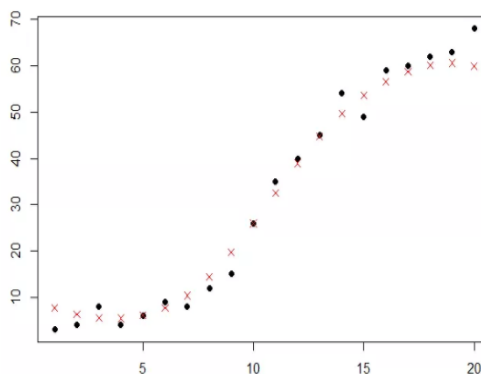


Figura 4 Regresión no lineal univariante

Fuente: (Kowalczyk, 2014)

Por lo general se tienen escenarios en que el caso lineal es común contar con una varianza que resulta independiente de la media, pero en un escenario donde se encuentran presenten regresores no lineales se omite este tipo de varianza (Walpole, Myers, Myers, & Ye, 2007).

2.3. Regresión multivariable

El propósito general de la regresión multiparamétrica o multivariable es mejorar la relación entre varias variables independientes o predictoras con una variable dependiente (StafSoft, 2013), por tanto el análisis multivariable ya no está definido

por la ecuación de la recta en el plano, más bien se define un hiperplano en un espacio multidimensional (León Velásquez, 2015).

Para el caso de una variable dependiente y y con dos variables x_1, x_2 , se necesitan tres ejes para poder representar el diagrama de dispersión como se observa en la figura 5. Pero si se presentan tres variables independientes x_1, x_2, x_3 sería necesario un espacio de cuatro ejes para poder representar el diagrama de dispersión lo cual no es posible visualizar (León Velásquez, 2015). Entonces si un modelo de regresión planteado presenta más de dos variables independientes, la representación gráfica por medio del diagrama de dispersión resulta ser poco intuitiva, razón por la cual se parte de la ecuación (1) del modelo de regresión lineal donde un modelo multivariable se expresa en la ecuación (5) (León Velásquez, 2015).

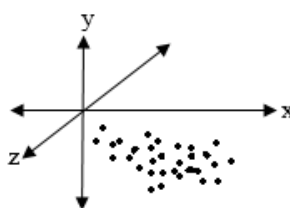


Figura 5 Diagrama de dispersión en 3 ejes

$$Y = a_1X_k + a_2X_2 + \dots + a_kX_k + b + \epsilon. \quad (5)$$

Donde la variable dependiente Y se interpreta como una combinación lineal de un conjunto de k variables independientes, en las cuales cada una de las variables va acompañado de un coeficiente a , que indica el peso relativo de la variable independiente en la ecuación. El componente aleatorio ϵ , que representa los errores o residuos que las variables independientes no son capaces de relacionar, de la misma forma que en el caso de regresión univariante, el error del modelo se encuentra con una distribución normal y media igual a cero, y varianza constante σ^2 .

En el caso de regresión lineal constaban los datos en pares de observaciones (x_i, y_i) de dos variables cuantitativas, para el caso de regresión multivariable se tiene varias variables independientes para cada variable dependiente con varios experimentos, representado por n y p respectivamente en la tabla 1. Por tanto la

variable dependiente Y viene a ser un vector de datos de n experimentos resultante de p variables.

Tabla 1

Base de datos organizada

1	x_{11}	x_{12}	\dots	x_{1p}	y_1
2	x_{21}	x_{22}	\dots	x_{2p}	y_2
$:$					
N	x_{n1}	x_{n2}	\dots	x_{np}	y_n

Fuente: (León Velásquez, 2015)

Partiendo del modelo de regresión lineal múltiple (5) se puede determinar la respuesta media, donde la respuesta media (u_y) (6) viene dada por una función lineal de las variables regresoras, además que está relacionada con la estimación de la variabilidad, denotada por $\sigma^{2'}$ en la ecuación 3 (León Velásquez, 2015).

$$u_Y = a_0 + a_1X_1 + a_2X_2 + \dots + a_kX_k \quad (6)$$

Al estar relacionada con la estimación de la variabilidad se podría predecir la media de la variable dependiente Y a través de la predicción de los parámetros de regresión u y $\sigma^{2'}$, similar a regresión lineal simple, en regresión multivariable se usa para la estimación el método de los mínimos cuadrados, pero en regresión multivariable la notación es matricial.

En la actualidad existe varias técnicas de Regresión entre las que más destacan son las siguientes (Microsoft, 2017):

- Bosque de decisión
- Árbol de decisión impulsado
- Random Forest
- Red Neuronal
- Regresión de Poisson
- Support vector regression
- Principal component regression

- Partial least squares regression
- Multiple linear regression
- Ordinary least squares

En la presente investigación se prestará principal interés en *Support Vector Regression* (SVR) ya que este modelo tiene mejor respuesta y predicción con regresiones no lineales (Kobayashi, Konishi, Fujita, & Ishigaki, 2005). Además se intenta determinar una fórmula que puede describir cómo los elementos en un vector de variables dependientes responden simultáneamente a los cambios de las variables independientes.

2.4. Support Vector Regression

Support Vector Regression (SVR) es un modelo que tiene origen en *Support Vector Machine* (SVM). SVM son máquinas de aprendizaje que implementan el principio inductivo de minimización del riesgo estructural para obtener una buena generalización en un número limitado de patrones de aprendizaje (Basak, Pal, & Chandra Patranabils, 2007).

SVR mantiene las características principales que son propias del algoritmo SVM, con pocas diferencias; por ejemplo SVR conserva el margen de tolerancia conocido también como ϵ (Sayad, 2010), además utiliza otra herramienta muy útil característica de SVM denominada *Kernels* necesaria para poder resolver modelos no lineales que más adelante es visible en la ecuación (10).

SVR lineal está representado en la figura 6, y definido por la ecuación 9; donde toma la forma de la ecuación de la recta (1) que hace referencia a la Regresión Lineal.

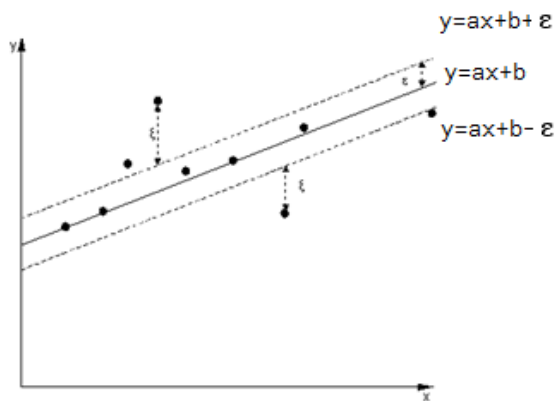


Figura 6 SVR lineal

Fuente: (TRIPOD, 2017)

La variable independiente x se mapea primero en un espacio de características m -dimensional usando algún mapeo fijo (no lineal), y luego se construye la ecuación (8) que además incluye la ecuación (9).

$$F(x, a) = \sum_{j=1}^N a_j g_j(x) + b \pm \epsilon \quad (8)$$

$$g_j(x), j = 1, \dots, N \quad (9)$$

Donde $F(x, a)$ es un modelo lineal en el espacio de características, el parámetro a viene a ser la pendiente puesto que como mencionamos $F(x, a)$ es un modelo lineal, y b es el término de “ajuste”, a menudo, se supone que los datos de b tienen una media igual a cero, por lo que el término de ajuste se elimina (Shi, Stumtner, Hao, & Quirchmayr, 2007). El Parámetro ϵ controla el ancho de la zona - insensible, que se utiliza para ajustar los datos de entrenamiento que además es visible en la figura 6. El valor de ϵ puede afectar el número de vectores de soporte utilizados para construir la función de regresión. Cuanto mayor sea, menos vectores de soporte se seleccionan. Por otro lado, los valores más grandes producen estimaciones más "planas", al decir planas se hace referencia al hiperplano representado por la figura 8 b). (TRIPOD, 2017). La función $g(x)$ representa un conjunto de transformaciones no lineales desde $j=1$ hasta N , donde N representa el número de muestras.

Si la variable independiente o regresora x se mapea en un espacio de características m -dimensional donde se aproxima un modelo no lineal en este espacio

de características, representado por la figura 7, el modelo lineal necesita de una optimización para la solución de modelos no lineales, por tanto se plantea la ecuación (10) que propone la optimización del modelo.

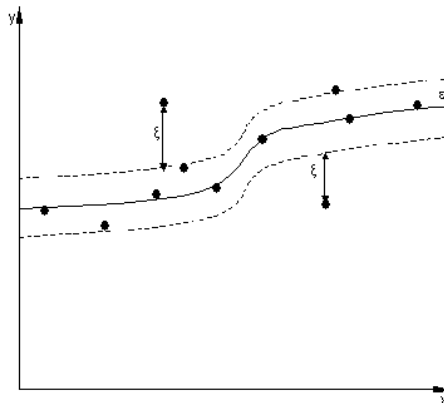


Figura 7 SVR No lineal

Fuente: (TRIPOD, 2017)

$$f(X) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) K(x, x_i) \quad (10)$$

Donde N es el número de Vectores de Soportes, α_i y α_i^* se encuentran dentro de los rangos: $0 \leq \alpha_i \leq C$, $0 \leq \alpha_i^* \leq C$. En los rangos planteados el parámetro C determina la compensación entre la complejidad del modelo y el grado en que las desviaciones son mayores que las toleradas en la ecuación (10). Por consiguiente si C es demasiado grande (infinito) el objetivo es minimizar el riesgo empírico de obtener un producto igual a cero así: $(\alpha_i - \alpha_i^*) K(x, x_i) = 0$; tomando en cuenta que si C es demasiado grande la complejidad del modelo en la formulación de *LaGrange* también se vería afectada. Por lo que C se usa con el menor valor para no afectar la complejidad de regularización en la formulación de *LaGrange* (TRIPOD, 2017).

Finalmente la función de *kernel* $K(x, x_i)$ está dada por la ecuación (11), además la función de *kernel* incluye la función $g_j(x)$ dada por la ecuación (9).

$$K(x, x_i) = \sum_{j=1}^m g_j(x) g_j(x_i) \quad (11)$$

El uso de *kernels* hace posible realizar la separación lineal, esto debido a que si el problema a ser analizado se encuentra en una dimensión determinada D tal como se puede ver en la figura 8(a), el algoritmo logra trabajar en una dimensión $D + 1$ representado por la figura 8(b), entonces las funciones del kernel transforman los datos en una característica espacial de mayor dimensión (Ruckstuhl, 2010).

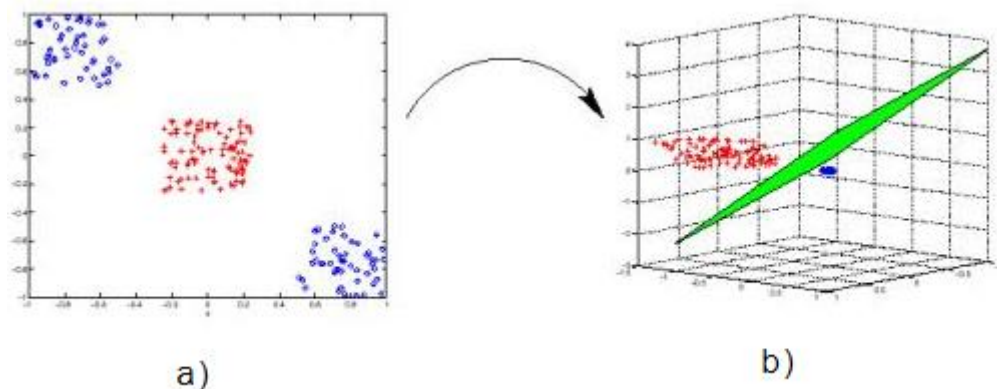


Figura 8 Separación lineal por sobredimensión

Fuente: (TRIPOD, 2017)

Para poder separar el conjunto de datos en SVR no lineal mapeamos los datos del espacio de la variable regresora x dentro de un espacio g , representado con la ecuación (12) y la figura 9; posteriormente se puede modificar la hipótesis para mapear implícitamente los datos del espacio x con otro espacio de funciones, que está representado por las ecuaciones (13) y (14)

$$x = (x_1, \dots, x_{11}) \rightarrow g(x) = (g_1(x), \dots, g_N(x)) \quad (12)$$

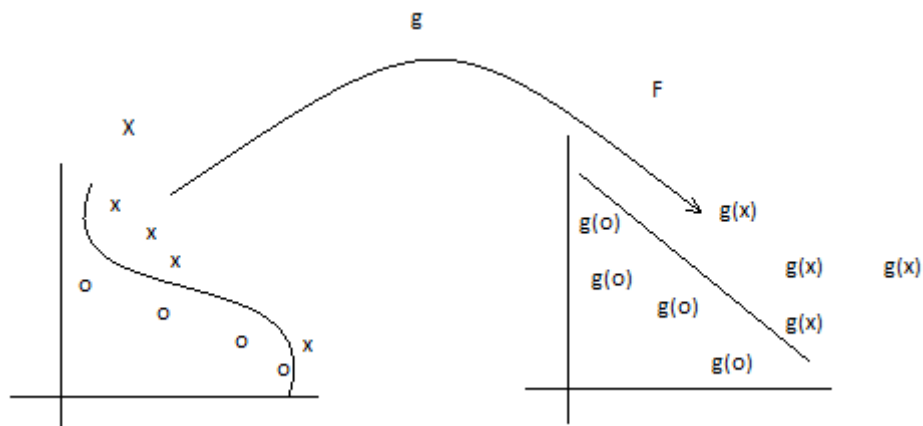


Figura 9 Espacio de funciones inducido por el kernel

Fuente: (Shi, Stumtner, Hao, & Quirchmayr, 2007)

$$F(x) = \sum_{i=1}^l a_i g_i(x) + b \quad (13)$$

$$F(x) = \sum_{i \in SV} y_i \alpha_i^* [g_i(x_i) \cdot g_i(x)] + b \quad (14)$$

Por tanto el *kernel* viene a ser una función K denotada por la ecuación (15), tal que para todo $(x, x_i \in X)$, siendo x_i la variable de la dimensión adicional.

$$K(x, x_i) = (g(x) \cdot g(x_i)) \quad (15)$$

Entre las funciones de Kernel se pueden destacar las siguientes:

- Polinomial
- Función de base radial
- Kernels sigmoideos

Los beneficios de utilizar *Kernel* en SVR, es que hace posible trabajar con muchas dimensiones, aunque estas dimensiones estén implícitas matemáticamente dentro del modelo o computacionalmente si es un software con el que se trabaja. Para el presente trabajo se hará uso del *kernel Polinomial* expresado en la ecuación (15).

2.4.1. Validación de parámetros

Para la validación de parámetros a usar en un modelo estadístico existen dos métodos que representan mejores resultados: validación cruzada y validación *bootstrap*, en los cuales resultados experimentales han demostrado que *bootstrap* puede ser el mejor método para validación, ya que ha demostrado una mejor respuesta de varianza (Kohavi, 1995) pero *bootstrap* utiliza un coste computacional elevado, por lo que en el presente proyecto se dará principal interés a validación cruzada por el número de datos obtenidos en los experimentos.

Dentro de validación cruzada existen métodos para la selección de datos que se utilizarán en el modelo para entrenamiento o para validación y así evitar resultados

sobre optimistas. Los métodos de validación que existen son (López Hernández, 2016):

- *K-fold cross validation*: Para este método se divide todo el conjunto de datos en dos bloques: bloque de entrenamiento y bloque de validación. Entonces si por ejemplo el conjunto de datos es de 8000 muestras, y se decide tomar para el bloque de entrenamiento 2000 datos restarían 6000 datos para el bloque de validación, por lo que k sería igual a 4 así $k = \frac{8000}{2000} = 4$, contemplando un escenario de 4-fold haciendo referencia a la figura 10.

Las 2000 muestras para entrenamiento van rotando de 2000 en 2000 en cada iteración hasta cumplir con todo el conjunto de datos de 8000 muestras.

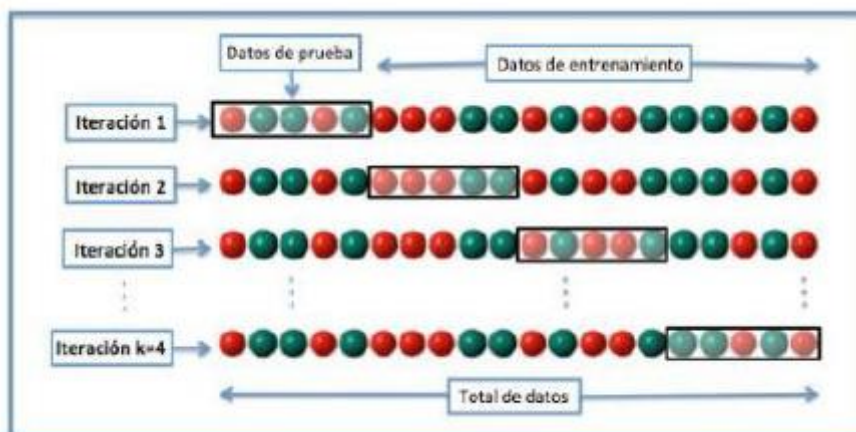


Figura 10 Método de k-fold cross validation

Fuente: (López Hernández, 2016)

- *Random cross validation*: Este método es similar al método anterior diferenciándose que para el bloque de datos de entrenamiento y validación se toma valores aleatorios además de no tener un número de iteraciones fijo tal como se observa en la figura 11.

Por otro lado, este método es menos eficiente en relación al anterior puesto que al ser los bloques de datos aleatorios las muestras de prueba pueden repetirse así como pueden existir muestras que quedan sin evaluar.

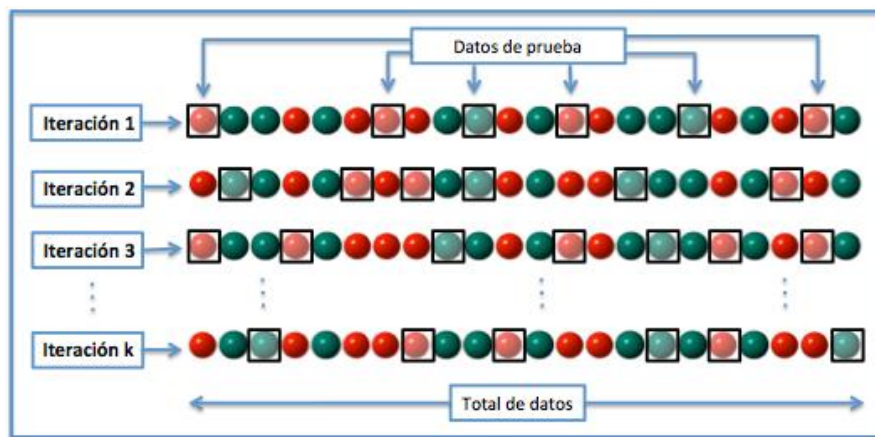


Figura 11 Método de random cross validation

Fuente: (López Hernández, 2016)

- *Leave one out cross validation*: En este método se utiliza una sola muestra para entrenamiento y las muestras restantes utilizadas para validación, obteniendo resultados muy buenos pero también con costos computacionales muy elevados, esto debido al número elevado de iteraciones, tantas como N muestras se tenga en el conjunto de datos. La figura 12 nos da una mejor percepción del número de iteraciones que se podrían tener hasta llegar al número N de muestras.

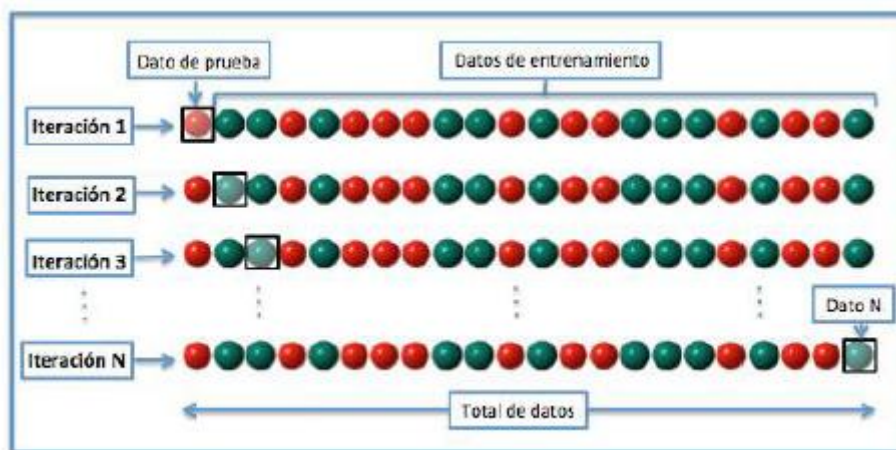


Figura 12 Método de leave one out cross validation

Fuente: (López Hernández, 2016)

CAPÍTULO III

ANÁLISIS DE DATOS PRECEDENTES

3.1.Introducción

En el presente capítulo se analizará los datos obtenidos en el proyecto de titulación realizado en la Universidad de las Fuerzas Armadas – ESPE (Bermeo & Guamán, 2017), enfatizando en primera instancia en los datos extraídos a través del hardware desarrollado que se conoce como la plantilla de presión plantar, puesto que es fundamental entender el funcionamiento del hardware junto con sus componentes para poder interpretar los valores extraídos del mismo, y posteriormente diseñar un nuevo protocolo de pruebas para la obtención de nuevos datos.

En segunda instancia se realizará el estudio del comportamiento de los sensores incluidos en la plantilla, al comprender de mejor manera la respuesta de los sensores se podrá entender los modelos obtenidos en el proyecto precedente y detectar las falencias en el modelo de calibración.

Una vez estudiada la respuesta de los sensores, se realizará una comparativa entre modelos de calibración univariante y modelos de calibración multivariable con los datos obtenidos en el protocolo de experimentos del proyecto precedente, esperando obtener mejores resultados en cuanto a parámetros de validación (*RMSE* y R^2).

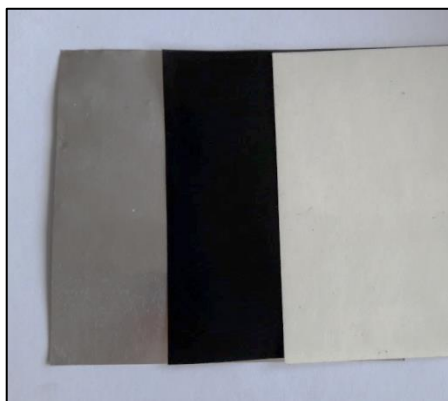
3.2.Plantilla de presión plantar precedente

La plantilla de presión plantar precedente fue elaborada con tres materiales principalmente, los cuales se citan en la tabla 2 y que además se puede observar en la figura 13 (Bermeo & Guamán, 2017). El objetivo de la plantilla es tener varios sensores distribuidos durante toda el área del pie, de manera que se pueda mapear la fuerza de la pisada del pie de una persona.

Tabla 2**Materiales de elaboración de plantilla precedente**

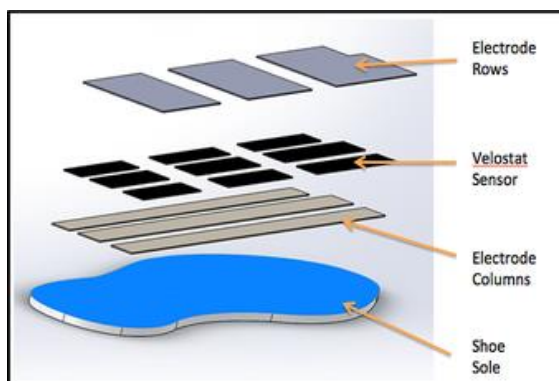
Materiales
Velostat
Mica de protección
Lámina de cobre

Fuente: (Bermeo & Guamán, 2017)

**Figura 13 Materiales plantilla de presión**

Fuente: (Bermeo & Guamán, 2017)

La distribución de los materiales por capas se muestra en la figura 14, la distribución está ubicada de manera que la lámina *velostat* quede introducida en medio de la cinta de cobre conductora y recubierta por la mica de protección, de este modo la plantilla puede resistir fuerzas de hasta 1200 *N* sin que se pueda alterar sus propiedades físicas (Bermeo & Guamán, 2017).

**Figura 14 Distribución de capas en la construcción del sensor**

Fuente: (Bermeo & Guamán, 2017)

Los sensores distribuidos en la plantilla construida en base a los tres materiales mencionados, toman el principio de funcionamiento de resistencias variables, es decir forman resistencias sensibles a la fuerza FSR, para poder variar la resistividad de la lámina velostat es necesario aplicar una fuerza perpendicular en cada sensor, o distribuir la presión de fuerza por áreas, para poder entender de mejor manera la construcción de una resistencia FSR se puede observar la figura 15,

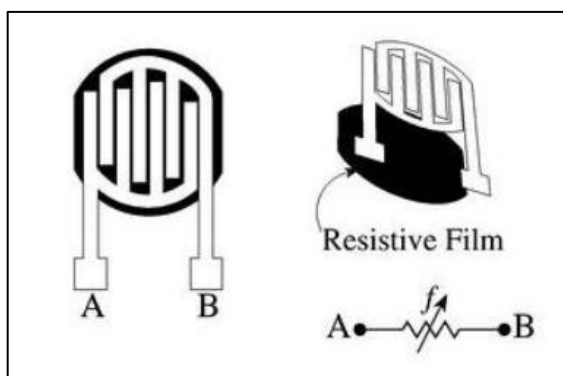


Figura 15 Estructura de un FSR

Fuente: (Bermeo & Guamán, 2017)

La plantilla mostrada en la figura 16 fue diseñada de manera que en el área de la plantilla se pueda contener 63 sensores distribuidos en filas y columnas, los 63 sensores adoptan la forma geométrica de un cuadrado de dimensiones $10 \times 10 \text{ mm}$ y separados entre ellos 4 mm aproximadamente distribuidos en forma matricial, principalmente está diseñada para que se aplique una componente de fuerza vertical soportando el peso corporal, logrando mapear la fuerza de la pisada del pie en toda el área de la plantilla (Bermeo & Guamán, 2017), al estar varios sensores distribuidos se puede discriminar también por áreas la presión plantar, lo cual es muy útil para casos puntuales en futuros trabajos.

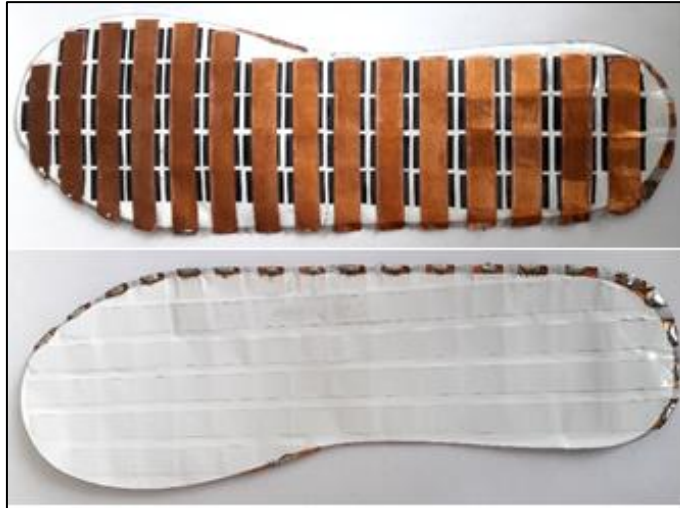


Figura 16 Prototipo de plantilla para medir la fuerza de la pisada

Fuente: (Bermeo & Guamán, 2017)

3.3.Respuesta no lineal de sensores piezoresistivos

Al realizar la extracción de datos de cada uno de los sensores piezoresistivos ejerciendo una fuerza puntual creciente F en un intervalo de tiempo t , se observó que la respuesta de voltaje del sensor ante la fuerza aplicada tiende a ser no lineal, representada por la figura 17. La respuesta del sensor está representada por dos curvas debido a que se aplica una fuerza creciente hasta un pico máximo (alrededor de 1200 N), luego de eso la fuerza decrece nuevamente hasta llegar a cero.

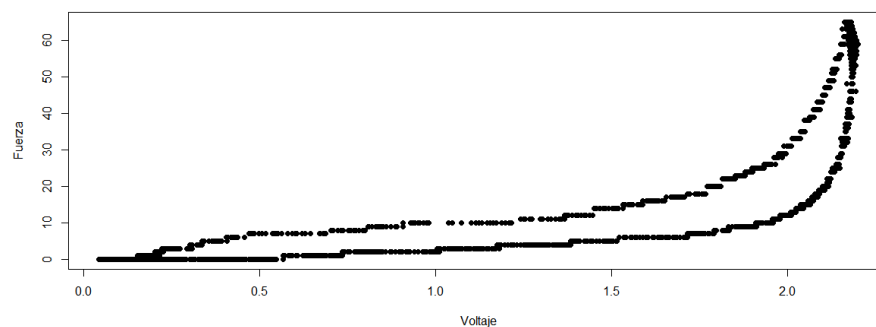


Figura 17 Respuesta del sensor fuerza vs voltaje

Es fundamental entender que al tener una respuesta no lineal de los sensores se aplicará técnicas para generar modelos de regresión no lineales. Al realizar pruebas en cada sensor y posteriormente generar un modelo de entrenamiento se va a tener que el modelo de validación será menos confiable puesto que los sensores no son

réplicas exactas existiendo variabilidad en la respuesta de un sensor con respecto a otro es decir la respuesta de los sensores no es reproducible. Entonces el modelo de validación muy difícilmente podrá predecir datos de “otros” sensores por lo que se ve en la necesidad de generar un modelo de entrenamiento que conozca los valores de todos los sensores para que el modelo de validación pueda predecir o estimar la respuesta de los 63 sensores.

3.4. Correlación entre sensores adyacentes

Al obtener los resultados de cada uno de los sensores adicionalmente del problema de no linealidad se observó que los sensores adyacentes al sensor de pruebas x_n , reaccionaban a la fuerza aplicada en x_n , en la figura 18 se puede observar por ejemplo que el sensor de pruebas es el sensor x_5 y los sensores adyacentes x_1 , x_2 , x_3 y x_4 , responden indirectamente, esto se debe propiamente a la construcción de la plantilla ya que todos los sensores comparten la misma lámina conductora de cobre, así como la lámina *velostat*, es decir toman el principio de una conexión de elementos en serie, por lo que si la resistencia de un sensor varía notablemente, los sensores adyacentes también variarían su resistividad.

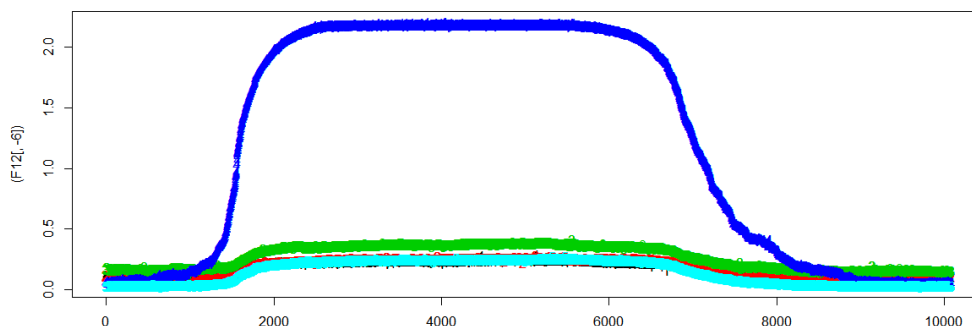


Figura 18 Correlación entre sensores adyacentes

Por el principio de conexión en serie existe una correlación entre sensores, puesto que los sensores están distribuidos en filas y columnas todos relacionados directamente con el voltaje de alimentación y con el material de construcción de la plantilla, los fabricantes (Bermeo & Guamán, 2017) describen el proceso de funcionamiento de la plantilla con sensores como un símil a un teclado matricial.

El problema de correlación no fue tomado en cuenta para generar el modelo de calibración de los sensores, influyendo en los resultados finales con una media de $R^2 = 87.27\%$ y $RMSE = 14,62\%$. Por las cifras obtenidas en los parámetros de

validación se tiene por ejemplo que una fuerza de $800N$ aplicada puntualmente en un sensor los sensores interpreten $683.04N$ (con error del $14,62\%$), entonces es necesario generar un modelo de regresión que tome en cuenta los valores que responden a la fuerza aplicada en un sensor determinado x_n y que mejore el R^2 y $RMSE$, de esta manera se puede obtener un modelo de regresión no lineal y un modelo de calibración más apegado a los valores reales aplicados ($error < 10\%$) involucrando a todos los valores de los sensores en el entrenamiento del modelo, y mejorar la apreciación de los resultados de los sensores en base a R^2 y $RMSE$.

3.5. Análisis de resultados con calibración univariante

En (Bermeo & Guamán, 2017) se trabajó con escala semilogarítmica representado por la figura 19(a) para tratar de solventar el problema de no linealidad del sistema pues como se observa en la figura 19(b) en la escala real es una respuesta no lineal. Adicionalmente se aplicó una regresión no lineal en cada una de las respuestas de los sensores obteniendo los resultados representados por la figura 20, donde se puede apreciar que los valores predichos no se ajustan a la distribución de valores reales, se dice que no se ajustan porque la distribución de valores predichos es totalmente distinta a la distribución de valores reales justificada con $RMSE=22.63\%$ obtenido en el sensor de la comuna 4 de fila 1.

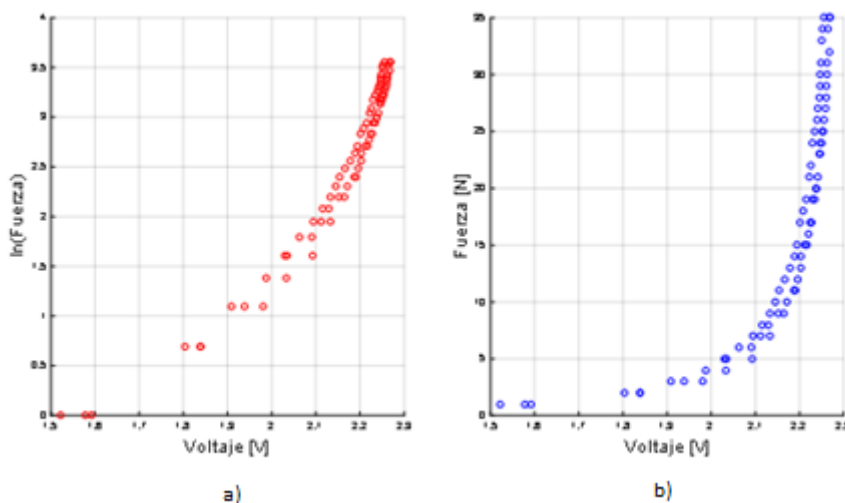


Figura 19 Curva de comportamiento del sensor de la columna 4 de fila 1 (x_4)

a) Escala semilogarítmica, b) Escala real

Fuente: (Bermeo & Guamán, 2017)

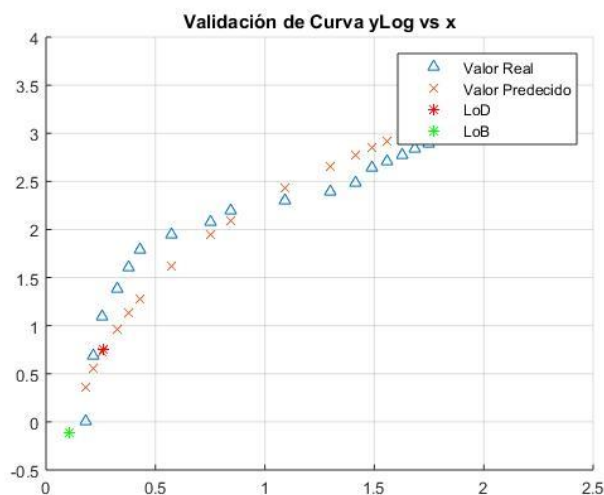


Figura 20 Regresión univariante - sensor C4

Fuente: (Bermeo & Guamán, 2017)

Posteriormente, se realizó la calibración de cada sensor con la variable y' del modelo de regresión obtenido versus la variable y del modelo real, siendo y' la fuerza predicha del modelo de regresión y siendo y la fuerza real aplicada en los sensores, demostrado por la figura 21, donde aquí se puede observar también un problema de no linealidad con un ajuste de sensibilidad (R^2) de 88,37% en escala semilogarítmica, que se tratará de mejorar con calibración multivariante.

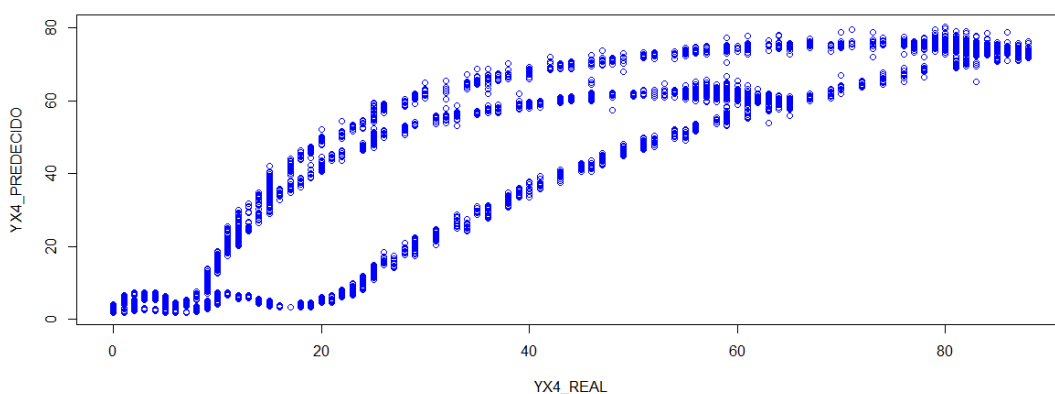


Figura 21 Modelo de calibración – sensor C4

Fuente: (Bermeo & Guamán, 2017)

3.6. Análisis de resultados con calibración multivariable

Durante el proceso de análisis de datos previos se generó un modelo de regresión multivariante con SVR, de esta manera se puede realizar una comparativa con los modelos precedentes y visualizar la mejora de linealidad en el modelo de calibración, ya que en el nuevo modelo multivariante se toma en cuenta la correlación entre sensores adyacentes.

En (Bermeo & Guamán, 2017) realizó un diseño de experimentos, en donde se ha realizado 3 experimentos por cada sensor de la plantilla con una fuerza progresiva de 0 a 1200 *N*. Para generar el modelo multivariante se tomó dos experimentos para entrenar el modelo multivariable y el experimento restante se utilizó para realizar la validación del modelo multivariable, de tal manera que los experimentos vendrían a ser la variable independiente x .

Los datos para el análisis multivariable de los modelos de experimentos precedentes fueron los datos de los experimentos obtenidos en la primera fila de sensores de la plantilla de presión plantar; en total se han realizado nueve experimentos distribuidos en los cinco sensores x_1, x_2, x_3, x_4, x_5 de la siguiente manera, los cinco sensores ubicados en la plantilla son visibles en la figura 22.

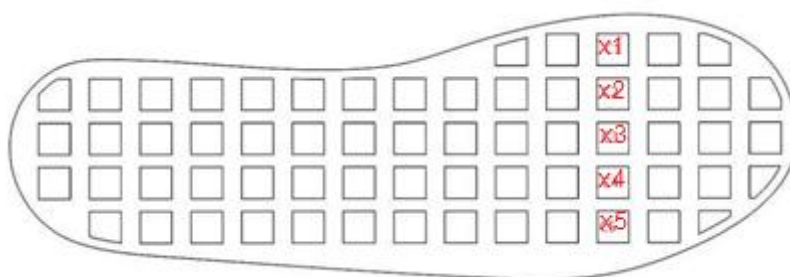


Figura 22 Sensores distribuidos en la plantilla de presión plantar

(Bermeo & Guamán, 2017)

Experimentos en el sensor x_4

Los tres primeros experimentos fueron realizados en el sensor x_4 , es decir que se aplicó fuerza en dicho sensor y por tanto se usará la información obtenida del sensor x_4 para ejemplificar el uso de modelos multivariantes, tomando en cuenta que los sensores adyacentes x_1, x_2, x_3, x_5 tienen un efecto de correlación al ejercer presión en el sensor x_4 .

La variable dependiente o regresora x del sensor de prueba x_4 para el modelo de entrenamiento se generará con los dos experimentos iniciales que contienen valores de los cinco sensores que han respondido a la fuerza aplicada en el sensor x_4 siendo visible en la ecuación (16).

$$y = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 \quad (16)$$

La variable independiente x para realizar el modelo de validación esta generado con el tercer experimento que también contiene los valores de los cinco sensores que han resultado a la fuerza ejercida en el sensor x_4 .

Los valores de fuerza aplicados en el sensor x_4 estarán representados como la variable dependiente y . Se ha utilizado la fuerza ejercida y de los dos primeros experimentos para entrenar el modelo, donde y también es visible en la ecuación (17). Mientras que para el modelo de validación del modelo se utilizará los valores de fuerza aplicados en el tercer experimento, que vendría a ser la variable y de validación.

Estos modelos han sido generados para los sensores x_2 y x_3 , de la misma manera se han realizado tres experimentos en cada uno de los sensores, dos pruebas para entrenar el modelo y la prueba restante para validación del modelo.

3.4.1. Construcción de matrices para entrenar y predecir el modelo

En la tabla 3 se puede observar cómo se generó la matriz de datos para poder generar el modelo de entrenamiento, así como la matriz de datos para la generación del modelo de validación. Además la tabla 3 muestra todo el conjunto de datos almacenado en Matriz $XCF1$, es decir los datos de los tres experimentos ($XC4, XC3, XC2$) realizados en el sensor x_4 aplicando una fuerza perpendicular creciente, donde en cada uno de los experimentos se almacenan los datos de la respuesta de los cinco sensores. Por tanto la matriz $XCF1$ representa los datos de las variables independientes o regresoras.

Tabla 3
Matriz XCF1

Datos	Número de experimentos	Descripción
XC4	Primer experimento	Primera respuesta de sensores
XC3	Segundo experimento	Segunda respuesta de sensores
XC2	Tercer experimento	Tercera respuesta de sensores

En la matriz de datos $YCF1$ visible en la tabla 4 se almacenó la fuerza aplicada en los tres experimentos realizados en el sensor x_4 , donde $YC4, YC3, YC2$ representa la fuerza aplicada en $XC4, XC3, XC2$ respectivamente, lo que significa que la fuerza aplicada corresponde a la variable dependiente de los modelos. Es necesario mencionar que para la generación de los modelos de entrenamiento y validación es fundamental que tanto para las variables independientes o regresoras x_n , como para las variables dependientes y_n , existan el mismo número de datos.

Tabla 4**Matriz YCF1**

Datos de fuerza	Descripción
YC4	Primera fuerza ejercida en el sensor x_4
YC3	Segunda fuerza ejercida en el sensor x_4
YC2	Tercera fuerza ejercida en el sensor x_4

Para la generación de modelos de entrenamiento, se tomará como datos de entrenamiento los dos primeros experimentos obtenidos ($XC4, XC3$) al ejercer presión ($YC4, YC3$) en el sensor x_4 como se especifica en la tabla 5 y tabla 6.

Tabla 5**Matriz XC4 para entrenamiento**

Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Descripción
x11	x21	x31	x41	x51	Primer experimento para entrenamiento
x12	x22	x32	x42	x52	Segundo experimento para entrenamiento

Tabla 6**Matriz YC4 para entrenamiento**

Datos para variable dependiente	Fuerzas ejercidas
Y41	Primera fuerza ejercida
Y42	Segunda fuerza ejercida

Finalmente para la generación del modelo de validación, se utilizará los datos obtenidos en el tercer experimento $XC2$ como datos de las variables, mientras que los datos utilizados para la variable dependiente está representado por $YC2$, para esto se ha generado las tablas 7 y tabla 8.

Tabla 7**Matriz XTF1 para validación**

Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Descripción
x13	x23	x33	x43	x53	Tercer experimento para predicción o validación

Tabla 8**Matriz YTF1 para validación**

Datos para variable dependiente	Fuerzas ejercidas
Y43	Tercera fuerza ejercida

3.4.2. Validación cruzada del modelo multivariante

El tipo de validación que se usará en el modelo multivariante es *k-fold cross validation* tomando un bloque de datos de entrenamiento y un bloque de datos de validación, el tamaño de cada bloque está definido según el valor que estimemos para k , para el presente caso se eligió un valor de $k = 3$ por lo que será necesario realizar 3 iteraciones y por consiguiente dividir el total de datos para 3 para poder cumplir con las 3 iteraciones, tomando en cuenta que el conjunto de datos de entrenamiento va rotando en cada iteración como se explicó en la figura 10.

La técnica de *machine learning* SVR tiene como parámetros de funcionamiento sigma σ y épsilon ϵ que son necesarios optimizar realizando validación cruzada, los valores que puede tomar sigma son: $0 < \sigma < 1$, mientras que los valores que puede

tomar ϵ son: $0 < \epsilon < 0.6$. A través de validación cruzada también se puede obtener el mejor conjunto de datos que minimice el error cuadrático medio *RMSE*.

Entonces el modelo multivariable se entrena con cada uno de los valores de los parámetros con los que se puede tomar σ y ϵ . En la tabla 9 se puede observar que para entrenar el modelo se realizó las validaciones con un ϵ igual a 0.1 que se incrementa hasta llegar a 0,5, mientras que σ toma valores desde 0.1 hasta llegar a 0.9, para los dos parámetros se realiza incrementos de 0.1, además se resalta el valor de σ igual a 0.6 que es el valor con el que mejor *RMSE* ha respondido el modelo entrenado.

Tabla 9

Optimización de parámetros con datos precedentes

$\epsilon=0.1$		$\epsilon=0.2$		$\epsilon=0.3$	
σ	RMSE	σ	RMSE	σ	RMSE
0.1	7.02	0.1	7.46	0.1	6.93
0.2	6.21	0.2	6.13	0.2	6.88
0.3	5.50	0.3	6.61	0.3	6.74
0.4	5.01	0.4	6.27	0.4	7.91
0.5	4.81	0.5	5.42	0.5	8.42
0.6	4.68	0.6	4.86	0.6	8.02
0.7	4.82	0.7	5.07	0.7	7.18
0.8	5.04	0.8	5.03	0.8	6.55
0.9	4.95	0.9	5.34	0.9	7.77
$\epsilon=0.4$		$\epsilon=0.5$			
σ	RMSE	σ	RMSE		
0.1	7.68	0.1	11.07		
0.2	8.73	0.2	6.89		
0.3	5.76	0.3	7.65		
0.4	8.32	0.4	6.34		
0.5	9.97	0.5	8.90		
0.6	10.39	0.6	11.98		
0.7	10.61	0.7	12.34		
0.8	10.86	0.8	12.64		
0.9	10.86	0.9	12.80		

Después de realizar la validación cruzada del modelo se determina que el mejor valor para ϵ y para σ en efecto son 0.1 y 0.6 respectivamente; por tanto con estos datos se procede a entrenar y a predecir el modelo multivariante de la Fila 1. Obteniendo un modelo de predicción con datos de entrenamiento y un modelo de predicción con datos de validación. En la figura 23 se visualiza la gráfica de calibración con datos de entrenamiento mientras que en la figura 24 se visualiza la gráfica de calibración con datos de validación; obteniendo un $RMSE$ de 5.18 % y 7.70% respectivamente. Entonces al comparar la calibración con modelo multivariable frente a la calibración con modelo univariante que ha resultado con un $RMSE$ de 22,63%, se puede diferenciar una mejora en la calibración de los sensores con la utilización de un modelo multivariable. En términos de ajuste si comparamos la figura 21 que viene a ser el modelo de calibración univariante con un $R^2 = 88,37\%$ (en escala semilogarítmica) con la figura 23 y 24 donde representan los modelos de calibración con modelos multivariables con un $R^2 = 93.41$ y 92.76% (en escala real) respectivamente, se puede concluir que en la calibración multivariable el término de ajuste o determinación mejora notablemente.

$$RMSE = 5.18\% \text{ y } R^2 = 93.41\%$$

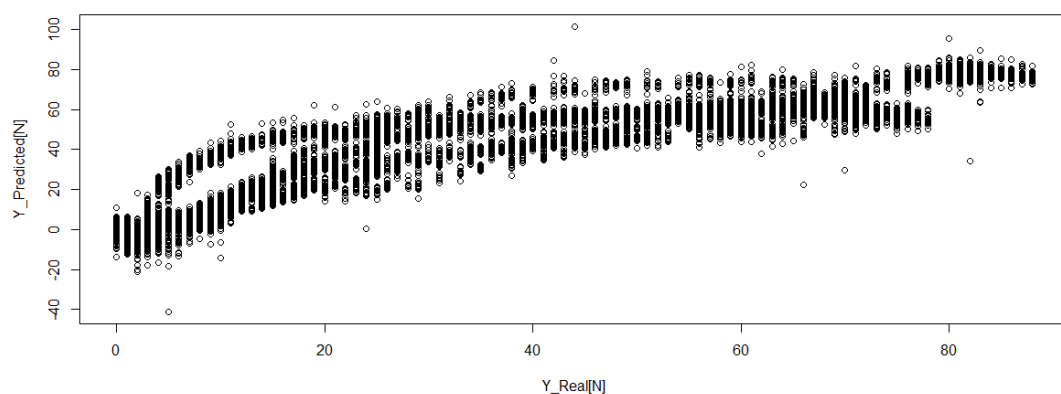


Figura 23 Modelo multivariable fila 1 fuerza predicha vs fuerza real con datos de entrenamiento

$$RMSE = 7.70\% \text{ y } R^2 = 92.76\%$$

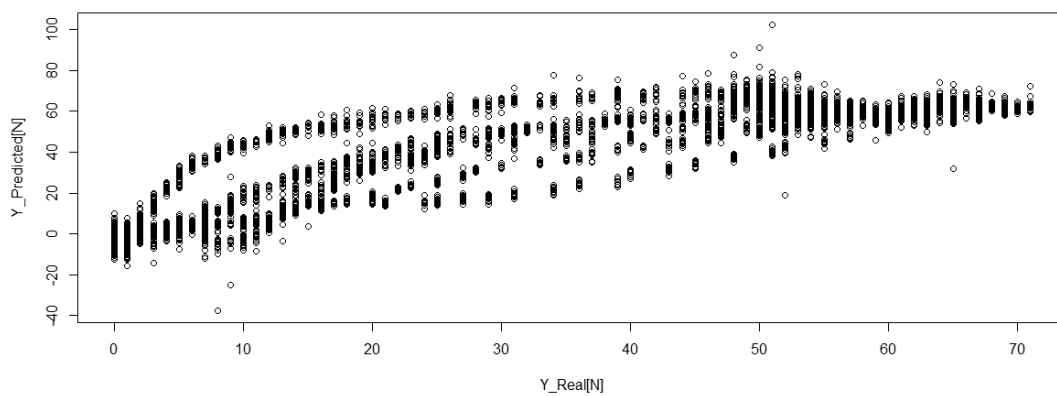


Figura 24 Modelo multivariable fila 1 fuerza predicha versus fuerza real con datos de validación

CAPÍTULO IV

MODELOS MULTIVARIABLES DE ÁREAS CON MAYOR PRESIÓN PLANTAR

4.1. Introducción

En el presente capítulo se presentará el procedimiento que se realizó para la obtención de nuevos datos, comenzando por la delimitación de las áreas en donde al realizar una pisada se obtiene umbrales de presión elevados. En consecuencia es necesario diseñar un protocolo de experimentos para ejercer presión en toda el área delimitada.

Para el diseño de experimentos se establecen diseños de prototipos para poder imponer presión uniforme distribuida en el área de presión deseada. Una vez ejercido presión es necesario procesar la información que arrojan los sensores, estableciendo un programa de adquisición de datos junto con un hardware que permita interpretar la información de los 63 sensores para luego poder construir modelos multivariantes por cada área delimitada.

En la construcción de modelos multivariantes se procesará la información de los sensores y se irá estableciendo los bloques de datos que servirán para generar los modelos de entrenamiento con la técnica de Machine Learning – SVR, los modelos de entrenamiento se generarán para las nueve áreas delimitadas. Durante la generación de los modelos de entrenamiento se procesará toda la información de los sensores y se usará la técnica SVR en el software estadístico Rstudio.

4.2. Delimitación de áreas con mayor presión plantar

En base a investigaciones previas (Pérez García, 2004) se adopta la distribución de nueve áreas con mayor presión plantar para la plantilla de (Bermeo & Guamán, 2017) distribuidas como se observa en la figura 25. Necesariamente se debe hacer énfasis en la distribución de presión plantar por áreas puesto que en estas áreas están determinados los huesos que indican una afectación puntual por umbrales de presión elevados en el pie; siendo los 9 huesos los siguientes:

- Hueso Hallux
- Huesos dedos de pie menores
- 1 Cabeza metatarsiana
- 2 Cabeza metatarsiana
- 3 Cabeza metatarsiana
- 4 Cabeza metatarsiana
- 5 Cabeza metatarsiana
- Hueso de Pie medio
- Hueso Heel

Con esta distribución se podrá calibrar la fuerza aplicada en los sensores de cada área de la plantilla, para el posterior análisis de la curva de calibración, tal que se pueda predecir con más exactitud la fuerza aplicada que se distribuye por toda el área de presión.

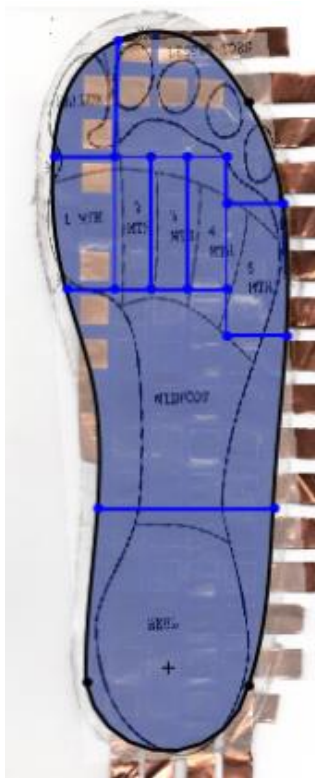


Figura 25 Plantilla con 9 áreas de presión plantar

4.3. Diseño de Protocolo de Experimentos

La generación de modelos multivariables por cada área de presión resuelve el problema de calibración del sistema al tomar en cuenta la respuesta de sensores adyacentes (correlación) en la generación de cada modelo. Para cada área se distribuye un cierto número de sensores que ayudarán a generar un modelo multivariable en base a cinco experimentos por cada área delimitada. Los experimentos se los realizará ejerciendo una fuerza puntual en bloques diseñados y contruidos en impresión 3D con la forma de cada área requerida capaces de soportar hasta una fuerza de 1200 *Newtons*. Se diseñó los bloques con el fin de generar los modelos de calibración con las mismas características al aplicar la misma fuerza en todos los sensores; además de simular la presión que ejercerá un pie humano en futuros trabajos de investigación realizando pruebas de validación del prototipo. Los bloques diseñados para cada área se muestran a continuación en la figuras 26, 27, 28, 29, 30, 31, 32, 33, 34; que en total son nueve bloques representado cada bloque el área correspondiente a cada hueso con umbrales de mayor presión plantar .

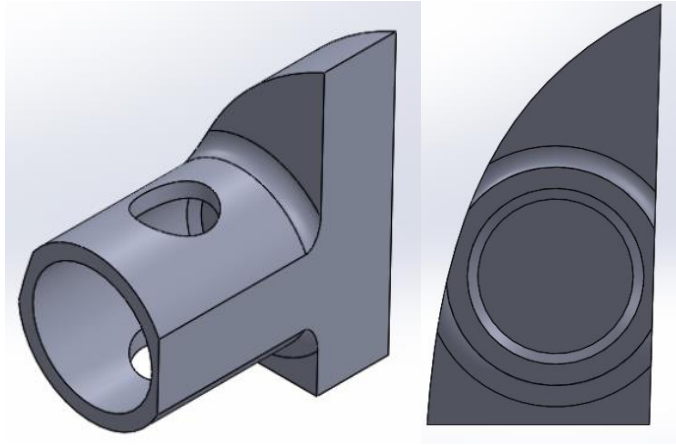


Figura 26 Bloque hueso hallux

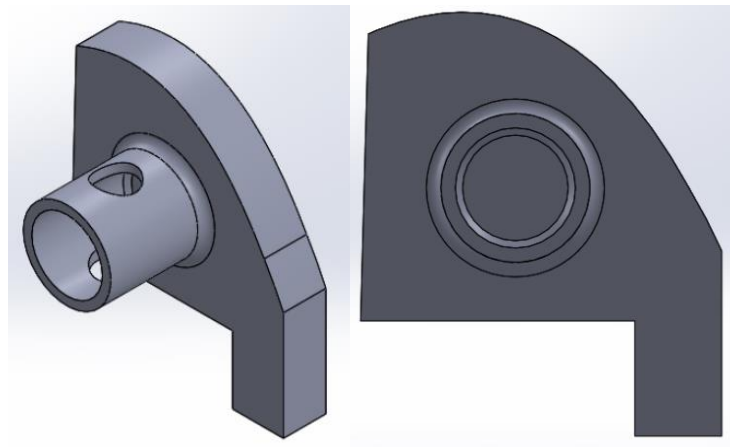


Figura 27 Bloque huesos dedos de pie menores

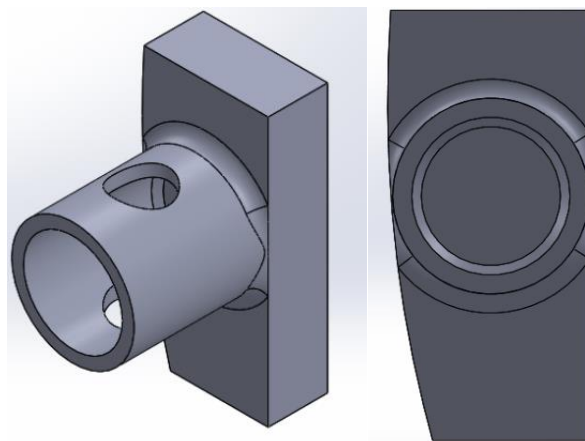


Figura 28 Bloque 1 cabeza metatarsiana 1

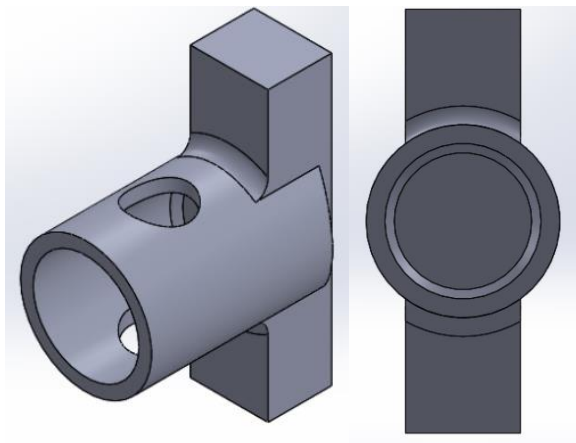


Figura 29 Bloque 2 cabeza metatarsiana 2

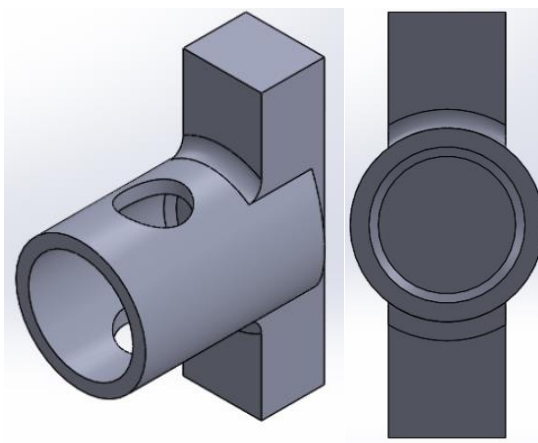


Figura 30 Bloque 3 cabeza metatarsiana 3

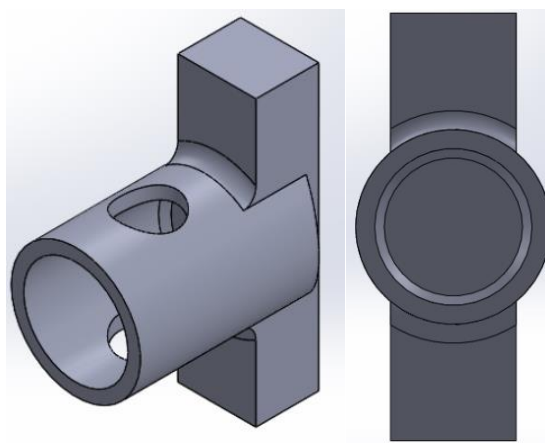


Figura 31 Bloque 4 cabeza metatarsiana 4

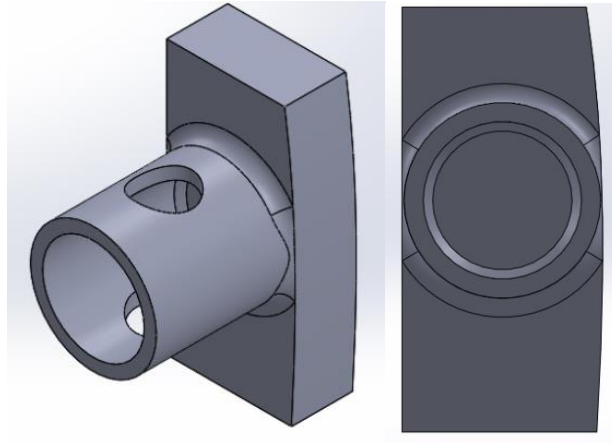


Figura 32 Bloque 5 cabeza metatarsiana 5

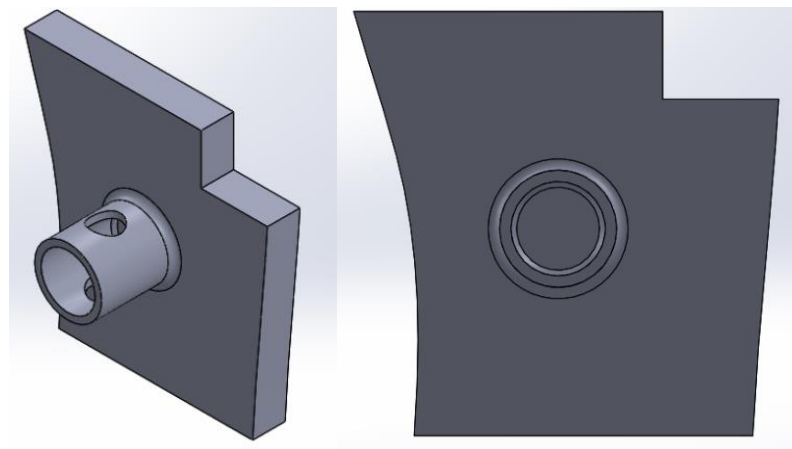


Figura 33 Bloque hueso pie medio

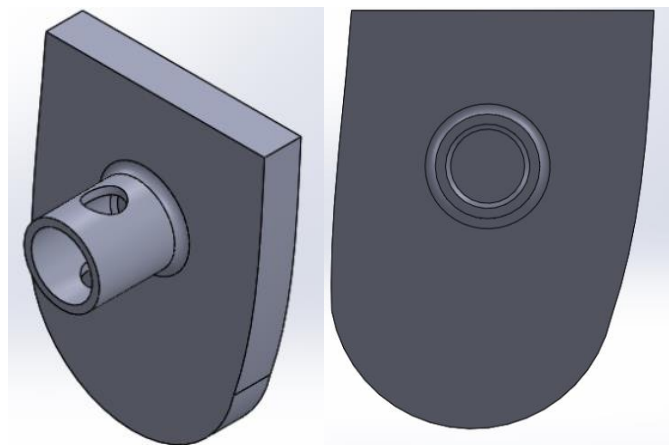


Figura 34 Bloque hueso heel

En la figura 35 se puede observar los prototipos de los bloques ya impresos en 3D. La toma de datos se la realizó en el laboratorio de mecánica de los materiales de la Universidad de las Fuerzas Armadas ESPE con la utilización de una máquina que

ejerce una fuerza puntual vertical, controlada a través de una interfaz *touch screen* como se puede ver en la figura 36.



Figura 35 Nueve Prototipos impresos en 3D

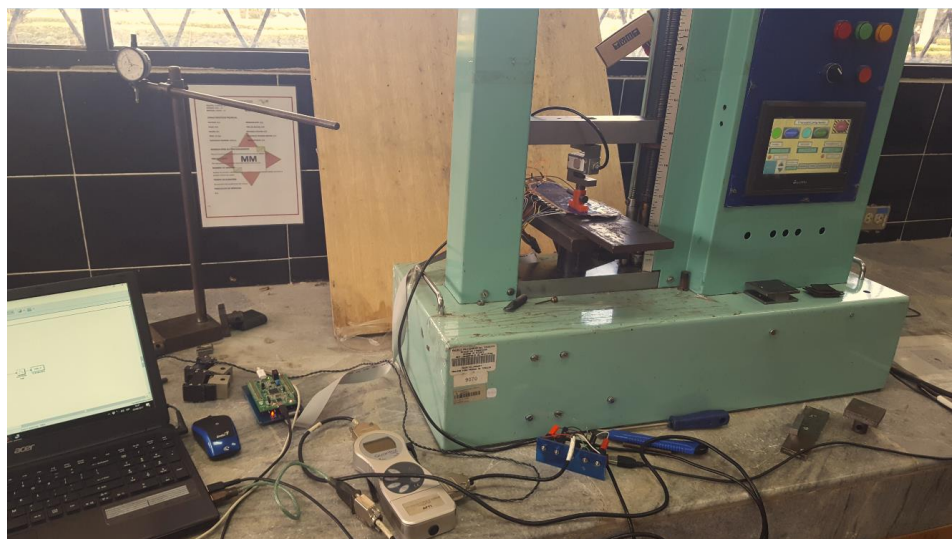


Figura 36 Pruebas finales con máquina de presión puntual

La adquisición de información que otorgan los sensores es realizada a través de una tarjeta Printed Circuit Board (PCB) diseñada en el software Proteus 8 Profesional que además el diseño es basado en el diseño realizado por Bermeo y Guamán (2016), el diseño final obtenido se lo puede ver en la figura 37. Una vez realizado el diseño final es necesario realizar un proceso químico para poder obtener la tarjeta final física que será parte del hardware a utilizarse en las pruebas finales, la tarjeta PCB final obtenida se la puede observar en la figura 38. Adicional el diseño de la tarjeta PCB cuenta con buses de datos para poder trabajar con la tarjeta de procesamiento STM32 lo cual resulta muy útil para poder ocupar las herramientas de

la tarjeta STM32, como por ejemplo la conexión transmisor-receptor asíncrono universal expresado en sus siglas en inglés por UART.

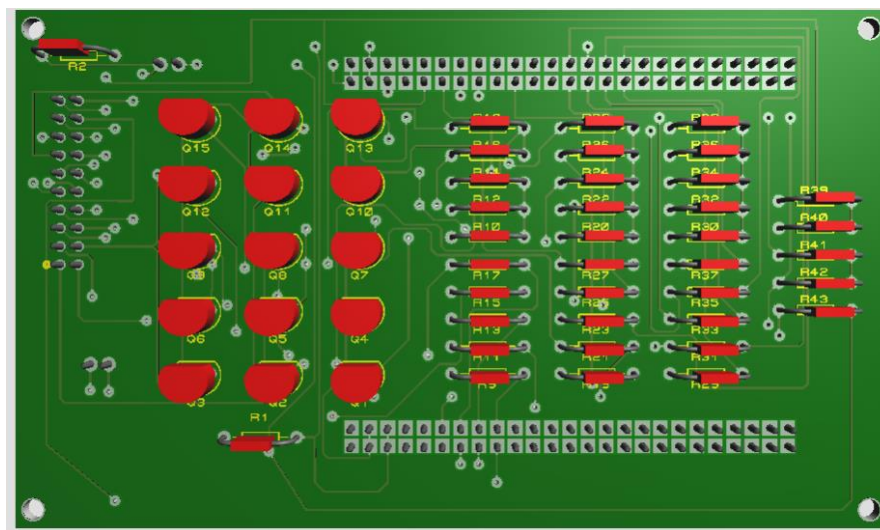


Figura 37 Diseño de Hardware para interpretación de datos de sensores



Figura 38 Hardware terminado para interpretación de datos de sensores

La conexión UART es utilizada para realizar la recepción de información de la tarjeta PCB en una PC por comunicación serial a una velocidad de 115200 baudios por segundo, que en este caso la información es los datos otorgados por los sensores de la plantilla. En la figura 39 se puede observar el esquema de recepción de la información de la tarjeta PCB en una PC.

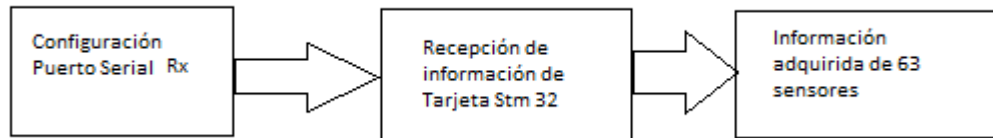


Figura 39 Recepción de datos de PCB por puerto de comunicación COM 7

Mientras que en la figura 40 se puede observar la respuesta de los sensores a través de transmisión serial aplicando un presión puntual en el área heel.

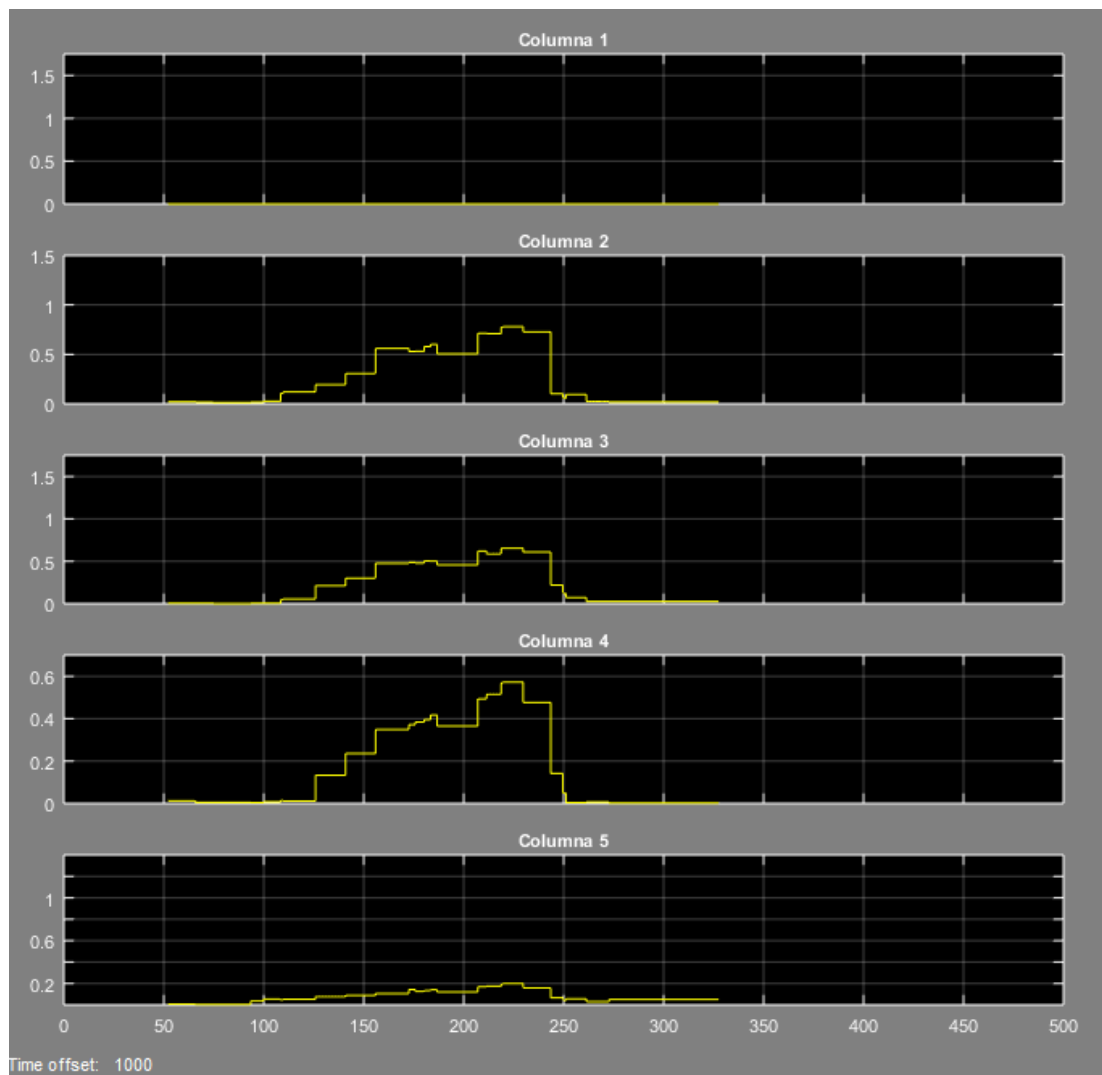


Figura 40 Respuesta de sensores por comunicación serial alámbrica

Una vez hecha la adquisición de información de los sensores a través de la tarjeta PCB es necesario procesar esta información para poder generar los modelos deseados de regresión. Para el procesamiento de información se utilizó el *software* estadístico Rstudio que adicional cuenta con varias técnicas de *Machine Learning*, entre las cuales en el presente proyecto se utilizó la técnica *Support Vector Regression*.

4.4. Generación de Modelos de Regresión Multivariantes

Previo a la generación de los modelos de regresión multivariantes por cada área de presión plantar se procesa la información adquirida de los sensores para determinar los datos que corresponden a cada sensor. La información obtenida se almacenó en una matriz de datos de 10000 filas por 63 columnas, donde las filas representan el número de muestras de la respuesta de los sensores mientras que las columnas representan el número de sensores en la plantilla. En la figura 41(b) se puede observar cómo se distribuyeron los sensores en cada área de presión. En consecuencia el número de sensores por cada área se puede observar en la tabla 14.

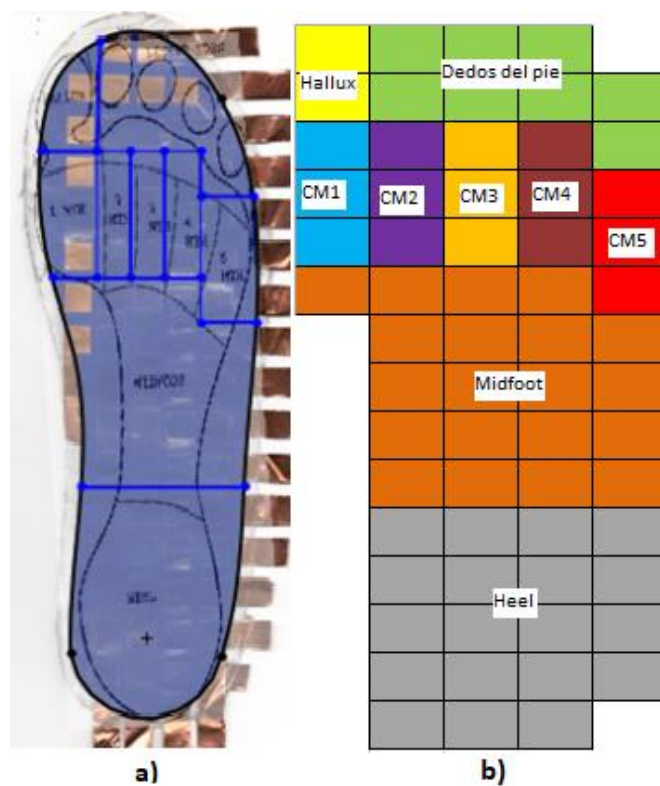


Figura 41 Sensores distribuidos en cada área con mayor presión plantar

a) Plantilla con 9 áreas de presión plantar b) Número de sensores por áreas

Al ejercer presión en cada área delimitada se obtiene una respuesta inmediata de todo el grupo de sensores correspondientes a cada área. Esta respuesta es interpretada como datos de los sensores y almacenados en una variable independiente x_n por cada sensor. En total se tendrían 63 variables independientes en donde se ira ordenando cada grupo de variables según la delimitación de cada área. Cada área viene a ser un modelo de regresión multivariable que en total serían nueve modelos de regresión. En la tabla 14 se puede observar las variables regresoras que se tiene por cada modelo de regresión.

Tabla 10

Número de sensores distribuidos en cada área o modelo de regresión

Área o modelo de regresión	Número de sensores	Número de variables independientes (x_n)
Hallux	2 sensores	2 variables
Dedos del Pie	8 sensores	8 variables
Cabeza Metatarsiana 5	3 sensores	3 variables
Cabeza Metatarsiana 4	3 sensores	3 variables
Cabeza Metatarsiana 3	3 sensores	3 variables
Cabeza Metatarsiana 2	3 sensores	3 variables
Cabeza Metatarsiana 1	3 sensores	3 variables
Midfoot	20 sensores	20 variables
Heel	19 sensores	19 variables
Total	63 sensores	63 variables

4.5. Entrenamiento de Modelos de Regresión Multivariables

De las cinco pruebas obtenidas en el laboratorio de mecánica de los materiales, se seleccionaron tres pruebas para la generación de los modelos de entrenamiento de regresión multivariables. En cada modelo de regresión se aplicó *k-fold cross validation*, con un $k=10$, es decir que se hizo diez iteraciones por modelo. En cada iteración se rota continuamente el bloque de datos para entrenamiento, siendo el bloque de datos del mismo tamaño en cada iteración para poder cumplir con las diez iteraciones del total del conjunto de datos. En la figura 42 se puede observar las diez iteraciones que se realizó rotando el bloque de datos de entrenamiento.

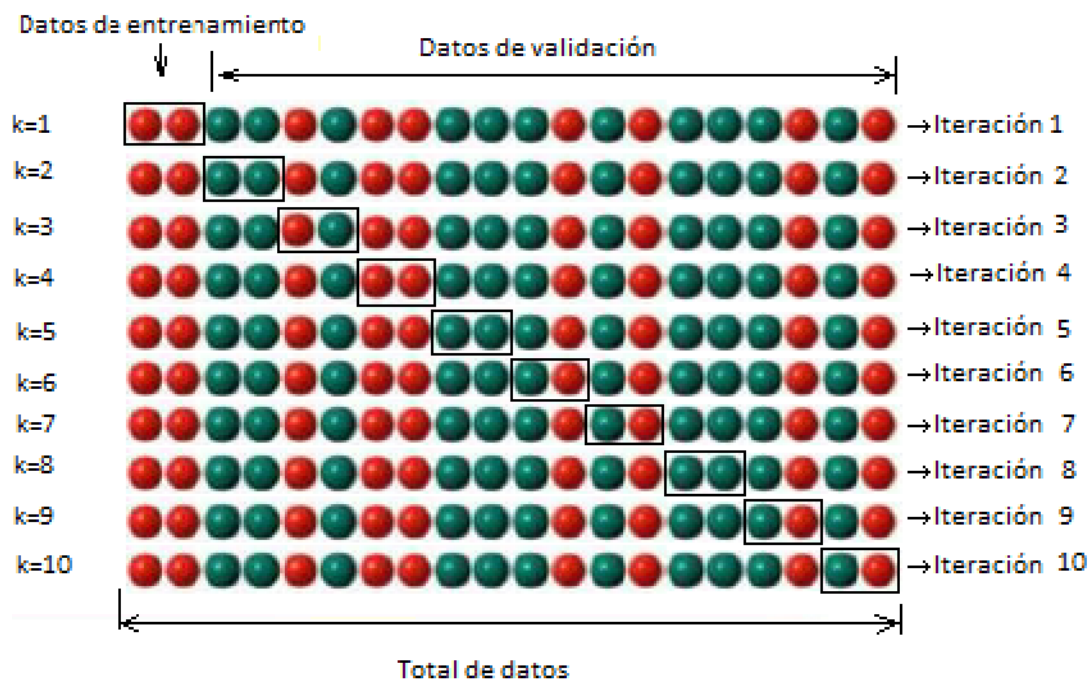


Figura 42 K-fold cross validation con k=10

K-fold cross validation es utilizado para optimizar los parámetros que serán indispensables en el modelo de regresión para poder determinar el mejor error cuadrático medio (RMSECV), en el siguiente apartado se hará la optimización de parámetros en base a *k-fold cross validation*.

4.6. Optimización de Parámetros con Cross Validation

En la optimización de parámetros se conserva el valor de ϵ igual a 0.1 pues se toma en consideración las recomendaciones de los autores del algoritmo de 'kernlab' (Karatzoglou, Smola, & Hornik, 2016) que viene dado por *default*, mientras que σ está dado por valores desde 0.1 hasta 0.9. En consecuencia procedemos con la optimación usando *k-fold cross validation* con $k=10$ visible en la figura 41. En las tablas desde la 11 hasta la 19 se encontrarán los datos correspondientes a la optimización de parámetros de los modelos de regresión que representarán las diferentes áreas de presión. En las tablas desde la 11 hasta la 19 se incluyó además el valor de error cuadrático medio de validación cruzada (RMSECV) que corresponde a la media de todos los valores de RMSE obtenidos en cada iteración k , este parámetro es usado como guía para obtener el valor más "óptimo" del parámetro σ . En

conclusión los parámetros ϵ y σ para entrenar el modelo *heel* con la técnica *support vector regression* están resaltados de color amarillo en las tablas mencionadas

Tabla 11

Optimización con k-fold cross validation para modelo hallux

$\epsilon=0.1 \sigma=0.1$		$\epsilon=0.1 \sigma=0.2$		$\epsilon=0.1 \sigma=0.3$		$\epsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	27.45	1	27.45	1	27.45	1	27.45
2	7.40	2	8.94	2	9.99	2	10.80
3	20.35	3	20.37	3	20.17	3	19.99
4	6.66	4	6.17	4	5.89	4	5.60
5	7.00	5	11.28	5	15.60	5	18.82
6	15.49	6	20.29	6	17.61	6	14.28
7	5.21	7	8.05	7	10.23	7	12.40
8	18.04	8	21.13	8	21.39	8	21.48
9	14.63	9	13.93	9	14.32	9	14.87
10	28.48	10	24.95	10	22.60	10	21.85
RMSECV	15.071	RMSECV	16.256	RMSECV	16.525	RMSECV	16.754
$\epsilon=0.1 \sigma=0.5$		$\epsilon=0.1 \sigma=0.6$		$\epsilon=0.1 \sigma=0.7$		$\epsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	27.45	1	27.45	1	27.45	1	27.45
2	11.52	2	12.39	2	12.82	2	12.91
3	19.96	3	19.95	3	19.92	3	19.94
4	5.42	4	5.38	4	5.37	4	5.30
5	20.78	5	21.93	5	22.80	5	23.15
6	13.24	6	13.28	6	13.60	6	13.92
7	13.39	7	13.31	7	13.29	7	12.84
8	21.53	8	21.84	8	22.08	8	22.24
9	15.15	9	15.59	9	16.15	9	16.83
10	21.67	10	21.49	10	21.77	10	22.11
RMSECV	17.011	RMSECV	17.261	RMSECV	17.525	RMSECV	17.669
$\epsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	27.45						
2	13.23						
3	19.98						
4	5.28						
5	23.34						
6	14.22						
7	12.36						
8	22.35						
9	16.73						
10	22.35						
RMSECV	17.729						

Tabla 12

Optimización con k-fold cross validation para modelo dedos del pie

$\varepsilon=0.1 \sigma=0.1$		$\varepsilon=0.1 \sigma=0.2$		$\varepsilon=0.1 \sigma=0.3$		$\varepsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	24.18	1	24.48	1	24.46	1	24.48
2	12.07	2	13.47	2	14.09	2	14.46
3	27.00	3	23.92	3	22.87	3	22.39
4	6.51	4	6.61	4	6.05	4	5.46
5	17.43	5	16.41	5	17.03	5	17.82
6	13.16	6	22.31	6	22.13	6	19.98
7	5.57	7	4.18	7	3.67	7	3.52
8	21.58	8	22.37	8	22.76	8	22.90
9	6.97	9	8.85	9	10.59	9	11.84
10	26.96	10	26.56	10	26.20	10	26.03
RMSECV	16.143	RMSECV	16.916	RMSECV	16.985	RMSECV	16.888
$\varepsilon=0.1 \sigma=0.5$		$\varepsilon=0.1 \sigma=0.6$		$\varepsilon=0.1 \sigma=0.7$		$\varepsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	24.48	1	24.47	1	24.50	1	24.52
2	14.72	2	14.91	2	15.05	2	15.17
3	22.12	3	21.85	3	21.64	3	21.55
4	4.96	4	4.58	4	4.29	4	4.09
5	18.50	5	19.21	5	19.85	5	20.43
6	18.86	6	18.19	6	17.68	6	17.50
7	3.56	7	3.63	7	3.70	7	3.77
8	22.97	8	23.02	8	22.95	8	22.83
9	12.71	9	13.07	9	12.86	9	12.82
10	25.88	10	25.81	10	25.79	10	25.79
RMSECV	16.876	RMSECV	16.874	RMSECV	16.831	RMSECV	16.847
$\varepsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	24.54						
2	15.27						
3	21.50						
4	3.97						
5	20.97						
6	17.38						
7	3.85						
8	22.81						
9	12.80						
10	25.81						
RMSECV	16.89						

Tabla 13

Optimización con k-fold cross validation para modelo CM1

$\varepsilon=0.1 \sigma=0.1$		$\varepsilon=0.1 \sigma=0.2$		$\varepsilon=0.1 \sigma=0.3$		$\varepsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	13.86	1	15.97	1	17.24	1	18.19
2	13.38	2	11.43	2	10.05	2	9.51
3	24.91	3	22.25	3	20.59	3	19.98
4	7.07	4	7.41	4	7.62	4	7.73
5	8.55	5	7.83	5	7.74	5	7.72
6	13.02	6	15.50	6	15.58	6	15.50
7	7.03	7	7.71	7	7.78	7	7.86
8	12.86	8	13.27	8	13.99	8	14.65
9	12.75	9	12.62	9	12.34	9	11.39
10	29.06	10	26.60	10	26.08	10	25.18
RMSECV	14.249	RMSECV	14.059	RMSECV	13.901	RMSECV	13.771
$\varepsilon=0.1 \sigma=0.5$		$\varepsilon=0.1 \sigma=0.6$		$\varepsilon=0.1 \sigma=0.7$		$\varepsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	18.83	1	19.35	1	19.75	1	20.02
2	9.35	2	9.29	2	9.16	2	8.97
3	20.08	3	20.27	3	20.24	3	20.34
4	7.75	4	7.73	4	7.64	4	7.52
5	7.77	5	7.87	5	8.09	5	8.34
6	15.68	6	15.91	6	16.12	6	16.28
7	7.95	7	7.92	7	7.82	7	7.70
8	15.06	8	15.49	8	15.92	8	16.28
9	10.76	9	10.28	9	10.01	9	9.82
10	24.42	10	24.01	10	23.59	10	23.36
RMSECV	13.765	RMSECV	13.812	RMSECV	13.834	RMSECV	13.863
$\varepsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	20.23						
2	8.93						
3	20.45						
4	7.39						
5	8.57						
6	16.37						
7	7.56						
8	16.47						
9	9.71						
10	23.18						
RMSECV	13.886						

Tabla 14

Optimización con k-fold cross validation para modelo CM2

$\varepsilon=0.1 \sigma=0.1$		$\varepsilon=0.1 \sigma=0.2$		$\varepsilon=0.1 \sigma=0.3$		$\varepsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	19.75	1	20.17	1	20.49	1	20.46
2	5.74	2	6.59	2	7.00	2	7.20
3	10.30	3	13.53	3	14.08	3	14.87
4	7.32	4	6.61	4	6.59	4	6.29
5	7.30	5	6.61	5	6.93	5	7.04
6	15.95	6	13.57	6	13.09	6	12.71
7	11.79	7	16.67	7	14.21	7	12.03
8	17.16	8	17.35	8	17.87	8	18.46
9	14.26	9	12.98	9	12.19	9	11.81
10	21.71	10	20.65	10	20.07	10	20.13
RMSECV	13.128	RMSECV	13.473	RMSECV	13.252	RMSECV	13.1
$\varepsilon=0.1 \sigma=0.5$		$\varepsilon=0.1 \sigma=0.6$		$\varepsilon=0.1 \sigma=0.7$		$\varepsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	20.60	1	20.74	1	20.86	1	20.94
2	7.35	2	7.52	2	7.64	2	7.69
3	15.39	3	15.68	3	15.81	3	15.86
4	6.12	4	6.25	4	6.38	4	6.48
5	7.32	5	8.01	5	9.31	5	10.49
6	12.40	6	12.14	6	11.97	6	11.88
7	10.62	7	9.80	7	8.98	7	8.64
8	18.75	8	19.03	8	19.41	8	19.65
9	11.49	9	11.14	9	10.85	9	10.67
10	20.37	10	20.73	10	21.05	10	21.26
RMSECV	13.041	RMSECV	13.104	RMSECV	13.226	RMSECV	13.356
$\varepsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	21.01						
2	7.66						
3	15.85						
4	6.54						
5	11.80						
6	11.77						
7	8.23						
8	19.90						
9	10.51						
10	21.35						
RMSECV	13.462						

Tabla 15

Optimización con k-fold cross validation para modelo CM3

$\varepsilon=0.1 \sigma=0.1$		$\varepsilon=0.1 \sigma=0.2$		$\varepsilon=0.1 \sigma=0.3$		$\varepsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	19.21	1	19.66	1	19.75	1	19.78
2	7.04	2	6.28	2	6.46	2	6.88
3	16.97	3	15.34	3	14.99	3	14.91
4	5.91	4	5.18	4	4.53	4	4.30
5	10.89	5	11.16	5	11.73	5	12.22
6	14.98	6	13.71	6	11.85	6	10.90
7	3.86	7	8.87	7	8.61	7	7.40
8	14.12	8	14.78	8	15.18	8	15.50
9	12.10	9	11.11	9	10.32	9	9.88
10	8.98	10	18.67	10	20.31	10	20.03
RMSECV	11.406	RMSECV	12.476	RMSECV	12.373	RMSECV	12.18
$\varepsilon=0.1 \sigma=0.5$		$\varepsilon=0.1 \sigma=0.6$		$\varepsilon=0.1 \sigma=0.7$		$\varepsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	19.80	1	19.81	1	19.83	1	19.86
2	7.66	2	8.43	2	9.02	2	9.74
3	14.84	3	14.81	3	14.80	3	14.75
4	4.16	4	4.07	4	3.98	4	4.04
5	12.67	5	13.08	5	13.52	5	13.91
6	10.49	6	10.63	6	10.87	6	11.02
7	6.69	7	6.77	7	6.61	7	6.83
8	15.80	8	16.07	8	16.29	8	16.54
9	9.89	9	9.95	9	9.93	9	9.89
10	19.93	10	19.95	10	20.07	10	20.16
RMSECV	12.193	RMSECV	12.357	RMSECV	12.492	RMSECV	12.674
$\varepsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	19.87						
2	10.49						
3	14.74						
4	4.12						
5	14.27						
6	11.19						
7	7.03						
8	16.69						
9	9.85						
10	20.26						
RMSECV	12.851						

Tabla 16

Optimización con k-fold cross validation para modelo CM4

$\varepsilon=0.1 \sigma=0.1$		$\varepsilon=0.1 \sigma=0.2$		$\varepsilon=0.1 \sigma=0.3$		$\varepsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	20.00	1	22.39	1	23.10	1	23.30
2	10.24	2	10.33	2	10.34	2	10.39
3	10.74	3	11.96	3	13.12	3	13.91
4	5.26	4	4.82	4	6.27	4	7.71
5	17.24	5	18.60	5	19.94	5	21.01
6	16.78	6	14.75	6	12.88	6	11.99
7	27.17	7	19.08	7	14.10	7	13.33
8	24.58	8	24.58	8	24.58	8	24.58
9	6.98	9	9.19	9	11.86	9	13.18
10	22.98	10	24.32	10	25.83	10	26.29
RMSECV	16.197	RMSECV	16.002	RMSECV	16.202	RMSECV	16.569
$\varepsilon=0.1 \sigma=0.5$		$\varepsilon=0.1 \sigma=0.6$		$\varepsilon=0.1 \sigma=0.7$		$\varepsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	23.38	1	23.53	1	23.69	1	23.84
2	10.60	2	10.76	2	11.13	2	11.37
3	14.52	3	14.91	3	15.27	3	15.55
4	8.58	4	9.08	4	9.34	4	9.51
5	21.69	5	22.21	5	22.58	5	22.82
6	11.51	6	11.17	6	10.90	6	10.77
7	13.20	7	13.48	7	13.50	7	13.46
8	24.58	8	24.58	8	24.58	8	24.58
9	13.43	9	13.46	9	13.19	9	13.06
10	25.71	10	25.44	10	25.30	10	25.21
RMSECV	16.72	RMSECV	16.862	RMSECV	16.948	RMSECV	17.017
$\varepsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	23.93						
2	11.48						
3	15.78						
4	9.61						
5	23.00						
6	10.73						
7	13.47						
8	24.58						
9	12.83						
10	25.11						
RMSECV	17.052						

Tabla 17

Optimización con k-fold cross validation para modelo CM5

$\varepsilon=0.1 \sigma=0.1$		$\varepsilon=0.1 \sigma=0.2$		$\varepsilon=0.1 \sigma=0.3$		$\varepsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	24.18	1	24.56	1	24.83	1	24.95
2	8.95	2	11.23	2	12.40	2	13.51
3	17.55	3	18.31	3	18.65	3	18.85
4	10.51	4	10.44	4	10.34	4	10.30
5	17.74	5	18.78	5	19.72	5	20.30
6	14.56	6	14.25	6	13.85	6	13.85
7	11.27	7	13.82	7	13.72	7	13.85
8	22.91	8	24.51	8	23.79	8	23.25
9	19.22	9	14.71	9	13.14	9	12.69
10	26.85	10	26.86	10	26.82	10	27.24
RMSECV	17.374	RMSECV	17.747	RMSECV	17.726	RMSECV	17.879
$\varepsilon=0.1 \sigma=0.5$		$\varepsilon=0.1 \sigma=0.6$		$\varepsilon=0.1 \sigma=0.7$		$\varepsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	25.00	1	25.03	1	25.04	1	25.06
2	14.97	2	15.90	2	16.58	2	17.09
3	19.02	3	19.18	3	19.30	3	19.45
4	10.24	4	10.19	4	10.04	4	9.89
5	20.70	5	21.11	5	21.46	5	21.78
6	13.88	6	14.02	6	14.14	6	14.22
7	10.69	7	9.83	7	9.43	7	9.35
8	23.00	8	22.95	8	22.98	8	23.05
9	12.71	9	12.67	9	12.73	9	12.72
10	27.31	10	27.07	10	26.98	10	26.76
RMSECV	17.752	RMSECV	17.795	RMSECV	17.868	RMSECV	17.937
$\varepsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	25.08						
2	17.47						
3	19.61						
4	9.74						
5	21.97						
6	14.24						
7	9.34						
8	23.09						
9	12.73						
10	26.60						
RMSECV	17.987						

Tabla 18

Optimización con k-fold cross validation para modelo midfoot

$\varepsilon=0.1 \sigma=0.1$		$\varepsilon=0.1 \sigma=0.2$		$\varepsilon=0.1 \sigma=0.3$		$\varepsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	22.29	1	22.41	1	22.43	1	22.55
2	18.24	2	17.85	2	17.80	2	17.80
3	17.14	3	19.20	3	19.84	3	19.70
4	9.50	4	9.64	4	9.85	4	10.12
5	18.97	5	20.29	5	20.85	5	21.09
6	13.66	6	13.23	6	13.61	6	13.91
7	8.19	7	6.62	7	8.20	7	9.44
8	21.79	8	22.48	8	22.59	8	22.58
9	12.73	9	12.72	9	13.08	9	13.48
10	28.11	10	30.68	10	30.73	10	30.41
RMSECV	17.062	RMSECV	17.512	RMSECV	17.898	RMSECV	18.108
$\varepsilon=0.1 \sigma=0.5$		$\varepsilon=0.1 \sigma=0.6$		$\varepsilon=0.1 \sigma=0.7$		$\varepsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	22.66	1	22.73	1	22.74	1	22.78
2	17.83	2	17.87	2	17.89	2	17.93
3	19.64	3	19.51	3	19.44	3	19.39
4	10.38	4	10.61	4	10.81	4	10.98
5	21.22	5	21.27	5	21.32	5	21.37
6	13.91	6	13.95	6	14.05	6	14.10
7	10.08	7	10.52	7	10.91	7	11.21
8	22.55	8	22.52	8	22.49	8	22.48
9	13.81	9	14.06	9	14.32	9	14.55
10	30.08	10	29.79	10	29.53	10	29.38
RMSECV	18.216	RMSECV	18.283	RMSECV	18.35	RMSECV	18.417
$\varepsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	22.81						
2	17.96						
3	19.36						
4	11.12						
5	21.43						
6	14.14						
7	11.54						
8	22.46						
9	14.73						
10	29.27						
RMSECV	18.482						

Tabla 19

Optimización con k-fold cross validation para modelo heel

$\varepsilon=0.1 \sigma=0.1$		$\varepsilon=0.1 \sigma=0.2$		$\varepsilon=0.1 \sigma=0.3$		$\varepsilon=0.1 \sigma=0.4$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	29.43	1	29.53	1	29.59	1	29.59
2	19.46	2	21.66	2	22.26	2	22.56
3	24.51	3	23.92	3	23.28	3	22.24
4	9.66	4	9.67	4	9.83	4	10.02
5	27.09	5	25.98	5	26.13	5	26.46
6	13.32	6	15.40	6	16.03	6	16.78
7	7.42	7	6.98	7	6.96	7	7.01
8	27.01	8	27.81	8	28.03	8	28.04
9	18.02	9	18.84	9	18.79	9	18.90
10	37.23	10	33.73	10	32.66	10	32.32
RMSECV	21.31	RMSECV	21.352	RMSECV	21.356	RMSECV	21.39
$\varepsilon=0.1 \sigma=0.5$		$\varepsilon=0.1 \sigma=0.6$		$\varepsilon=0.1 \sigma=0.7$		$\varepsilon=0.1 \sigma=0.8$	
k	RMSE%	k	RMSE%	k	RMSE%	k	RMSE%
1	29.59	1	29.60	1	29.61	1	29.63
2	31.12	2	23.08	2	23.21	2	23.40
3	21.65	3	21.30	3	21.15	3	21.07
4	10.20	4	10.37	4	10.52	4	10.67
5	26.74	5	29.79	5	27.01	5	27.21
6	17.45	6	17.88	6	18.15	6	18.36
7	7.14	7	7.27	7	7.43	7	7.57
8	25.05	8	28.00	8	28.00	8	27.98
9	19.01	9	19.18	9	19.34	9	19.59
10	32.22	10	32.18	10	32.20	10	32.24
RMSECV	22.01	RMSECV	21.86	RMSECV	21.66	RMSECV	21.77
$\varepsilon=0.1 \sigma=0.9$							
k	RMSE%						
1	29.65						
2	23.56						
3	21.13						
4	10.79						
5	27.40						
6	18.50						
7	7.69						
8	27.97						
9	19.74						
10	32.29						
RMSECV	21.87						

Desde la figura 43 hasta la figura 51 se exponen las gráficas correspondientes a los modelos de calibración los cuales se obtuvieron con los modelos de regresión generados con tres experimentos por cada área de presión plantar, que a la vez los mismos tres experimentos fueron utilizados para realizar el test del modelo y ver cuán efectivo puede llegar a ser. Los datos de RMSE y R^2 pueden llegar a variar con experimentos nuevos que se usarán para estimar la fuerza aplicada.

Área dedo hallux

$$\text{RMSE} = 2.47\% \text{ y } R^2 = 98.44\%$$

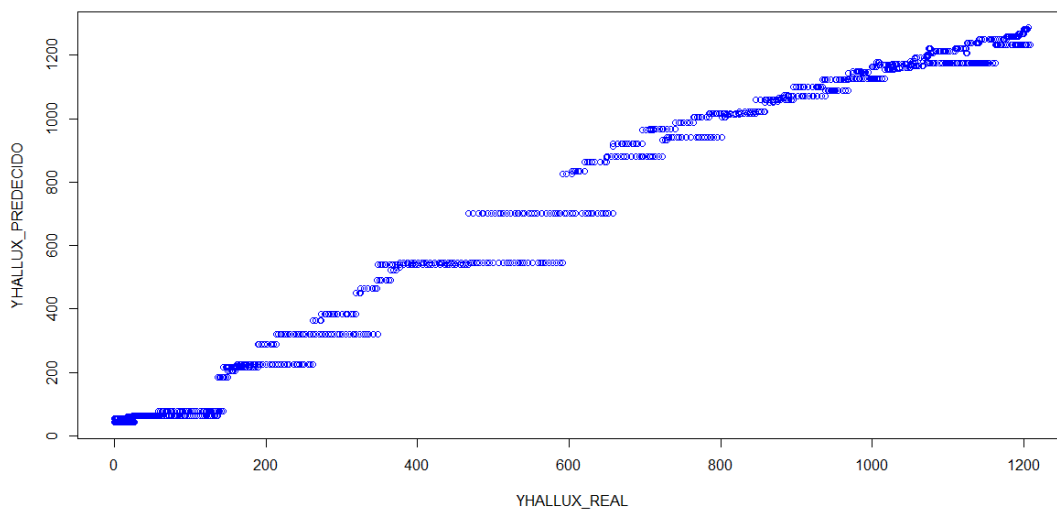


Figura 43 Fuerza predicha vs. fuerza real aplicada en área hallux con datos de entrenamiento

Área dedos menores del pie

$$RMSE = 1.62\% \text{ y } R^2 = 98.69\%$$

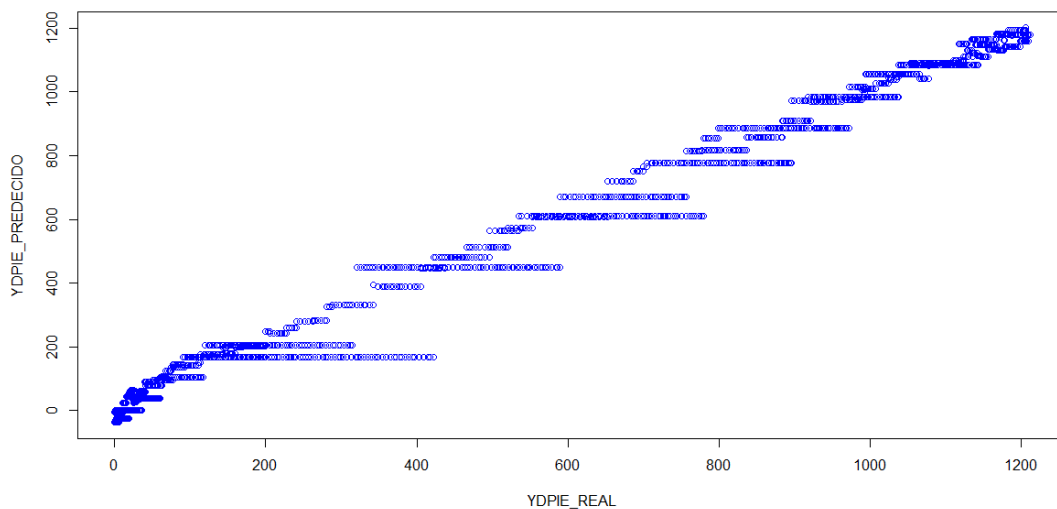


Figura 44 Fuerza predicha vs. fuerza real aplicada en área dedos menores del pie con datos de entrenamiento

Área cabeza metatarsiana 1

$$RMSE = 2.09\% \text{ y } R^2 = 98.53\%$$

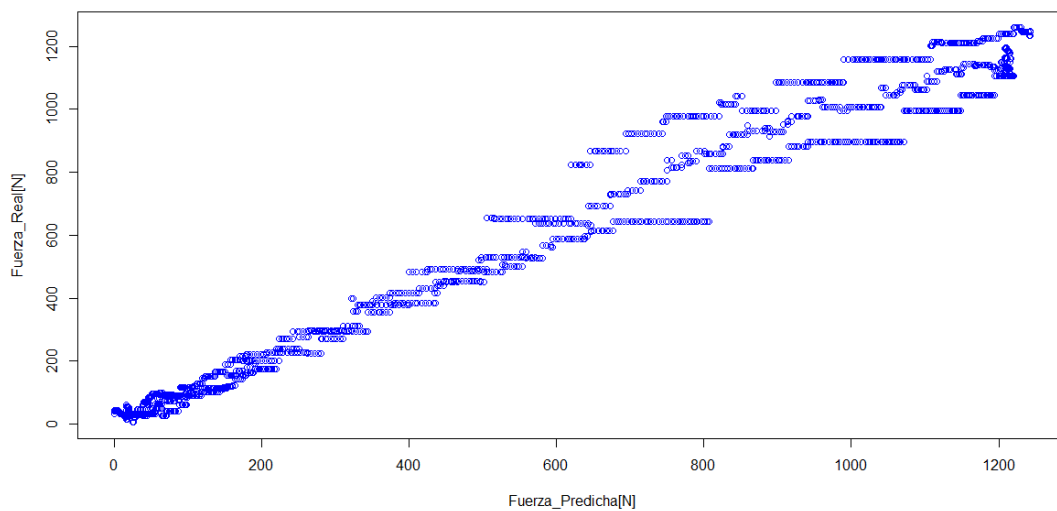


Figura 45 Fuerza predicha vs. fuerza real aplicada en área CM1 con datos de entrenamiento

Área cabeza metatarsiana 2

$$RMSE = 1.29\% \text{ y } R^2 = 98.80\%$$

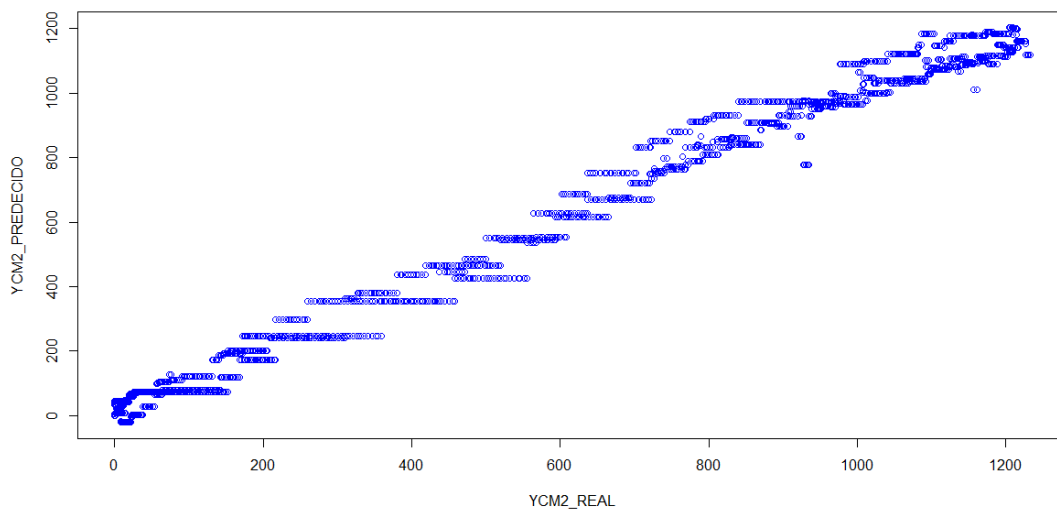


Figura 46 Fuerza predicha vs. fuerza real aplicada en área CM2 con datos de entrenamiento

Área cabeza metatarsiana 3

$$RMSE = 1.45\% \text{ y } R^2 = 98.99\%$$

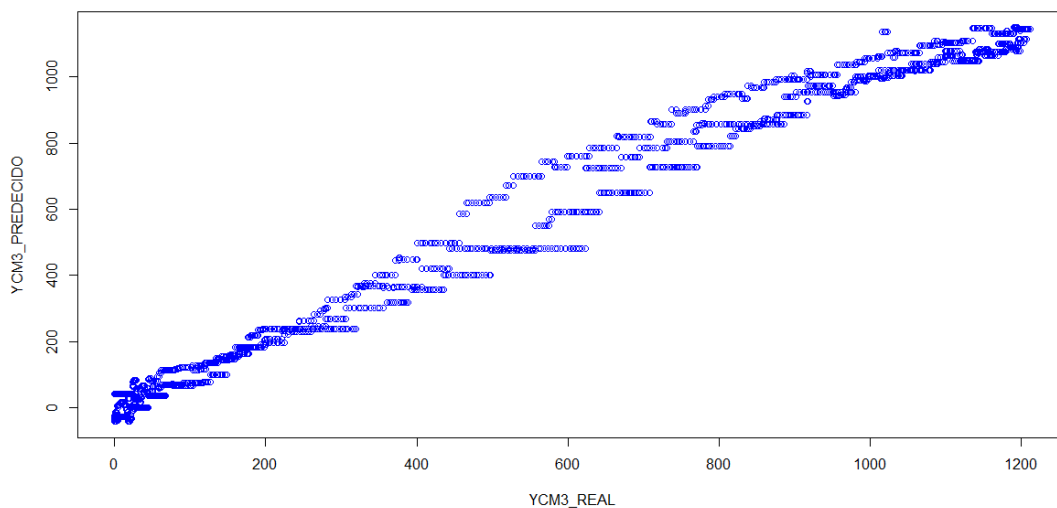


Figura 47 Fuerza predicha vs. fuerza real aplicada en área CM3 con datos de entrenamiento

Área cabeza metatarsiana 4

$$RMSE = 2.40\% \text{ y } R^2 = 97.95\%$$

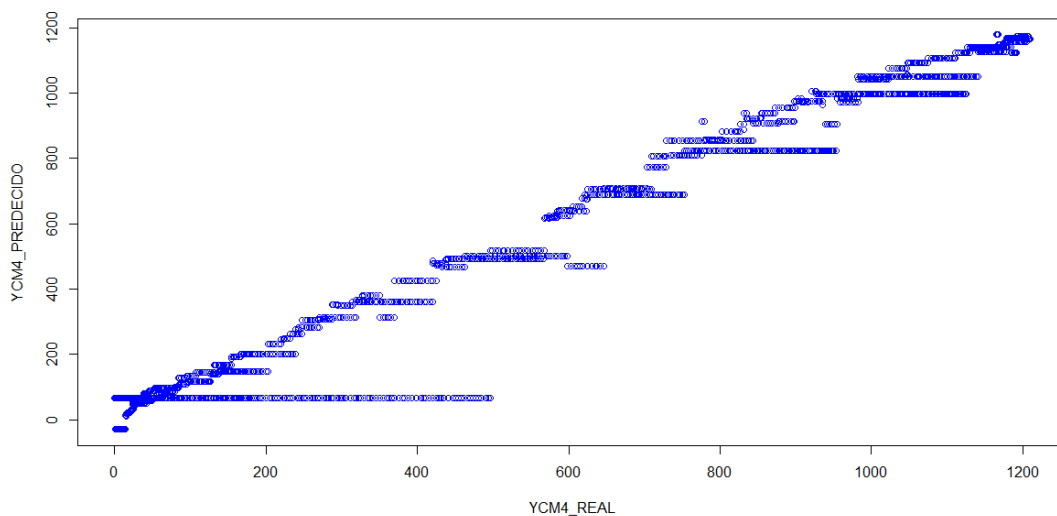


Figura 48 Fuerza predicha vs. fuerza real aplicada en área CM4 con datos de entrenamiento

Área cabeza metatarsiana 5

$$RMSE = 2.36\% \text{ y } R^2 = 97.28\%$$

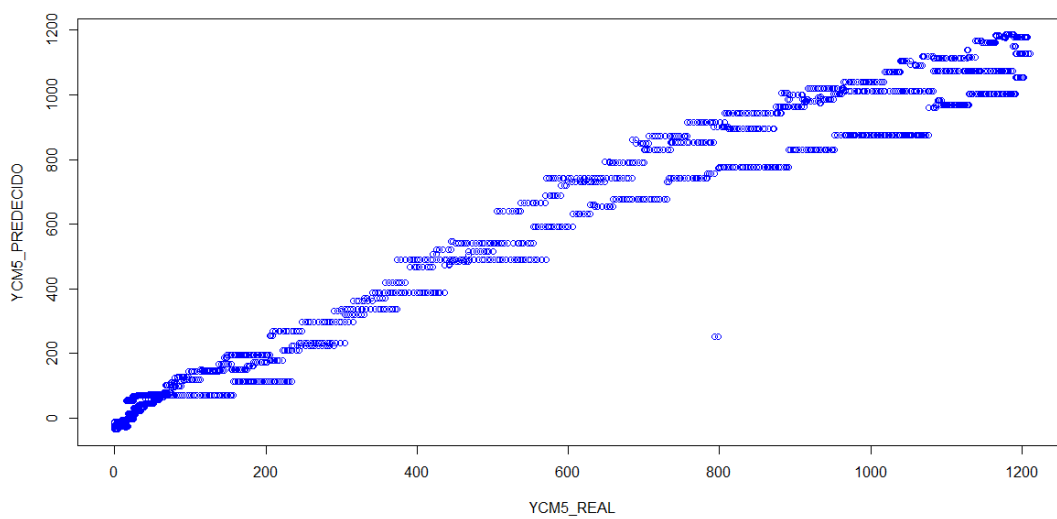


Figura 49 Fuerza predicha vs. fuerza real aplicada en área CM5 con datos de entrenamiento

Área midfoot

$$RMSE = 1.27\% \text{ y } R^2 = 99.23\%$$

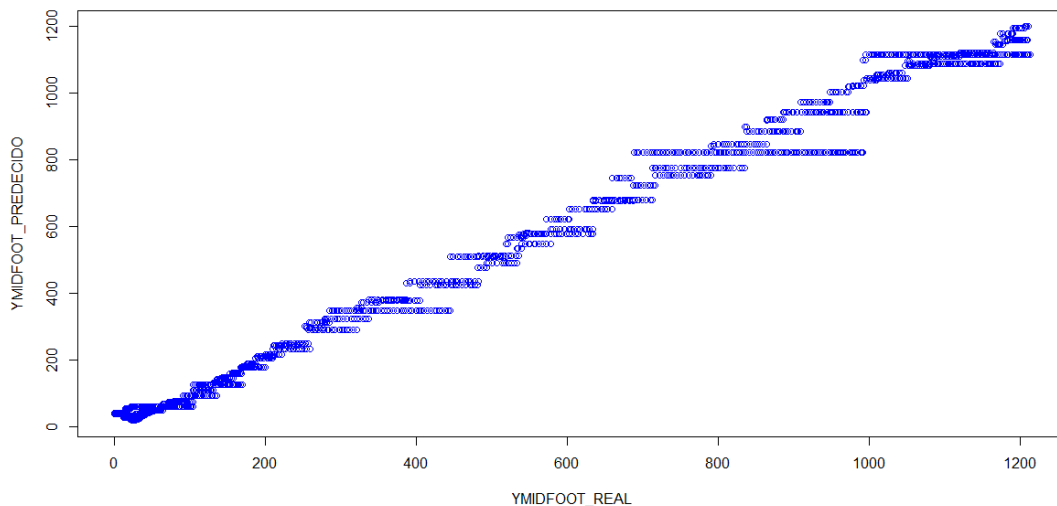


Figura 50 Fuerza predicha vs. fuerza real aplicada en área midfoot con datos de entrenamiento

Área heel

$$RMSE = 3.53\% \text{ y } R^2 = 95.95\%$$

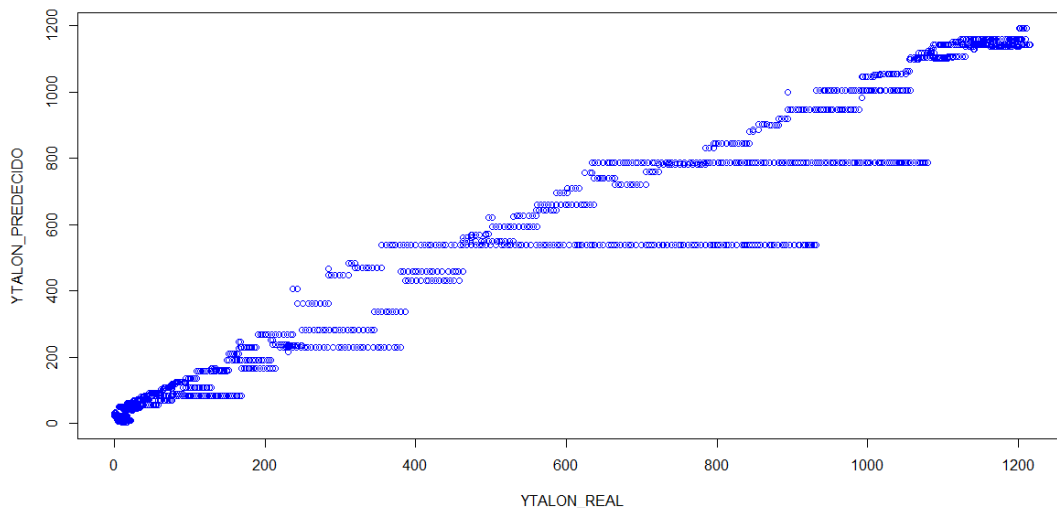


Figura 51 Fuerza predicha vs. fuerza real aplicada en área heel con datos de entrenamiento

CAPÍTULO V

VALIDACIÓN DE MODELO MULTIVARIANTE

5.1. Introducción

En el presente capítulo se mostrarán la capacidad de predicción de cada modelo de regresión, para la predicción se usarán los experimentos obtenidos en el diseño de protocolo de experimentos que no se usaron en el modelo de entrenamiento. En el entrenamiento de modelos de predicción serán utilizados los valores de los parámetros óptimos obtenidos en el capítulo anterior. Se basará en valores de $RMSE$ y R^2 para determinar la capacidad de predicción de cada modelo de regresión.

Adicional se establece un protocolo de pruebas con la pisada real de una persona para poder comprobar la aplicabilidad de los modelos multivariantes, junto con una interfaz de comunicación inalámbrica entre el prototipo de plantilla de presión plantar y la PC que será indispensable para el procesamiento de información y construcción de modelos multivariantes.

5.2. Predicción y pruebas de modelo multivariantes

A continuación se exponen los resultados obtenidos en la predicción y pruebas de validación de los modelos multivariantes. Los valores $RMSE$ y R^2 muestran una guía de la capacidad de predicción que tiene cada modelo expuesto desde la figura 52 hasta la figura 60 correspondientes a cada área de presión. Cabe recalcar que las pruebas de predicción se han obtenido con los dos experimentos finales obtenidos en el diseño de protocolo de experimentos.

Área dedo hallux

$$RMSE = 5.00 \% \text{ y } R^2 = 98.22\%$$

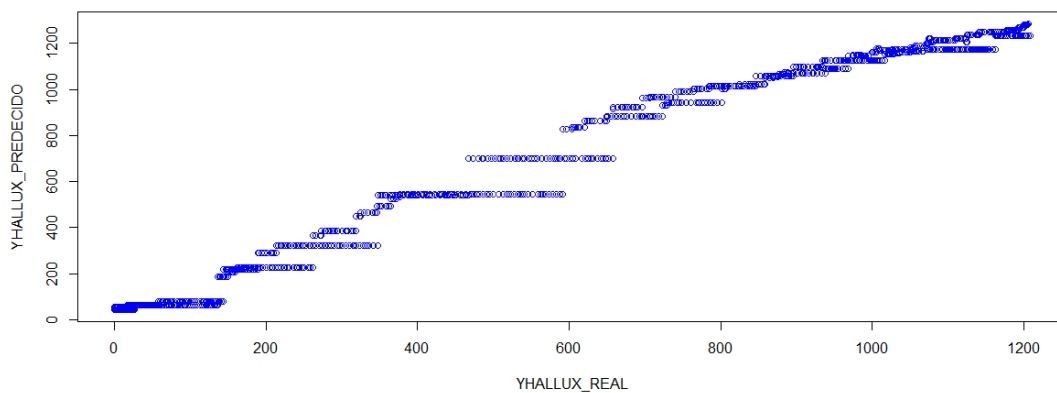


Figura 52 Fuerza predicha vs. fuerza real aplicada en área hallux con datos de validación

Área dedos menores del pie

$$RMSE = 3.69\% \text{ y } R^2 = 97.63\%$$

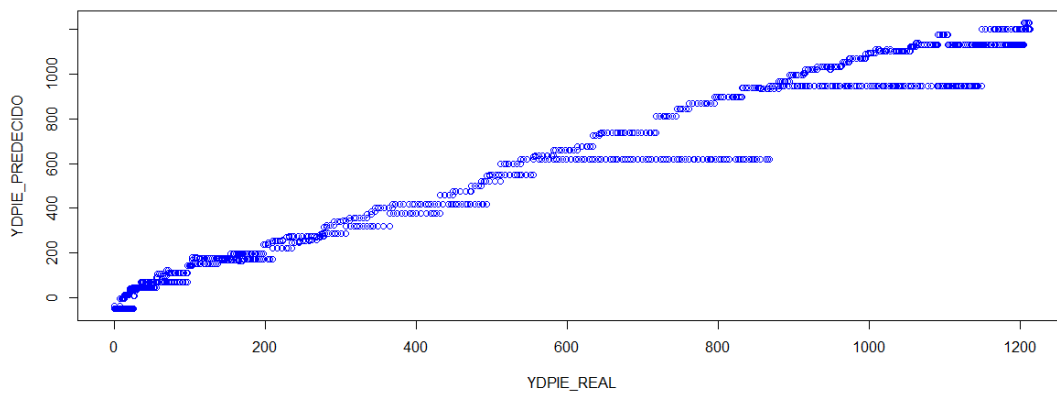


Figura 53 Fuerza predicha vs. fuerza real aplicada en área dedos menores del pie con datos de validación

Área cabeza metatarsiana 1

$$RMSE = 2.66\% \text{ y } R^2 = 98.20\%$$

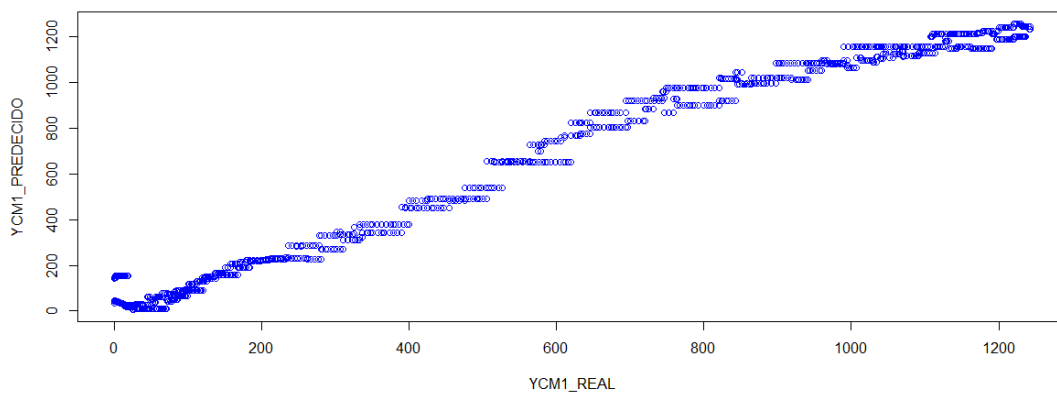


Figura 54 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 1 con datos de validación

Área cabeza metatarsiana 2

$$RMSE = 4.20\% \text{ y } R^2 = 98.32\%$$

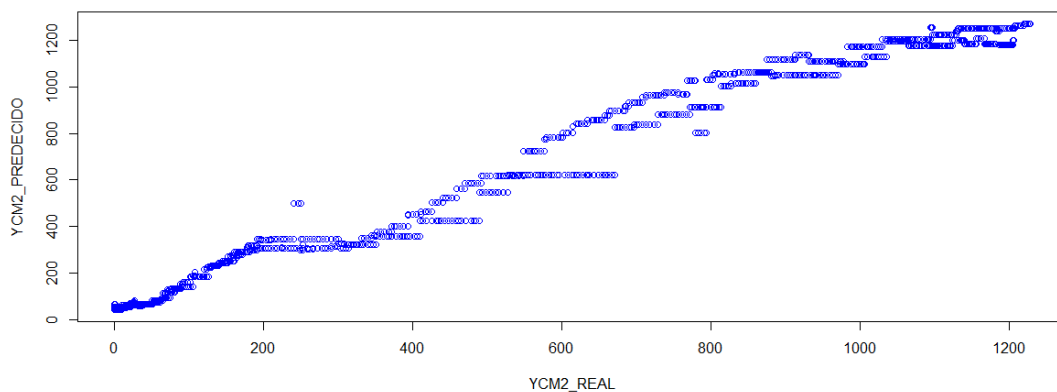


Figura 55 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 2 con datos de validación

Área cabeza metatarsiana 3

$$RMSE = 2.12\% \text{ y } R^2 = 98.99\%$$

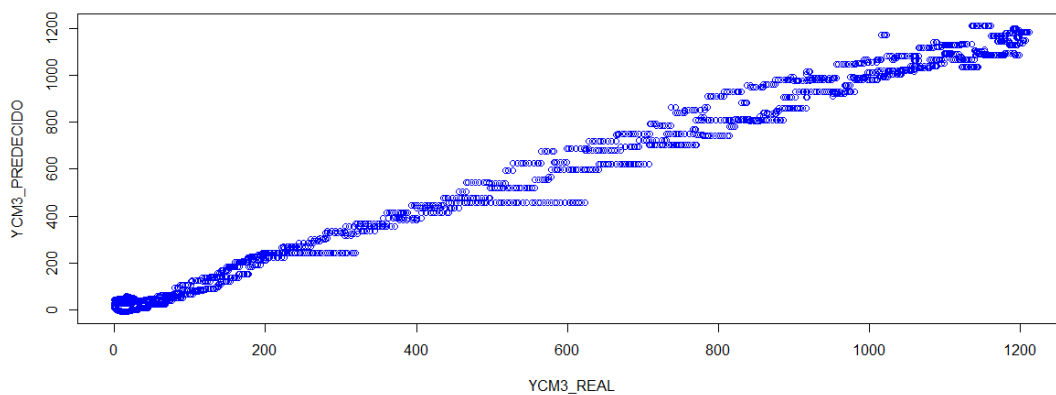


Figura 56 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 3 con datos de validación

Área cabeza metatarsiana 4

$$RMSE = 3.34\% \text{ y } R^2 = 97.59\%$$

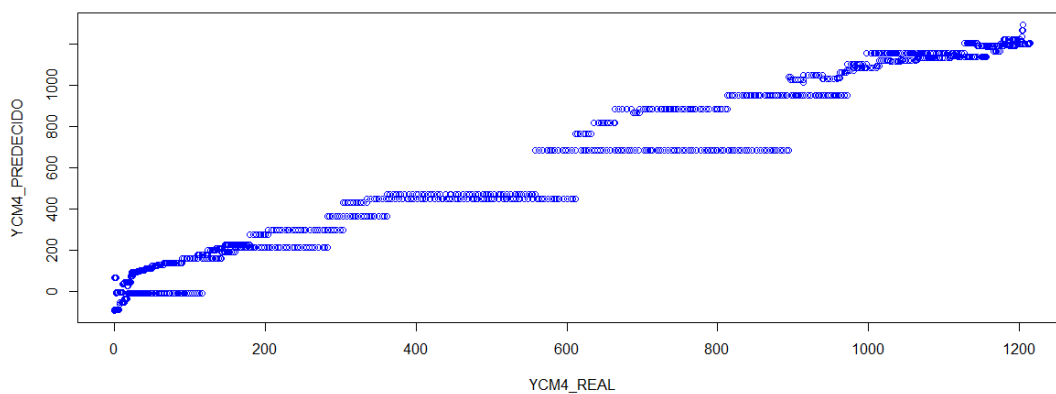


Figura 57 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 4 con datos de validación

Área cabeza metatarsiana 5

$$RMSE = 5.65\% \text{ y } R^2 = 94.60\%$$

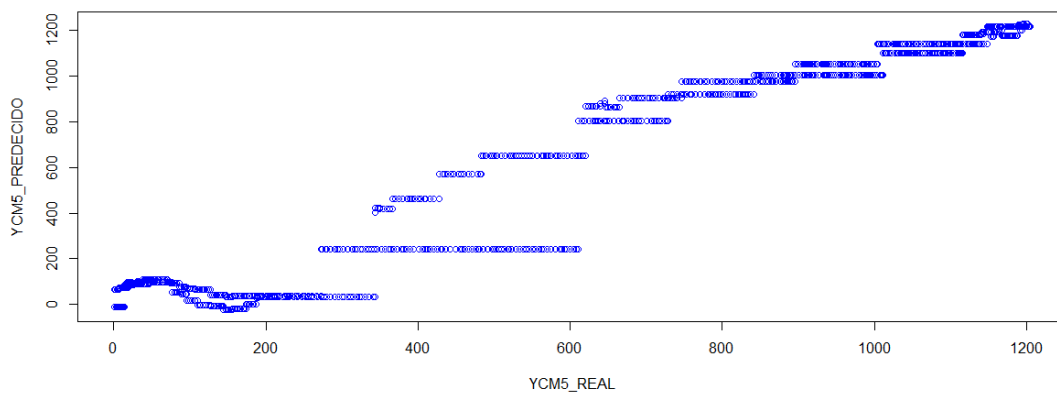


Figura 58 Fuerza predicha vs. fuerza real aplicada en área cabeza metatarsiana 5 con datos de validación

Área midfoot

$$RMSE = 6.33\% \text{ y } R^2 = 93.90\%$$

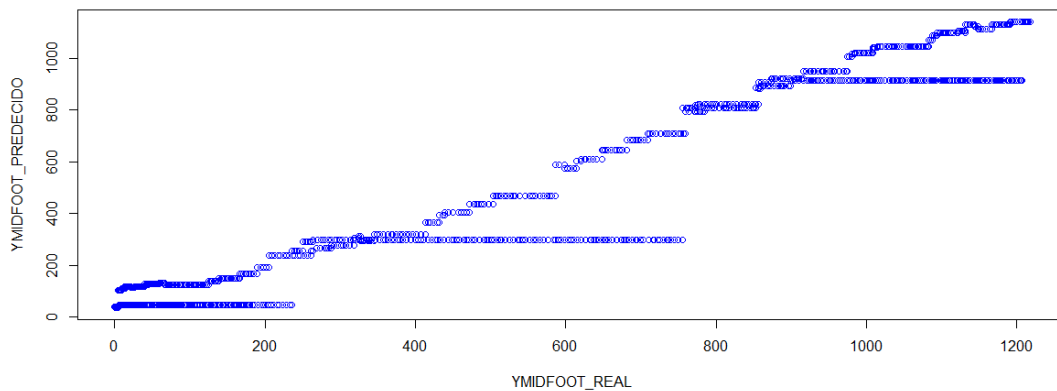


Figura 59 Fuerza predicha vs. fuerza real aplicada en área midfoot con datos de validación

Área heel

$$RMSE = 7.11\% \text{ y } R^2 = 90.91\%$$

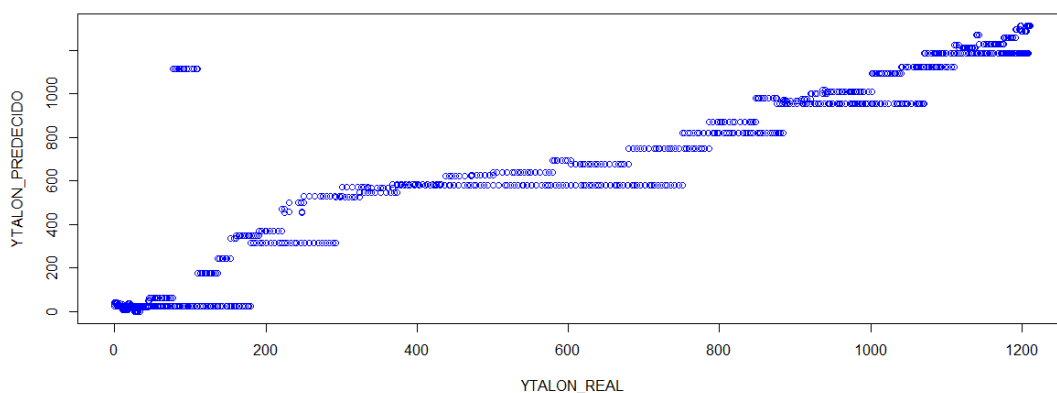


Figura 60 Fuerza predicha vs. fuerza real aplicada en área heel con datos de validación

En la tabla 20 se podrá apreciar de mejor manera la capacidad de predicción y ajuste de cada uno de los nueve modelos multivariados que se han obtenido en el desarrollo del presente proyecto y que fueron expuestos gráficamente en las figuras anteriores.

Tabla 20

Resultados de RMSE Y R^2 de Modelos Multivariados

Modelo	RMSE %	R^2 %
Área Hallux	5.00	98.22
Área Dedos Menores del Pie	3.69	97.63
Área CM1	2.66	98.20
Área CM2	4.20	98.32
Área CM3	2.12	98.99
Área CM4	3.34	97.59
Área CM5	5.65	94.60
Área Midfoot	6.33	93.90
Área Talón	7.11	90.91

Haciendo un análisis de la tabla 20 se puede concluir que las áreas que cuentan con menor número de sensores son las áreas que menor $RMSE$ y mejor R^2 presentan. Estos resultados están relacionados con el nivel de correlación que cada área presenta, puesto que a mayor número de sensores, mayor nivel de correlación presentará. Otra característica que se relaciona con el tamaño del área o el número de sensores es por la respuesta al voltaje que cada área presenta, esto debido a que dentro del área comparten una sola caída de voltaje y al encontrarse en serie todos los sensores, este valor de voltaje se divide. En consecuencia mientras más sensores existan en un área menos voltaje se alimentará para cada sensor, por lo que nos lleva a determinar que los modelos responden mejor cuando los sensores trabajan en mayores rangos de voltaje.

5.3. Interfaz de comunicación PC-plantilla de presión plantar

La comunicación entre la PC y la plantilla se ha establecido desde el proyecto precedente de manera alámbrica, quitando opciones de portabilidad de hardware por lo que se optó por establecer una interfaz de comunicación inalámbrica con el fin de aumentar portabilidad y poder brindar movilidad para trabajos futuros.

Para el establecimiento de comunicación inalámbrica se ha diseñado a partir del estándar IEEE 802.15.1 como una red *wireless personal area network* WPAN que utiliza tecnología *bluetooth*. Al utilizar tecnología *bluetooth* es fundamental poder contar con el *hardware* apropiado que permita la configuración para poder establecer comunicación; es por eso que se ha optado por la elección del módulo HC-06 por su fácil configuración como un puerto de comunicación serial, como se puede apreciar en la figura 61 en los pines 4 y 5 (de arriba hacia abajo) del módulo HC-06 presenta los pines de transmisión y recepción de comunicación serial. Mientras que en la figura 62 se ilustra el prototipo de plantilla de presión junto con el módulo de comunicación *bluetooth* conectado.



Figura 61 Módulo Bluetooth HC-06

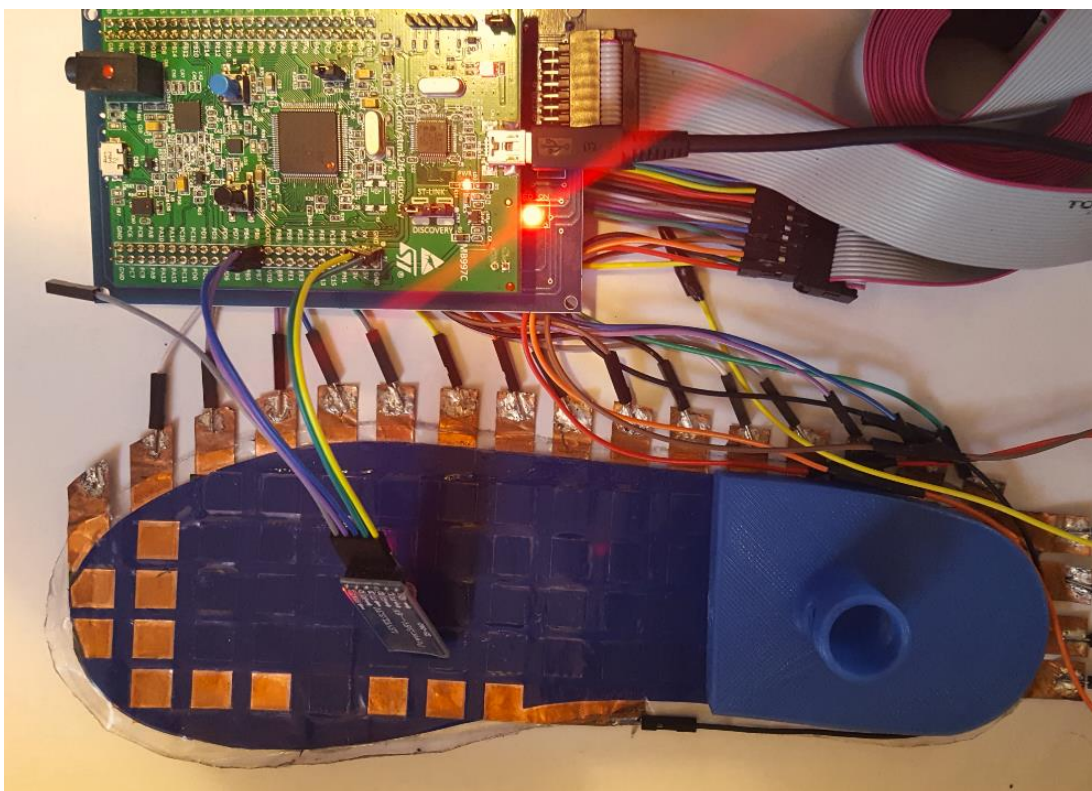


Figura 62 Plantilla de presión plantar con módulo bluetooth

El módulo HC-06 presenta una velocidad de funcionamiento de 9600 baudios por segundo, la cual es una velocidad menor a comunicación serial por cable la cual presentaba una velocidad de 115200 baudios por segundo. La figura 63 muestra la respuesta de los sensores aplicando una presión puntual en el área Heel a través de comunicación serial por bluetooth en la cual se puede observar un retardo en la adquisición de información si la comparamos con transmisión serial por cable en la figura 40.

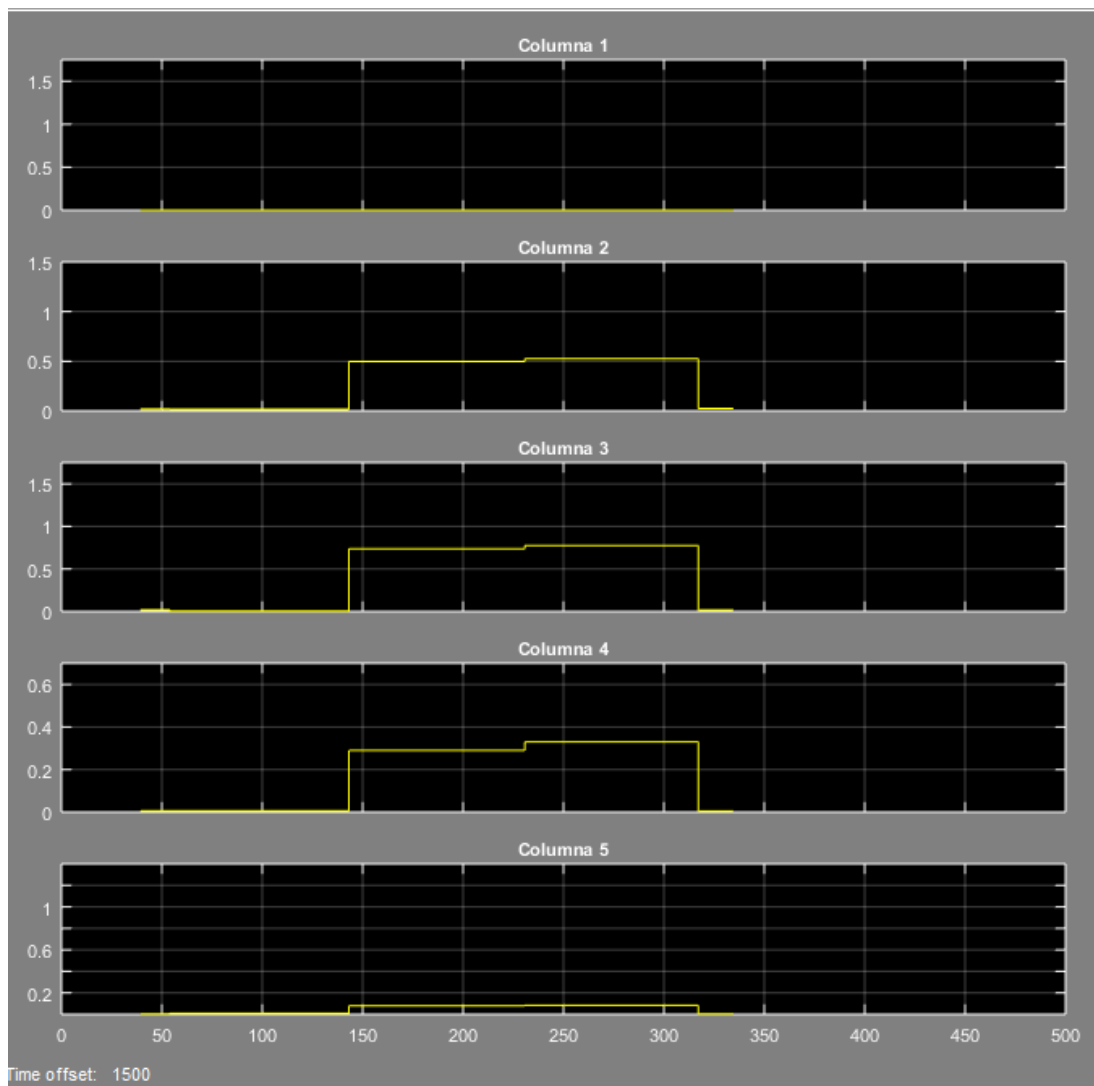


Figura 63 Respuesta de sensores por comunicación serial inalámbrica

El retardo mencionado se debe a diferentes factores, pero entre ellos se puede destacar dos principalmente: el canal de comunicación y la velocidad de transmisión de la tecnología *bluetooth* (Transmisor y Receptor).

El canal de comunicación es considerado debido a que al utilizar tecnología *bluetooth*, el canal de comunicación es un medio no guiado (aire), en el cual existen más pérdidas que en un medio guiado o cableado, además de las interferencias que se presentan en el ambiente. La velocidad de transmisión de *bluetooth* está dada por 9600 bps la cual es 12 veces menor a la velocidad que presenta el medio cableado, existiendo un retardo considerable en la transmisión y recepción de información.

5.4. Pruebas con pisada normal

Finalmente se procede a realizar pruebas con una pisada normal de un sujeto de prueba en los modelos de calibración multivariantes obtenidos con un tiempo de muestreo igual a cinco segundos. A modo de ejemplo se muestra en la figura 64 los valores en voltaje arrojados por las variables independientes del área dedo hallux que ingresarán al modelo de regresión y se procederá con la predicción de la fuerza. De manera similar son las gráficas de los datos de las variables independientes que ingresan en el resto de modelos.

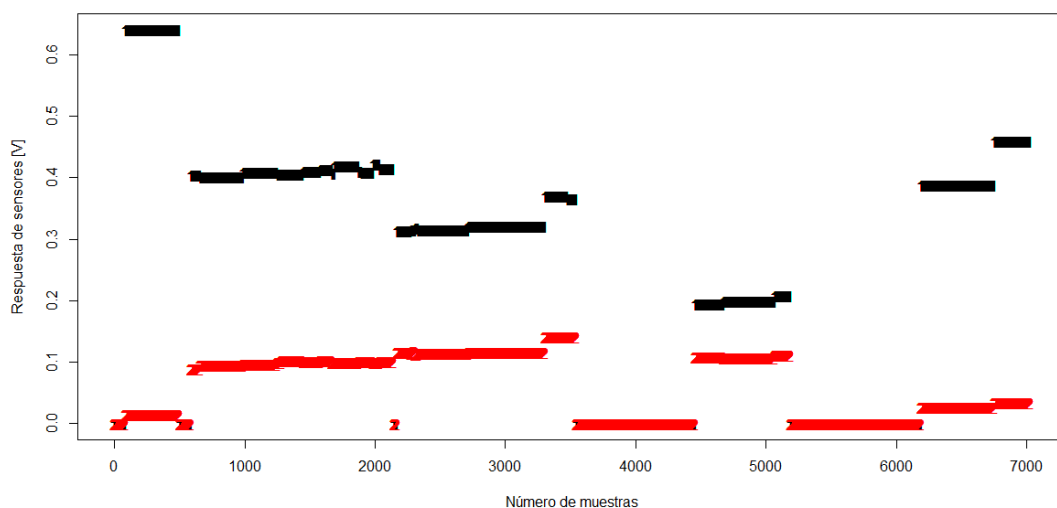


Figura 64 Variables Independientes de pisada normal del área Hallux

Con los datos obtenidos de las pruebas mencionadas en el párrafo anterior se procede con la predicción de fuerza en los nueve modelos de regresión. En cada área de presión se obtiene una fuerza predicha media y una fuerza predicha máxima, las dos fuerzas expresadas en Nétwtones, las cuales se pueden observar desde la figura 64 hasta la figura 72.

Área dedo hallux

$$F_{media} = 38.73 \text{ N}$$

$$F_{max} = 100.11 \text{ N}$$

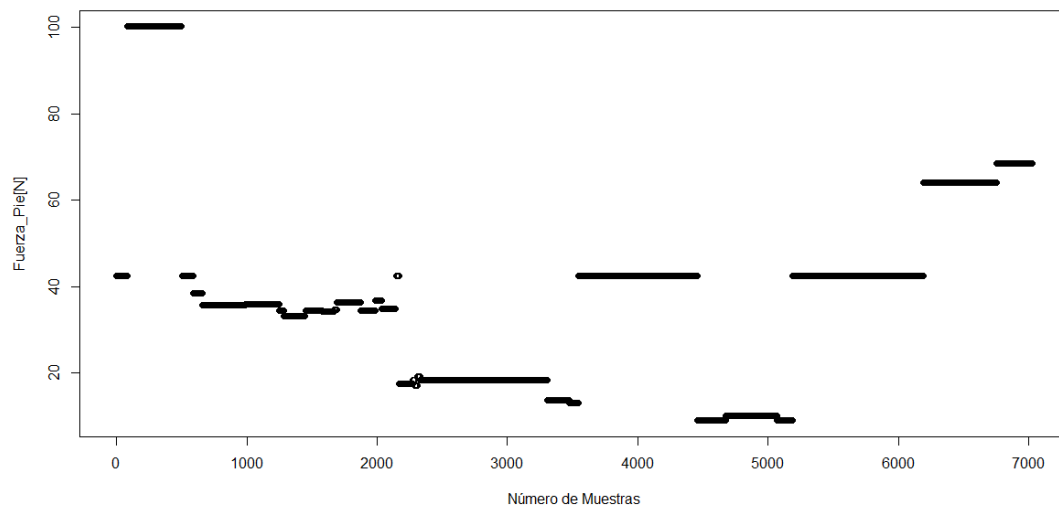


Figura 65 Fuerza predicha en área hallux

Área dedos menores del pie

$$F_{media} = 59.22 \text{ N}$$

$$F_{max} = 300.53 \text{ N}$$

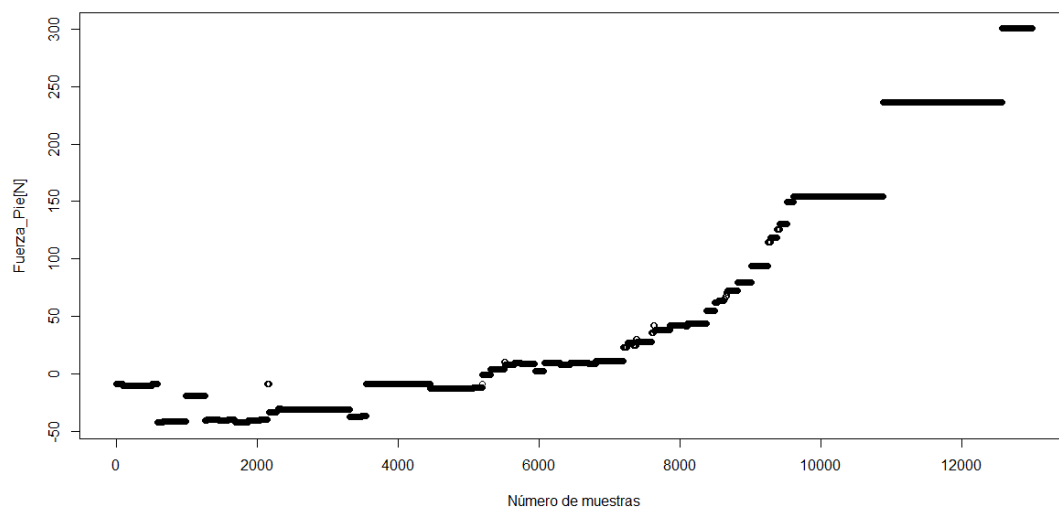


Figura 66 Fuerza predicha en área dedos menores del pie

Área cabeza metatarsiana 1

$$F_{media} = 111.03 \text{ N}$$

$$F_{max} = 199.77 \text{ N}$$

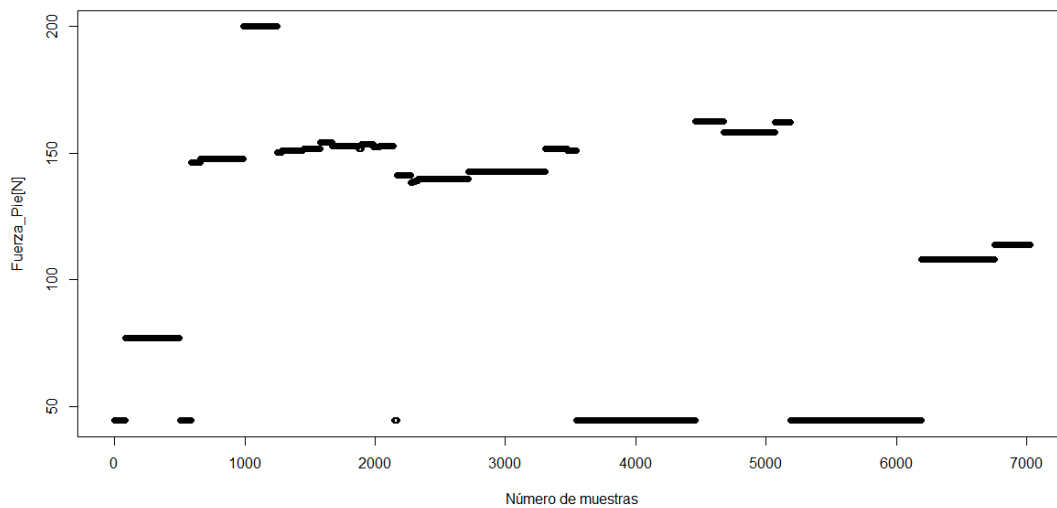


Figura 67 Fuerza predicha en área cabeza metatarsiana 1

Área cabeza metatarsiana 2

$$F_{media} = 156.83 \text{ N}$$

$$F_{max} = 277.21 \text{ N}$$

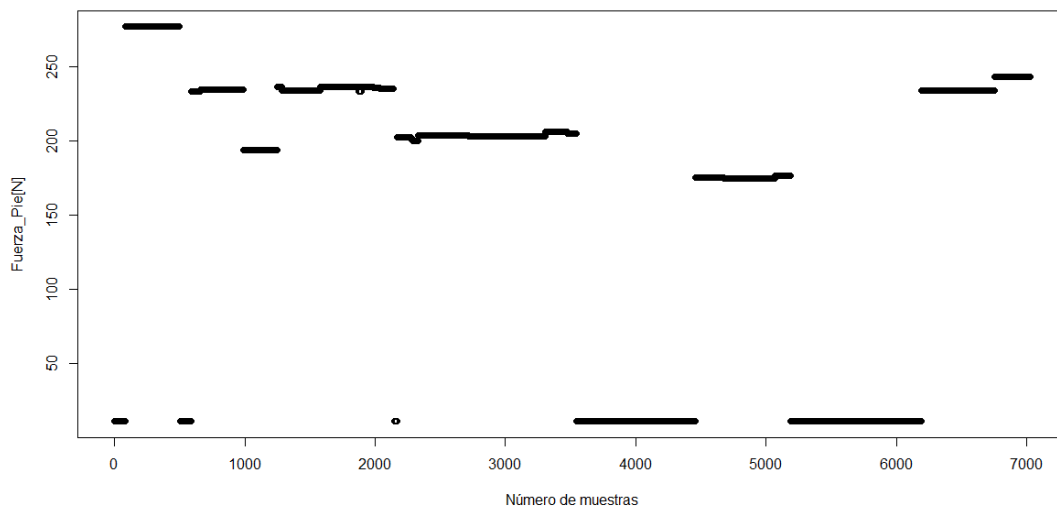


Figura 68 Fuerza predicha en área cabeza metatarsiana 2

Área cabeza metatarsiana 3

$$F_{media} = 142.86$$

$$F_{max} = 228.69$$

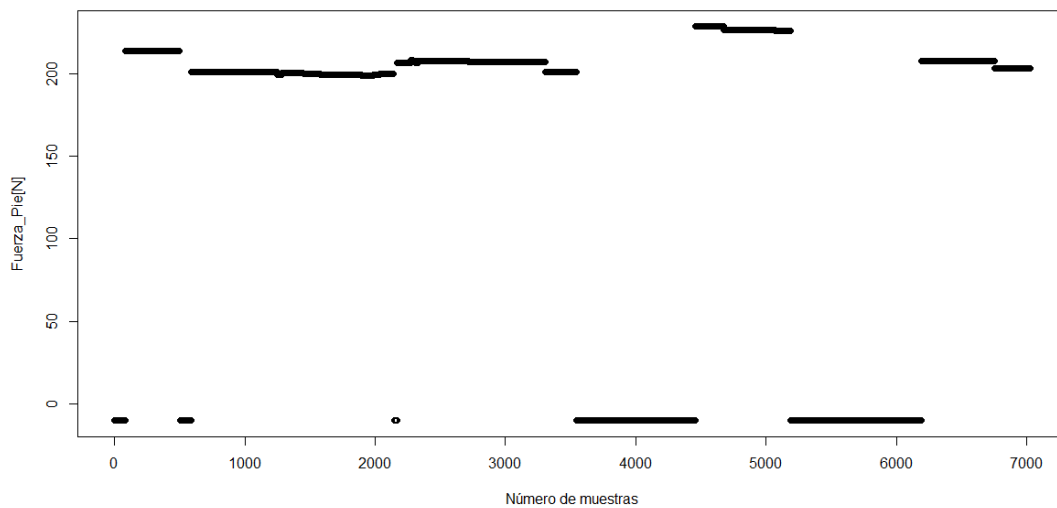


Figura 69 Fuerza predicha en área cabeza metatarsiana 3

Área cabeza metatarsiana 4

$$F_{media} = 194.57$$

$$F_{max} = 317.11$$

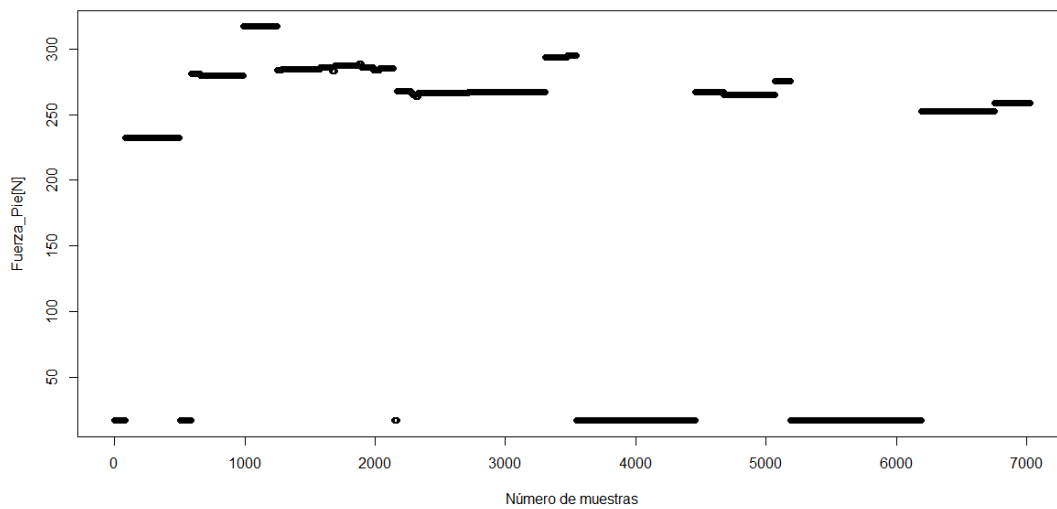


Figura 70 Fuerza predicha en área cabeza metatarsiana 4

Área cabeza metatarsiana 5

$$F_{media} = 60.26$$

$$F_{max} = 116.29$$

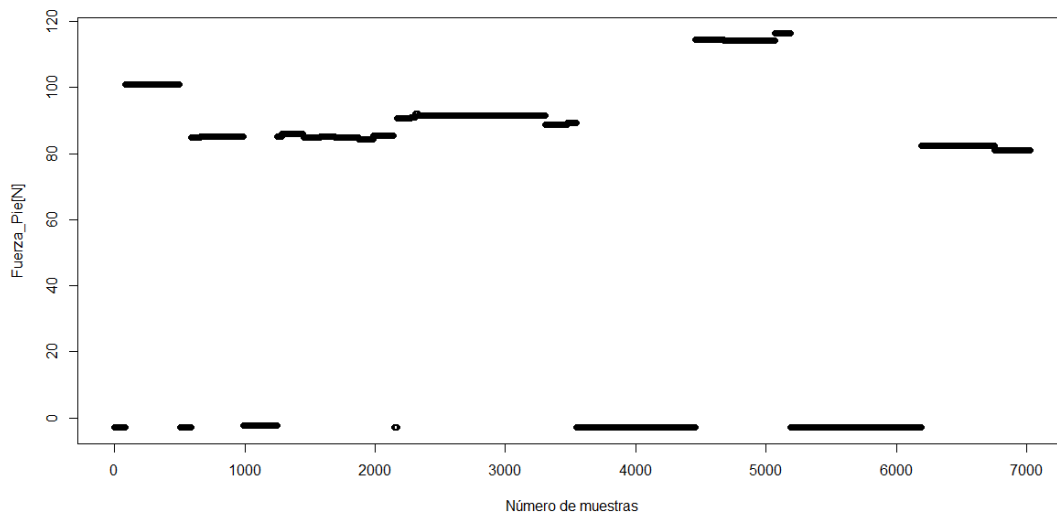


Figura 71 Fuerza predicha en área cabeza metatarsiana 5

Área Midfoot

$$F_{media} = 225.14$$

$$F_{max} = 351.97$$

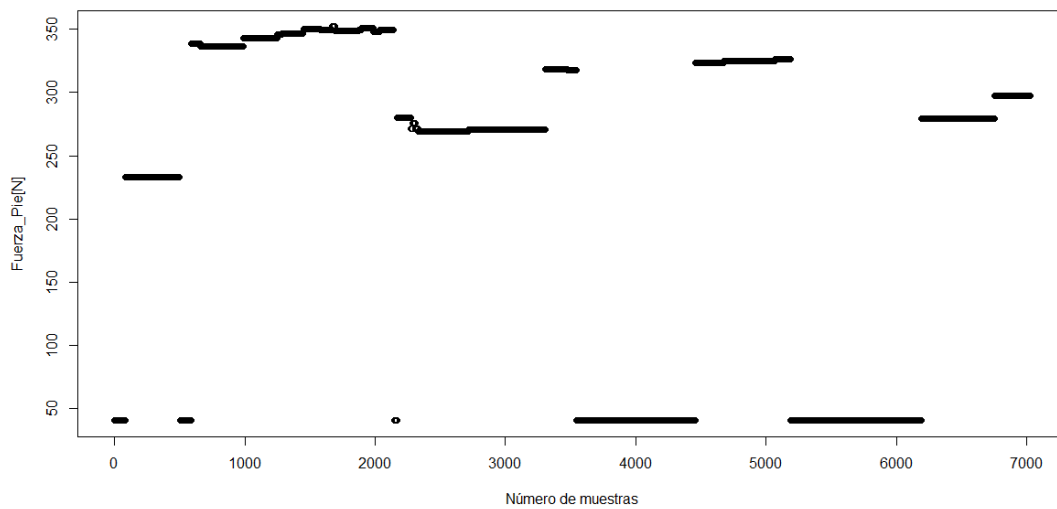


Figura 72 Fuerza predicha en área midfoot

Área heel

$$F_{media} = 228.38$$

$$F_{max} = 341.76$$

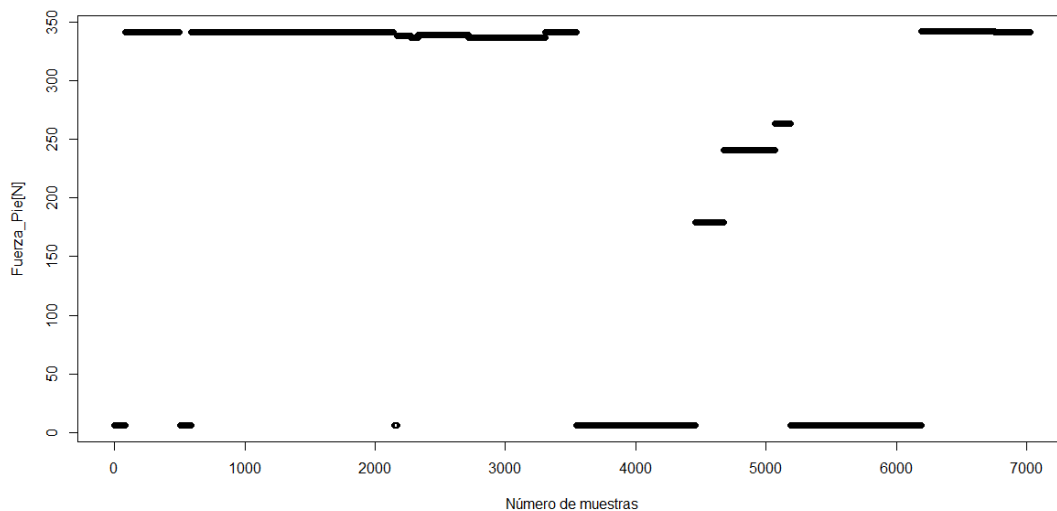


Figura 73 Fuerza predicha en área heel

En la tabla 21 se puede visualizar los valores obtenidos en las pruebas anteriores. En la tabla se puede visualizar que en las áreas de presión Midfoot y Talón se encuentran con mayor fuerza predicha, esto debido a que se encuentra en un área fisiológica donde mayor presión vertical recae al realizar una pisada. En las áreas de las cabezas metatarsianas (CM1, CM2, CM3, CM4, CM5) la presión se distribuye de manera más equitativa puesto que las cinco áreas se encuentran a la misma altura del pie, en consecuencia la fuerza predicha también es equitativa. Mientras que en las áreas de presión hallux y dedos menores del pie la presión ejercida durante una pisada normal es menor a las siete áreas restantes, por lo que la fuerza predicha resulta ser menor en comparación con el resto de áreas.

Tabla 21**Resultados finales de fuerzas predichas con pruebas de pisada normal**

Modelo	Fuerza Media [N]	Fuerza Máxima [N]
Área Hallux	38.73	100.11
Área Dedos Menores del Pie	59.22	300.53
Área CM1	111.03	199.77
Área CM2	156.83	277.21
Área CM3	142.86	228.69
Área CM4	194.57	317.11
Área CM5	60.26	116.29
Área Midfoot	225.14	351.97
Área Talón	228.38	341.76

CAPÍTULO VI

CONCLUSIONES Y RECOMENDACIONES

6.1. Conclusiones

Se realizó nueve modelos de calibración multivariantes no lineales con resultados de *RMSE* hasta 7.11% en el caso de modelo multivariable con mayor correlación, y con un R^2 desde 90.91% en el peor ajuste y hasta 98.99 % en el mejor ajuste. Los nueve modelos de calibración obtenidos han servido para poder estimar la fuerza aplicada en cada área de presión.

Se definió el estado del arte de las áreas con presión plantar elevadas en donde se afirmó que se obtienen diferentes patrones de presión dependiendo del tamaño del área. Las áreas presión plantar elevadas para el presente proyecto se estableció en total nueve: hallux, dedos del pie menores, cabeza metatarsiana 1, cabeza metatarsiana 1, cabeza metatarsiana 1, cabeza metatarsiana 1, midfoot y talón-heel.

Se estableció el número de sensores para cada área definida con mayor presión plantar, el número de sensores dependió del tamaño de cada área definida con mayor presión plantar, en donde fue necesario establecer diferentes modelos de calibración para cada área de presión.

Se generó nueve modelos multivariantes diferentes entre ellos para cada área de presión, debido a que los 63 sensores no son réplicas exactas entre ellos, de tal manera que se vio en la necesidad de optimizar los parámetros (ϵ y σ) con k-fold cross validation. Gracias a la optimización de parámetros se pudo generar los modelos de entrenamiento más óptimos por cada área. Adicionalmente se concluyó que al involucrar diferente número de sensores por cada área existe diferentes patrones de correlación en cada área.

Se elaboró un algoritmo computacional en el Software estadístico R capaz de calibrar cada área de presión en la plantilla, en el cual se determinó que el costo computacional es elevado al involucrar varias variables regresoras en cada modelo.

Se diseñó una interfaz de comunicación inalámbrica entre la PC y el prototipo de plantilla de presión plantar a través de tecnología bluetooth dada por el estándar IEEE 802.15.1. Esta tecnología aportó con movilidad y portabilidad para el prototipo pero también se notó un retardo en la adquisición de información de los sensores, que ha diferencia con la interfaz inicial alámbrica casi ese retardo era despreciable.

Se determinó el protocolo de pruebas de los modelos multivariados, en donde se utilizó dos experimentos de los cinco totales obtenidos en el diseño de protocolo de experimentos, los cuales fueron de gran utilidad para poder validar los modelos de calibración multivariados. Adicional se determinó que al utilizar dos experimentos para la determinación de la capacidad de predicción de los modelos los resultados arrojados son bastante aceptables en término de calibración y ajuste.

6.2. Recomendaciones

Se recomienda que si se va a ser el diseño y construcción de una tarjeta de circuito impreso (PCB) se la realice en un lugar donde se tenga experiencia y fiabilidad por sus trabajos, reflejando demasiada diferencia en este detalle, puesto que hasta se podría decir que por un mal diseño y construcción de una PCB se puede fracasar en un proyecto que implique el uso de estas tarjetas.

Es recomendable siempre realizar validación cruzada para optimizar los diferentes parámetros que involucren en la generación de modelos de regresión. La selección puntual de una técnica de validación cruzada de entre distintas técnicas está sujeto en base a la investigación y experiencia del uso en trabajos anteriores.

Para trabajos futuros se recomienda hacer una aplicación que permita la obtención de información junto con el procesamiento de la calibración obtenida en el

software R del presente trabajo en tiempo real, por lo que se recomienda también hacer el uso de la tecnología WIFI dada por el estándar 802.11 a/b/g/n/ac para una interfaz de comunicación inalámbrica, puesto que esta tecnología permitirá la adquisición de información en tiempo real.

Se recomienda estudiar el estado del arte de optimización de parámetros que ayudará a comprender cuál es el procedimiento adecuado para optimizar los parámetros que influyen en el entrenamiento de modelos multivariados, además del procedimiento para validar los modelos generados.

Se recomienda el uso de presente proyecto para que oriente la realización de aplicaciones puntuales en futuros trabajos de investigación, que aportará con una herramienta más a los profesionales que involucren el estudio de anomalías en el pie por una mala distribución de presión en el área plantar.

Se recomienda utilizar el software R en trabajos futuros que ha sido de gran utilidad en el actual proyecto al brindar herramientas estadísticas necesarias para el análisis de modelos de regresión multivariados además de su amigable nomenclatura para su uso y programación.

BIBLIOGRAFÍA

- Basak, D., Pal, S., & Chandra Patranabils, D. (2007). *Neural Information Processing*.
- Bauman, J. H., & Brand, P. W. (1963). Measurement of pressure between foot and shoe, 629-632.
- Bermeo, J. D., & Guamán, E. F. (2017). *Diseño e implementación de la instrumentación para la caracterización de la biomecánica de la extremidad inferior en sujetos sanos*. Sangolqui.
- Boulton, A. J., Hardisty, C. A., Betts, R. P., Franks, C. I., Worth, R. C., & Ward, J. D. (1983). *Dynamic foot pressure and other studies as diagnostic and management aids in diabetic neuropathy*.
- Brand, P. W. (1979). *Management of the insensitive limb*.
- Caselli, A., Pham, H., Giurini, J. M., Armstrong, D. G., & Veves, A. (2002). *The forefoot-to-rearfoot plantar pressure ratio is increased in severe diabetic neuropathy and can predict foot ulceration*.
- Cavanagh, P. R., & Ulbrecht, J. S. (1994). *Clinical plantar pressure measurement in diabetes: rationale and methodology*.
- Chamberlain, G. (1982). *Multivariate Regression Models for Panel Data*.
- Ctercteko, G. C., Dhanendram, M., Hutton, W. C., & Le Quesne, L. P. (1981). *Vertical forces acting on the feet of diabetic patients with neuropathic ulceration*.
- El-Hilaly, R., Elshazly, O., & Amer, A. (2013). *The role of a total contact insole in diminishing foot pressures following partial first ray amputation in diabetic patients*.
- Femery, V., Moretto, P., Renaut, H., Thévenon, A., & Linsel, G. (2002). *Measurement of plantar pressure distribution in hemiplegic children: changes to adaptive gait patterns in accordance with deficiency*.
- García, A. M. (25 de 02 de 2008). *Plantilla de presión plantar*. Obtenido de Análisis de regresión: Casos univariante y multivariante: <http://ocw.uc3m.es/estadistica/aprendizaje-del-software-estadistico-r-un-entorno-para-simulacion-y-computacion-estadistica/analisis-de-regresion-casos-univariante-y-multivariante>.

- Gorgas García, J., Cardiel López, N., & Zamorano Calvo, J. (2011). *Regresión lineal*. Madrid.
- Hadi, A., Razak, A., Zayegh, A., Begg, R. K., & Wahab, Y. (2012). *Foot Plantar Pressure Measurement System: A Review*.
- Hinz, P., Henningsen, Q., Matthes, G., Jager, B., Ekkernkamp, A., & Rosenbaum, D. (2008). *The Foot*.
- Karatzoglou, A., Smola, A., & Hornik, K. (3 de Octubre de 2016). *Kernel-Based Machine Learning Lab*.
- Kati, R. E., Forrester, S., & Fleming, P. (2010). *Evaluation of pressure insoles during running*. Loughborough University.
- Kerrigan, D. C., Frates, E. P., Rogan, S., & Riley, P. O. (1999). *Am J Phys Med Rehabil*, 354-360.
- Kobayashi, M., Konishi, Y., Fujita, S., & Ishigaki, H. (2005). *Multivariate Linear Regression Models*. Rusia.
- Koch, M., Lunde, L. K., Ernst, M., Knardahl, S., & Veiersted, K. B. (2016). *Validity and reliability of pressure-measurement insoles for vertical ground reaction force assessment in field situations*”, *Applied Ergonomics*.
- Kohavi, R. (1995). *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*.
- Kowalczyk, A. (23 de Octubre de 2014). *SVM Tutorial*. Obtenido de Support Vector Regression with R: <https://www.svm-tutorial.com/2014/10/support-vector-regression-r/>
- León Velásquez, W. (2015). *Analisis de Regresión Multiple*. Publicaciones Universidad Nacional de San Marcos.
- Liaw, A., & Wiener, M. (2002). *Classification and Regression by randomForest*. new york: ISSN.
- López Hernández, I. P. (24 de Agosto de 2016). *Desarrollo de un prototipo electrónico de sensado químico, para la detección de trinitrotuleno (tnt) y pólvora base doble en un ambiente controlado*. Sangolquí, Pichincha, Ecuador.
- Martínez Assucena, A., Barrés Carsí, M., Pérez Lahuerta, Sánchez Ruiz, M. D., Guerrero Alonso, A., & Soler Gracia, C. (2003). *Rehabilitación (Madr)*, 240-251.

- Martínez Assucena, A., Pradas Silvestre, J., Sánchez Ruiz, M. D., & Peydro De Moya, M. F. (2005). *Plantillas instrumentadas. Utilidad clínica, Técnicas Instrumentales de Diagnóstico y Evaluación en Rehabilitación.*
- Martínez Assucena, A., Vivas Broseta, M. J., Sánchez Ruiz, M. D., Peydró de Mollà, F., Barberà Guillem, R., & Baydal Bertomeu, J. (2003). *Análisis presurométrico dinámico en hemipléjico.* Santander.
- Martínez Nova, A., Cuevas García, J. C., Pascual Huerta, J., & Sánchez Rodríguez, R. (2004). *BioFoot in-shoe system: Normal values and assessment of the reliability and repeatability.*
- Microsoft. (13 de 08 de 2017). *Microsoft Azure.* Obtenido de Cómo elegir algoritmos para Aprendizaje automático de Microsoft Azure: <https://docs.microsoft.com/es-es/azure/machine-learning/machine-learning-algorithm-choice>
- Morag, E., Pammer, S., Boulton, A., Young, M., Deffner, K., & Cavanagh, P. (1997). *Structural and functional aspects of the diabetic foot.*
- Morag, E; Cavanagh, P R. (1999). *Structural and functional predictors of regional peak pressures under the foot during walking.*
- Pataky, Z., Golay, A., Bounameaux, H., Bobbioni Harsch, E., & Assal, J. P. (2003). *Relationship between peripheral vascular disease and high plantar pressures in diabetic neuroischaemic patients.*
- Pecoraro, R. E., Rehntr, G. E., & Burgess, E. M. (1990). *Pathways to diabetic limb amputation, . Basis for prevention.*
- Pérez García, J. M. (2004). *Medición de la Presión Plantar durante la Marcha. Aplicación del sistema electrónico portátil PDM240 en medicina pericial y forense.*
- Pham, H., Armstrong, D. G., Harvey, C., Harkless, L. B., Giurini, J. M., & Veves, A. (2000). *Screening techniques to identify people at high risk for diabetic foot ulceration: a prospective multicenter trial.*
- Promotion, N. C. (1990). *United States Department of Health and Human Services Centers for Disease Control and Prevention.*
- Ruckstuhl, A. (2010). *Introduction to Nonlinear Regression.*
- Salinas, J. M. (2005). *Regresión y correlación lineal.* Granada.

- San Tsung, B. Y., Zhang, M., Tat Mak, A. F., & Nar Wong, M. W. (2004). *Effectiveness old insoles on plantar pressure redistribution. Jockey Club Rehabilitation Engineering Centre. The Hong Kong Polytechnic University.*
- Sayad, S. (2010). *Support Vector Machine - Regression (SVR).*
- Shi, Z., Stumptner, M., Hao, Y., & Quirchmayr, G. (2007). *Support Vector Machines.* Australia: University of South Australia.
- StatSoft. (2013). *TIBCO Statistica.* Obtenido de How To Find Relationship Between Variables, Multiple Regression: <http://www.statsoft.com/Textbook/Multiple-Regression>
- TRIPOD. (16 de Junio de 2017). Obtenido de Support Vector Machine Regression: <http://kernelsvm.tripod.com/>
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2007). *Probabilidad y estadística para ingeniería y ciencias.* México: Pearson Educación.
- Wrobel, J. S., Connolly, J. E., & Beach, M. L. (2004). *Associations between static and functional measures of joint function in the foot and ankle.*