



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y TRANSFERENCIA DE TECNOLOGÍA

MAESTRÍA EN INGENIERÍA DE SOFTWARE
CUARTA PROMOCIÓN
TESIS PRESENTADA PREVIO A LA OBTENCIÓN DEL TÍTULO
DE MAGISTER EN INGENIERÍA DE SOFTWARE

TEMA: “DISEÑO E IMPLEMENTACIÓN DE UN MODELO
SOFTWARE BASADO EN TÉCNICAS DE INTELIGENCIA
ARTIFICIAL, PARA PREDECIR EL ÍNDICE DE RADIACIÓN
SOLAR EN RIOBAMBA-ECUADOR”

AUTOR: VERÓNICA ELIZABETH MORA CHUNLLO

DIRECTOR: DR. LUIS ENRIQUE SÁNCHEZ

LATACUNGA

2015



**VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA
MAESTRÍA EN INGENIERÍA DE SOFTWARE
CUARTA PROMOCIÓN**

CERTIFICACIÓN

Certifico que el trabajo de titulación, **“DISEÑO E IMPLEMENTACIÓN DE UN MODELO SOFTWARE BASADO EN TÉCNICAS DE INTELIGENCIA ARTIFICIAL, PARA PREDECIR EL ÍNDICE DE RADIACIÓN SOLAR EN RIOBAMBA-ECUADOR”**, realizado por la Sra. Ing. Verónica Elizabeth Mora Chunllo, 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 la Sra. Ing. **VERÓNICA ELIZABETH MORA CHUNLLO**, para que lo sustente públicamente.

Latacunga, 7 de marzo de 2016

A handwritten signature in black ink, which appears to read 'Luis Enrique Sánchez', is written over two horizontal lines.

**Dr. Luis Enrique Sánchez
DIRECTOR**



**VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA**

**MAESTRÍA EN INGENIERÍA DE SOFTWARE
CUARTA PROMOCIÓN**

AUTORÍA DE RESPONSABILIDAD

Yo, Ing. **VERÓNICA ELIZABETH MORA CHUNLLO**, con cédula de identidad 060281198-6 declaro que este trabajo de titulación “**DISEÑO E IMPLEMENTACIÓN DE UN MODELO SOFTWARE BASADO EN TÉCNICAS DE INTELIGENCIA ARTIFICIAL, PARA PREDECIR EL ÍNDICE DE RADIACIÓN SOLAR EN RIOBAMBA-ECUADOR**”, 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.

Latacunga, 7 de marzo de 2016

Ing. Verónica Elizabeth Mora Chunllo

C.C.: 060281198-6



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

iv

**VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA
MAESTRÍA EN INGENIERÍA DE SOFTWARE
CUARTA PROMOCIÓN**

AUTORIZACIÓN

Yo, Ing. **VERÓNICA ELIZABETH MORA CHUNLLO**, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación **“DISEÑO E IMPLEMENTACIÓN DE UN MODELO SOFTWARE BASADO EN TÉCNICAS DE INTELIGENCIA ARTIFICIAL, PARA PREDECIR EL ÍNDICE DE RADIACIÓN SOLAR EN RIOBAMBA-ECUADOR”**, cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

Latacunga, 7 de marzo de 2016

Ing. Verónica Elizabeth Mora Chunllo

C.C.: 060281198-6

DEDICATORIA

A Dios por su amor incondicional hacia mí,
A mi princesa Sol por su comprensión.

Verónica

AGRADECIMIENTO

Al maestro Dr. Luis Enrique Sánchez por el apoyo desinteresado, y por la asesoría comprometida al mejoramiento continuo de sus estudiantes.

A mi bella hija Solange, por el tiempo de sacrificio autónomo en cada una de sus actividades, gracias hija mía.

Verónica

ÍNDICE DE CONTENIDO

PORTADA.....	i
CERTIFICACIÓN.....	ii
AUTORÍA DE RESPONSABILIDAD	iii
AUTORIZACIÓN	iv
DEDICATORIA.....	v
AGRADECIMIENTO.....	vi
INDICE DE CONTENIDO	vii
ÍNDICE DE TABLAS	xi
ÍNDICE DE FIGURAS	xii
RESUMEN.....	xiv
SUMMARY	xvi

CAPÍTULO I

1. INTRODUCCIÓN.....	1
1.1. Identificación del problema	4
1.2 Aproximación a la solución	7

CAPÍTULO II

TÉCNICAS DE INTELIGENCIA ARTIFICIAL CON APLICACIÓN A PREDICCIÓN DE SERIES TEMPORALES NO LINEALES	10
2. Introducción	10
2.1 Enfoques inteligencia artificial.....	11
2.2 Aplicación a predicción de series temporales no lineales con IA.....	14
2.3 Argumentación de criterios para aplicación de técnicas IA	19

CAPÍTULO III

ASIGNACIÓN DE TÉCNICAS IA PARA PREDICCIÓN DEL ÍNDICE DE RADIACIÓN SOLAR EN RIOBAMBA.....	22
3. Introducción	22
3.1 Modelo biológico	23
3.1.1 Elementos de una red neuronal artificial	24
3.2 Regla de aprendizaje	26
3.3 Modelo perceptrón	27
3.3.1 Perceptrón multicapa mpl (redes feedforward).....	29
3.4 Algoritmo NN Backpropagation	30
3.4.1 Estructura y aprendizaje de la Red Backpropagation.....	31
3.4.2 Pasos para aplicar el algoritmo de entrenamiento nn Backpropagation	34
3.4.3 Número de capas ocultas	36

3.5	Aplicación de NN Backpropagation para predicción	37
3.5.1	Metodología utilizada	38
3.5.2	Algoritmo de retro propagación predictivo	41
3.5.3	Configuración del modelo predictivo	42
3.5.4	Resultados del modelo predictivo	44
3.6	Máquinas de soporte vectorial (SVM)	46
3.6.1	La idea básica de las máquinas de soporte vectorial	47
3.6.2	Método kernel	49
3.7	Tipos de funciones kernel (Núcleo)	50
3.8	Aplicaciones de las funciones kernel.....	51
3.8.1	Procedimiento clasificación.....	53
3.8.2	Método del kernel	56
3.9	Support vector regression (SVR)	58
3.10	Conclusión	63
3.11	Comparación ANN y SVM.....	64
3.12	Comparación de trabajos de predicción de radiación solar	65

CAPÍTULO IV

DISEÑO DEL MODELO SOFTWARE PARA PREDICCIÓN DEL INDICE DE RADIACIÓN SOLAR.....67

4.	Introducción	67
4.1	Lenguaje unificado de modelado (UML).....	67
4.2	Proceso unificado de desarrollo (Pud)	68
4.2.1	Ciclo de vida del proceso unificado	68
4.2.2	Flujos de trabajo	69
4.2.3	Fases.....	71
4.2.4	La plataforma de desarrollo .Net	72
4.2.5	Desarrollo e implementación con Visual Studio.....	73
4.2.6	El lenguaje de programación C#	73
4.2.7	Sql server.....	73
4.3	Iteraciones y evolución del proyecto	73
4.3.1.	Iteración 1: planificación del proyecto	73
4.3.2.	Iteración 2: diseño y creación de la base de datos.....	74
4.3.3.	Iteración 3: construcción de la herramienta.....	75
4.3.4.	Iteración 4: entrega	75
4.4.	Resultados – Construcción de la herramienta web infosolar	76
4.4.1.	Especificación de Requisitos.....	76
4.4.2	Ámbito y límites del proyecto	76

4.4.3	Análisis de los actores implicados.....	76
4.4.4	Análisis de objetivos o requisitos que debe cumplir el software	76
4.5	Costos en recursos y tiempos	78
4.6	Estimar riesgos e incertidumbres	78
4.7	Análisis	79
4.7.1	Arquitectura	80
4.7.2	Registro	80
4.7.3	Cambiar contraseña.....	82
4.7.4	Identificación.....	83
4.7.5	Gestión de organización	84
4.7.6	Gestión de usuario.....	84
4.7.7	Gestión de mediciones.....	86
4.7.8	Consultar informe.....	86
4.7.9	Consultar información	87
4.8	Diseño	88
4.8.1	Capas de la arquitectura	88
4.8.2	Diagrama de clases	89
4.8.3	Diagrama de paquetes.....	89
4.8.4	Diagrama de estado.....	90
4.8.5	Base de datos en Visual Studio	90
4.8.6	Aspectos a señalar	91
4.8.7	Interfaz gráfica	91
4.8.8	Informes y disponibilidad en internet.....	92
4.9	Implementación del software y hardware para predicción del IUV.....	96
4.9.1	Sistema de adquisición de datos.....	97
4.9.2	Sensores de la radiación UV.....	98
4.9.3	El UVI-01 detector de rayos ultravioleta	100
4.9.4	Simulación en proteus para medición del índice de radiación solar.....	101
4.10	Aproximación del PCB en altium para medición del IUV	102
4.10.1	Software del dispositivo electrónico de predicción del IUV.....	103
4.10.2	Selección del punto de referencia en la medición	103
4.10.3	Predicción de la radiación solar bajo el modelo kernel gaussiano.....	104
4.10.4	Clasificación.....	104
4.10.5	Pre procesamiento de datos	108
4.10.6	Regresión en función del tiempo.....	112
4.11	Modelos implementados con Matlab	114
4.11.1	Modelo general Model gauss1	115
4.11.2	Modelo general model gauss2:	115

4.11.3	Modelo general model fourier1:.....	116
4.11.4	Modelo General Model Fourier2:.....	116
4.12	Corrección del modelo del voltaje del sensor.	117

CAPÍTULO V

VALIDACIÓN DEL MODELO SOFTWARE.....	121
5. Introducción	121
5.1 Criterios de evaluación del modelo	121
5.2 Validación de la exactitud de los resultados	122
5.3 Validación del software del sistema electrónico.	123
5.4 Validación de la predicción del iuv en la epoch	125
5.5 Error de predicción.....	125
6 CONCLUSIONES	130
7. RECOMENDACIONES.....	132
8. REFERENCIAS BIBLIOGRÁFICAS.....	133
9. ANEXOS	138

INDICE DE TABLAS

Tabla 1	Ciudades con un alto Índice de Radiación Solar	6
Tabla 2	Comparativa entre Redes Biológicas Y Redes Artificiales.....	24
Tabla 3	Configuración de Los Patrones	40
Tabla 4	Valores del Coeficiente de determinación Mse.....	42
Tabla 5	Comparación Técnica Ann y Técnica Svm	64
Tabla 6	Comparación de Trabajos de Predicción Del Iuv.....	65
Tabla 7	Resumen de La Iteración 1	74
Tabla 8	Resumen de La Iteración 2	74
Tabla 9	Resumen de La Iteración 3	75
Tabla 10	Resumen de La Iteración 4	75
Tabla 11	Descripción Textual de Gestión De Usuario	80
Tabla 12	Descripción Textual de Nuevo Registro	81
Tabla 13	Descripción Textual de Cambiar Contraseña	82
Tabla 14	Descripción Textual de Identificación	83
Tabla 15	Descripción Textual de Usuarios.....	85
Tabla 16	Descripción Textual de Baja Usuarios.....	85
Tabla 17	Resumen Coste Temporal del Proyecto.....	95
Tabla 18	Formato de Recolección de Datos	105
Tabla 19	Valores más altos de Niuv obtenidos en Enero	108
Tabla 20	Promedio de los datos Niuv Medidos	110
Tabla 21	Algunos datos del Voltaje migrados a excel.	113
Tabla 22	Resultados de las pruebas de Simulación.....	123
Tabla 23	Respuesta y resultado de la Validación del Software	124
Tabla 24	Descripción del error en la predicción del Iuv	128

INDICE DE FIGURAS

Figura 1	Radiación Solar Mundial	3
Figura 2	Influencia En El Alto Potencial De Radiación Solar En Latinoamérica	5
Figura 3	Mapa Geográfica Y Estaciones Del Ecuador	5
Figura 4	Espectro Electromagnético	6
Figura 5	Inteligencia Artificial, Técnicas Ia	8
Figura 6	Neurona Biológica Y Conexiones Sinápticas	23
Figura 7	Esquema De Una Red Neuronal 3 Capas.....	25
Figura 8	Comparación Neurona Biológica Vs Neurona Artificial.....	25
Figura 9	Representación Función F Neuronas Interconectadas.....	26
Figura 10	Red Retropropagada Bajo Aprendizaje Supervisado	27
Figura 11	Modelo Perceptrón.....	28
Figura 12	Modelo Perceptrón.....	28
Figura 13	Frontera de decisión Perceptrón de dos Capas	28
Figura 14	Estructura General Perceptrón Multicapa(De 3 Capas)	29
Figura 15	Decisión reconocidas por distintos tipos de Perceptrones.....	29
Figura 16	Función de Transferencia, Representación Sigmoidal	32
Figura 17	Rna Mlp con tres Capas de Neuronas	39
Figura 18	Funciones de Activación Rampa Y Sigmoidal	39
Figura 19	Esquema del Rna de Predicción de series Temporales	39
Figura 20	Modelo de Pronóstico	40
Figura 21	Modelo de pronóstico con preprocesamiento de las Series de datos.	41
Figura 22	Evolución temporal de las Series	43
Figura 23	Evolución temporal de las series reales RNA.....	44
Figura 24	Evolución Temporal de las series reales Rna para experimento A.....	44
Figura 25	Gráficas de secuencias.....	45
Figura 26	Función Kernel: Polinomial-Homogénea	50
Figura 27	Función Kernel: Polinomial-Homogénea	51
Figura 28	Función de Base Radial Gaussiana	51
Figura 29	Interpolación Lineal En (X_0, X_1)	53
Figura 30	Hiperplano De Máxima Separación.....	54
Figura 31	Transformación del espacio de entrada al espacio B característico	58
Figura 32	Función de Regresión No-Lineal con la Banda E	59
Figura 33	Hiperplano en el espacio tras la transformación mediante Φ	60
Figura 34	Arquitectura de una SVM.....	60
Figura 35	Datos de prueba para el modelo de predicción a corto plazo.....	61
Figura 36	Gráficos de los Datos de prueba en Invierno y verano	62
Figura 37	Ciclo De Vida del Proceso Unificado de Desarrollo.....	69

Figura 38 Fases del Pud.....	69
Figura 39 Flujos de Trabajo	70
Figura 40 Arquitectura de Capas utilizada para el Sistema.....	71
Figura 41 Arquitectura Del .Net Framework.....	72
Figura 42 Diagrama General De Casos De Uso Del Sistema	79
Figura 43 Diagrama de Casos de uso de registro.....	80
Figura 44 Diagrama de casos de uso de cambiar contraseña.....	82
Figura 45 Diagrama de casos de uso de Identificación.....	83
Figura 46 Diagrama de casos de uso de Gestión de Organización.....	84
Figura 47 Diagrama de casos de Uso de Gestión de Usuario.....	84
Figura 48 Diagrama de Casos de Uso de Gestión de Mediciones	86
Figura 49 Diagrama de Casos de Uso de Consultar Informe	87
Figura 50 Diagrama de Casos de Uso de Consultar Información.....	87
Figura 51 Diagrama deneral de Arquitectura Multicapa en La Herramienta	88
Figura 52 Diagrama de Clases General de La Lógica de Dominio.....	89
Figura 53 Diagrama de Clases General de La Lógica de Dominio.....	89
Figura 54 Diagrama de Clases General de la Lógica De Dominio	90
Figura 55 Bases de Datos en Ms Visual Studio	90
Figura 56 Pantalla Inicial de Sesión de la Herramienta.....	92
Figura 57 Diagrama de Gantt (Mediciones)	93
Figura 58 Diagrama de Gantt Iteraciones 1 Y 2 De La Herramienta	93
Figura 59 Diagrama de Gantt Iteración 3 de La Construcción de La Herramienta..	94
Figura 60 Diagrama de Gantt Iteración 4 de La Construcción de La Herramienta..	94
Figura 61 Diagrama de Gantt Caso Estudio – Fin Pfc.....	95
Figura 62 Código Internacional de Colores para Comunicar El Iuv.	96
Figura 63 Diagrama de Bloques de un Sistema Sad.....	97
Figura 64 Circuitos de acondicionamiento de la Señal del Sensor Uv.....	98
Figura 65 Opciones de Selección del Fotodetector.	99
Figura 66 Características Seleccionadas del Fotodetector.	99
Figura 67 Respuesta Espectral de un Fotodetector.	100
Figura 68 Circuito de Aplicación para el Sad	101
Figura 69 Simulación Circuito Lógico del Sad.....	101
Figura 70 Simulación Circuito Físico del Sad.....	102
Figura 71 Conexión del Ldc A los Puertos del Arduino.	102
Figura 72 Construcción del Diagrama Circuitual del Sad	102
Figura 73 Ubicación Punto de Referencia Espoch.....	103
Figura 74 Índice Uv según La Nws Y Epa.....	104
Figura 75 Vni En Función de la Hora para Enero.....	106

Figura 76 Gráficos del Niuv en Función del Tiempo.....	109
Figura 77 Gráfico del Niuv En Función del tiempo para la Tabla 4.....	110
Figura 78 Gráficos Del Niuv A) Forma Deseada, B) Forma Obtenida	111
Figura 79 Gráficos Del Niuv A) Forma Deseada, B) Forma Obtenida	112
Figura 80 Variación Del Voltaje En Función del Tiempo (H.T).	114
Figura 81 Medidas del Voltaje Del sensor a Cielo Claro.	115
Figura 82 Alternativas de Modelos Matemáticos.....	116
Figura 83 Grafico del Modelo General Gauss 2.	117
Figura 84 El Programa de Simulación Corriendo Sobre Un Pc.	123
Figura 85 Captura De Pantalla Con El Mensaje del Sistema Al Usuario	124
Figura 86 Vni Medido En Función de la Hora.....	126
Figura 87 Vni Pronosticado en función de la Hora.	127
Figura 88 Representación de los Valores Máximos Vni	127

RESUMEN

La presente investigación tiene por objeto el diseño e implementación de un modelo software basado en técnicas de inteligencia artificial, para predecir el índice de radiación solar en Riobamba-Ecuador para lo cual se determinó la técnica IA apropiada para la predicción del índice ultravioleta (IUV), siendo las ANN Backpropagation y los SVM las técnicas que ofrecen mayor grado de precisión en la predicción de variables meteorológicas. Los datos se tomaron en base a una investigación previa durante los meses de enero 2012 a enero 2015 desde las 6h00 hasta las 18h30 a intervalos de 5 minutos, la cantidad de radiación UV se mide a través de un Sistema de Adquisición de datos y un detector Ultravioleta que mide los valores del número de índice (VNI) según The National Weather service (NWS) y la Environmental Protection Agency (EPA) propuesto en el 2004. El modelo software de predicción es un modelo local que predice el voltaje y su variación durante el día, se añade un factor de corrección tomando en cuenta el valor constante de la altura a nivel del mar. Los valores de las mediciones son guardados en la web en tiempo real y se puede acceder a través de infosolarespoch.com El pronóstico promedio proporciona el error típico de las predicciones del valor deseado, con las variables de entrada Hora y VNI se aplica el modelo de regresión lineal gaussiano simple que explica el comportamiento del VNI en función de la hora y se realice la estimación en las horas de mediciones definidas (6:00am a 6:30 pm) debido a que los días con o sin nubosidad en un año calendario a partir de las 6:00pm el índice IUV es de 1VNI constante.

PALABRAS CLAVE:

- **MODELO DE PREDICCIÓN**
- **TÉCNICAS IA**
- **MODELO SOFTWARE DE PREDICCIÓN**
- **RADIACIÓN SOLAR**

SUMMARY

This research aims to design and implement a software model based on artificial intelligence techniques to predict the rate of solar radiation in Riobamba, Ecuador for which the IA proper technique for predicting UV index (UVI) was determined being the Backpropagation ANN and SVM techniques that offer greater accuracy in predicting weather variables. Data were collected on the basis of previous research during the months of January 2012 to January 2015 from 6:00 to 18:30 at intervals of 5 minutes, the amount of UV radiation is measured through a data acquisition system and a detector UV measuring values index number (NIV) according to the National Weather service (NWS) and the Environmental Protection Agency (EPA) proposed in 2004. The model prediction software is a local model that predicts the voltage and its variation during the day, a correction factor taking into account the constant value of the height at sea level is added. The measurement values are stored on the web in real time and can be accessed through infosolarespoch.com The average forecast provides the standard error of the predictions of the desired value, the input variables Time and VNI model applies Simple Gaussian linear regression explaining the behavior of the NAV depending on the time and the estimation is performed in the hours defined measurements (6:00 am to 6:30 pm) because the day with or without clouds in a calendar year 6:00 pm UVI 1VNI rate is constant.

KEYWORDS:

- PREDICTION MODEL
- TECHNICAL IA
- MODEL PREDICTION SOFTWARE
- SOLAR RADIATION

CAPÍTULO I

EL PROBLEMA DE DESARROLLO DE MODELOS SOFTWARE PARA PREDICCIÓN.

1. Introducción

Desde hace varios años, las computadoras han sido de gran ayuda en la toma de decisiones, en un entorno globalizado en donde la información está a la orden del día, son muchas las actividades humanas especialmente referentes a salud y economía en donde la toma de decisiones son de vital importancia.

Las personas al realizar numerosas actividades requieren saber con antelación cierta información para planificar, adquirir recursos, realizar proyecciones, predecir eventos entre otros. Diversos son los Ej. que se pueden citar cuyo conocimiento previo sea importante, así por Ej. Conocer los riesgos a los que están sometidos los activos de las Pymes, sabiendo que la limitación de los recursos humanos y económicos no permiten realizar una adecuada gestión, predecir posibles heladas en una zona netamente agrícola en donde el sustento de miles de familias depende de información certera, la predicción del número de defectos residuales en el software de aerolíneas lograría evitar catástrofes, la predicción de enfermedades catastróficas salvaría vidas, entre otros:

que han sido aplicadas a la ingeniería del software desde sus inicios.

Y más recientemente, modelos gráficos, y modelos software combinan técnicas para la toma de decisiones en diversas áreas. Desde hace muchos años se ha tratado de conocer la forma de predecir acontecimientos futuros de manera que se puedan tomar acciones preventivas, Según la Organización Mundial de la Salud (OMS), el calentamiento global y la salud son los indicadores con más alto índice de crecimiento negativo desde el año 2000 en todo el mundo

Dentro de las enfermedades, el cáncer de piel es una enfermedad producida por una exposición directa a la radiación solar, fenómeno natural que nos afecta diariamente sobre la piel de las personas, produce además cataratas en los ojos, la disminución de su funcionamiento general en el organismo disminuye las defensas naturales de la piel, foto envejecimiento y cáncer de piel, reacciones alérgicas a ciertos medicamentos, y modifica el ADN

La radiación ultravioleta (sol) en porcentajes normales son beneficiosos para la salud humana puesto que ayuda en la fijación de ciertas vitaminas, sin embargo la exposición excesiva a la radiación solar puede dar lugar a efectos tanto agudos como crónicos en la salud, especialmente para la piel y los ojos, que son los órganos más expuestos, estos efectos incluyen cataratas, disminución de las defensas naturales de la piel, modificación del ADN y produce cáncer

En las últimas décadas alrededor del mundo se ha incrementado el cáncer de piel, solo hace 14 años, en el año 2000 se diagnosticaron a nivel mundial 200 mil casos de cáncer, se produjeron 65.000 muertes, y 2 millones de casos nuevos de cáncer de piel cada año todos debido al sol[1].

Debido a esta serie de problemas que genera el exceso de radiación solar presente en la salud, surge la necesidad de contar con Sistemas Inteligentes, que tengan la capacidad de poder predecir el índice de radiación solar e informar de forma anticipada a los ciudadanos, para que se pueda tomar la protección adecuada[2]

Los modelos de predicción han evolucionado en las últimas décadas, desde los sistemas inteligentes con reglas formales y teorías lógicas, hasta la aparición de las técnicas de inteligencia artificial que permiten plantear alternativas en el tratamiento de la información [3]

La predicción de la radiación solar permitiría tomar decisiones con el fin de mejorar la calidad de vida de las personas, y son diversas las técnicas de Inteligencia artificial tales como Heurística de Soluciones, Representación del Conocimiento, Sistemas expertos, lógica difusa, deducción Automática, Programación Simbólica, Redes Neuronales las que permiten desarrollar modelos fiables y robustos en la predicción de variables

Sin embargo algunos campos de la ciencia tienen gran dificultad de predicción y pronóstico exacto, parece ser que las técnicas IA no obedecen a los requerimientos del sistema.

Tomando en cuenta variables externas a la variable de medición como por Ej. nubosidad, lluvia etc., sin embargo la predicción de radiación solar no es una tarea fácil, la atmósfera es un sistema termodinámico en el cual las condiciones atmosféricas cambian constantemente, de tal manera que la predicción se ve condicionada por la calidad de los datos y la cantidad de mediciones (historial) de la potencia de radiación solar.

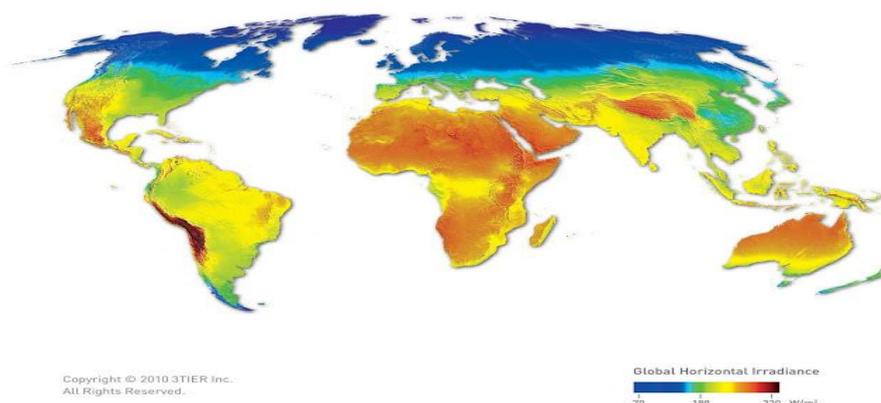


Figura 1 Radiación Solar Mundial

Fuente: (EXA, 2015)

Al ser una medición que incide directamente ante la salud de las personas y el cuidado de la piel las predicciones deben tener un margen de

error mínimo, El modelo software de predicción será un modelo regional que considera un área que cubre los 20 km.

La predicción define en detalle el índice de radiación solar local, pudiendo contrastar los resultados con la predicción global del Ecuador.

En tal virtud es de vital importancia identificar la técnica IA más adecuada que permita la recolección de los datos de entrada (potencial de radiación solar) en tiempos no lineales, en dos puntos de referencia de la ciudad de Riobamba, los valide los procese y al final permita obtener como resultado el modelo software de predicción del índice de radiación solar según los rangos de la OMS para determinar el margen de error en la predicción del sistema.

1.1. Identificación del problema

La OMS registra al cáncer de piel producida por el sol como un indicador que ha incrementado a nivel mundial, Según estadísticas son 200 mil casos de melanoma maligno, debido al sol, y se produjeron 65.000 muertes en el mundo, 18 millones de personas ciegas por cataratas y el 5 % fueron debido al sol, Países como China, Singapur, Australia e India son cuatro de los países con mayor potencial de energía solar en el mundo, en el mismo sentido México, Brasil, Chile Ecuador y Perú registran actualmente un potencial de radiación solar sobre los 7.5 Kw/m² por día cuando lo aceptable está sobre los 4Kw/m² por día

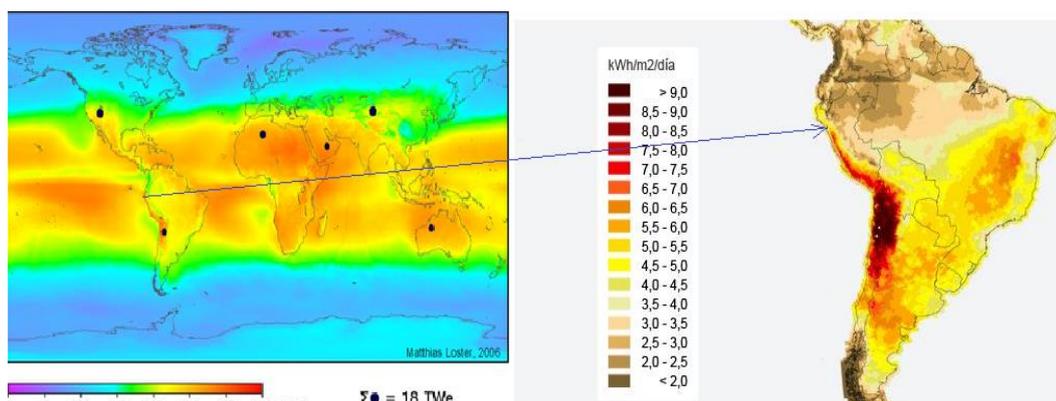


Figura 2 Influencia en el alto potencial de radiación solar en Latinoamérica

Fuente editada: (National renewable energy laboratory, 2012)

En Ecuador las ciudades interandinas como Quito (3000 metros), Riobamba (2750 metros), Ibarra (2225 metros), Loja (2100 metros) sobre el nivel del mar, han sido calificadas a lo largo de los años, como ciudades con un alto índice de radiación solar, como se ilustra en Tabla 1.

La radiación solar puede tener un efecto nocivo sobre la salud y potenciando la aparición del cáncer de piel [4].



Figura 3 Mapa geográfica y estaciones del Ecuador

Fuente editada: (EXA 2015)

Tabla 1
Ciudades con un alto índice de radiación solar

Índice de potencial solar región interandina-ecuador	Máximo kwh/m ² día	Mínimo Kwh/m ² día	Promedio Kwh/m ² día
Quito	4.8	4.4	4.6
Riobamba	4.6	4.3	4.4
Ibarra	4.5	4	4.25
Loja	4.5	4	4.25

La ubicación geográfica en los dos hemisferios, sus zonas meridionales, el daño de la capa de ozono, la línea ecuatorial y la altitud superior a los 2.700 metros sobre el nivel del mar, son los factores más incidentes para que Riobamba se ubique en uno de los lugares con mayor potencial e incidencia de energía solar y con mayor riesgo de radiación solar en el Ecuador, en relación a los últimos 5 años

La energía que emite el sol y que es recibida en la superficie de la tierra, es la fuente y variación de muchos fenómenos meteorológicos en el transcurso de los años. Es un proceso físico, a través del cual se transmite energía en forma de ondas electromagnéticas, a 300.000 km por segundo. Aproximadamente el 50% de la energía emitida por el sol está en la porción visible del espectro, el 40% en el infrarrojo y el restante en el espectro ultravioleta

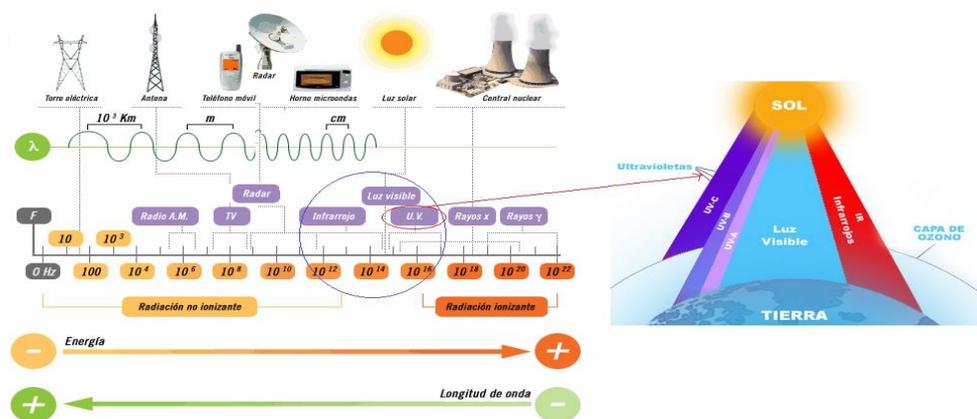


Figura 4 Espectro Electromagnético

Fuente Editada: (CONELEC, 2008)

Ecuador registra un potencial de radiación solar sobre los 7.5 Kw/m² por día, cuando lo aceptable está sobre los 4Kw/m² por día. En la ciudad de Riobamba el cáncer de piel ha aumentado en un 500% en relación a este significativo potencial. Es posible llegar a predecir fenómenos climáticos a través de modelos matemáticos, computacionales, estadísticos, modelos lineales o no lineales, y además basados en investigación y análisis de diversas series de tiempo, siendo bastante útiles en predicción a corto plazo.

La potencia de la radiación solar mediante modelos físicos y estadísticos, son modelos utilizados para estimar con una alta resolución de la radiación, siendo de vital importancia la identificación de la técnica IA permita predecir el índice de radiación solar según los rangos de la OMS [5].

Se determina que la complejidad de la presente investigación conlleva a la determinación de la mejor técnica de inteligencia artificial de tal forma que:

- No hay evidencia de aplicaciones de los procesos basados en técnicas IA en el Ecuador.
- La recolección de datos (Potencial de radiación solar) en 2 puntos de referencia, los datos de entrada a la técnica IA pueden resultar lento para la generación del modelo siendo una limitación en la investigación.
- La definición de la técnica IA incorrecta no permitirá obtener una predicción robusta y exacta
- El tiempo de la validación del modelo software

1.2 Aproximación a la solución

En la planificación de futuras necesidades del modelo de predicción, la exactitud es la característica primordial en cualquier pronóstico realizado, si la predicción es bastante cercana a la realidad bajo las series de tiempo no lineales podría proyectarse a varios años de predicción a futuro.

La predicción de la radiación solar permitiría tomar decisiones con el fin de mejorar la calidad de vida de las personas, El objetivo de que el sistema a desarrollar alcance un nivel de precisión adecuado, requiere de la aplicación de la técnica IA óptima, combinándose con exactitud en los valores medidos y una predicción acertada.

Así, las técnicas IA deben implementarse de forma jerárquica, para lograr un sistema robusto y exacto. Son varias las técnicas IA tales como Heurística de Soluciones, Representación del Conocimiento.

En el mismo sentido los sistemas expertos, lógica difusa, deducción Automática, Programación Simbólica, Redes Neuronales las que permiten desarrollar modelos fiables y robustos en la predicción de variables



Figura 5 Inteligencia artificial, Técnicas IA

Fuente:(Escolano, 2012)

Resolver problemas de predicciones en variables meteorológicas basados en técnicas de inteligencia artificial depende del desempeño de los modelos sobre los cuales se trabaje. Así por Ej. Algunas aplicaciones del uso de esta técnica de IA destacables son El Modelo para la predicción de la radiación solar a partir de redes neuronales artificiales capaces de realizar la estimación de la radiación solar horaria en la provincia de Córdoba, Argentina utilizando variables meteorológicas de sencilla adquisición [6].

De manera similar se propuso un modelo para el pronóstico de la radiación solar en un periodo de tiempo corto por medio del uso de una red neuronal base radial, mientras que Khatib et Al. (2012) propuso un método de predicción de la energía solar usando una red neuronal artificial.

Esta red llamada feedforward del tipo perceptrón multicapa, en la que los modelos fueron validados obteniendo resultados que bordean el 3, 4 y 5% de error respectivamente [7].

La aproximación del modelo Software en el proceso de análisis, consiste en la selección de la técnica más cercana a la predicción de las series temporales, Importantes investigaciones se han destacado en el uso de las técnicas IA, para predicción de series temporales en variables climáticas, así por Ej. el Modelo para la predicción del viento a partir de redes neuronales artificiales capaces de realizar la estimación de la variable en la provincia de Córdoba, Argentina utilizando variables meteorológicas de sencilla adquisición

Aplicaciones como La Red neuronal Madaline y una perceptrón multicapa para el aprendizaje de las variables climáticas radiación solar El controlador predictivo difuso simulado con la planta de colectores solares distribuidos desarrollado en la Universidad de Sevilla para la planta solar Acurex.

Bajo esta perspectiva la predicción de la radiación solar en Riobamba-ecuador estará basado en dos puntos de referencia (Norte y Sur de Riobamba) se evalúa la técnica IA óptima para el modelado, Los resultados obtenidos de la presente investigación pretenden servir de aporte local, y puede ser utilizada como línea base por el Ministerio de Salud en la prevención de cáncer de piel en la población de Riobamba, el Ministerio de Electricidad y Energía Renovable, y posteriormente pudriera extrapolarse a otras regiones del País.

CAPÍTULO II

TÉCNICAS DE INTELIGENCIA ARTIFICIAL CON APLICACIÓN A PREDICCIÓN DE SERIES TEMPORALES NO LINEALES

2. Introducción

Las aplicaciones informáticas y el desarrollo de herramientas tecnológicas y científicas son cada vez más necesarios en nuestras actividades diarias, a tal punto que no imaginaríamos como sería el mundo si no existieran. Las brechas digitales entre mundos, regiones, países y grupos humanos han sido definidos por la toma de decisiones complejas, y estas decisiones han sido motivadas por la gestión de la información generada día con día.

En este contexto la globalización de la información, ha permitido abrir las fronteras del conocimiento, dando paso a crear disciplinas que se ocupan de crear programas informáticos cada vez más adaptables a la realidad del ser humano, el esfuerzo por lograr que las computadoras realicen trabajos muy parecidos al ser humano ha obligado a muchos estudiosos a incursionar en la ejecución de operaciones como aprender, razonar, pensar, tomar decisiones comparables a las del ser humano contradiciendo la teoría de Ada Lovelace, (1843) en donde declara que una máquina solamente puede hacer lo que le decimos que haga.

La inteligencia artificial en su definición más aproximada según (RAE, 2010) afirma que: ~inteligencia artificial. f. Inform. Es una “Disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico” [8].

Según Feigenbaum y Barr (1980) [9] La Inteligencia Artificial es la parte de las Ciencias de la Computación que se ocupa del diseño de sistemas inteligentes, esto es sistemas que exhiben características que asociamos con la inteligencia en las conductas humanas.

Rich - Knight, (1991)[10] La Inteligencia Artificial es el estudio de cómo lograr que las Ciencias de la Computación realicen tareas ocupa del diseño de sistemas que por el momento, los humanos inteligentes, esto es sistemas que hacen mejor exhiben características que asociamos.

J. Mc Carthy, (1998) afirma que la Ciencia e Ingeniería de hacer computación que se ocupa de la máquinas inteligentes (especialmente automatización de la conducta programas). Esto está relacionado a la inteligente tarea de usar computadoras para entender la inteligencia humana, pero IA no tiene Luger y Stubblefield, 1993 que limitarse a métodos que son biológicamente observables

En (Iglesias, 1998) se define la Inteligencia Artificial (IA) como "*La ciencia que se centra en comportamientos inteligentes colectivos que son producto de la cooperación de diversos agentes*". Refiriéndose a los agentes a las entidades que colaboran.

La inteligencia artificial define al agente inteligente como principal unificador común entre el hombre y la máquina [11] De tal manera que es precisamente la inteligencia artificial una ciencia que estudia los diversos agentes artificiales (entidad física, virtual, racional o artificial) que toman o reciben percepciones del entorno para llevar a cabo diversas acciones.

2.1 Enfoques inteligencia artificial

La inteligencia Artificial en la actualidad abarca varias áreas que van desde áreas de propósito muy general como percepción, aprendizaje demostraciones, predicciones, pronóstico de enfermedades, automatizando

tareas intelectuales en campos como la medicina, agricultura, educación, ciencias sociales finanzas e ingeniería por citar algunas en las que tratan de crear nuevos enfoques en el tratamiento de la información y enmarcadas a eliminar las certezas que las mismas áreas la generan [12], en este contexto es potencialmente relevante para cualquiera de las áreas antes mencionadas en su contexto universal y aplicable a los paradigmas de Inteligencia Artificial

La inteligencia artificial basa sus estudios en los enfoques centrados en la racionalidad y los centrados a en el comportamiento humano. De ahí que se han creado varias técnicas que pueden ser usadas para solucionar problemas dentro de muchas categorías

Así por Ej. Las técnicas básicas como heurísticas, Representación del Conocimiento, Deducción Automática, Programación Simbólica (LISP) y Redes Neuronales, Las tecnologías como Robótica y Visión, Lenguaje Natural, Sistemas Expertos y las Aplicaciones: Diagnóstico, Predicción (Área del presente trabajo de investigación), Secuenciamiento, Diseño, Interpretación de datos.

Desde hace muchos años se ha tratado de dilucidar la forma de predecir acontecimientos futuros para tomar acciones preventivas. Es así que dentro de estos eventos futuros están las condiciones climáticas que afectan directa o indirectamente todos los días

El objetivo del presente trabajo es predecir de manera aproximada el valor de radiación solar que incide sobre la población de Riobamba – Ecuador tomando como referencia dos puntos de medición del sol. El estudio se direccionó principalmente a la definición de la técnica IA más apropiada para trabajar con aplicaciones a predicciones de series temporales no lineales (adquisición de valores de la radiación solar durante el día).

Características como: *La técnica IA óptima para la predicción de la radiación solar. Los datos recibidos de una fuente natural (Sol) deben ser almacenados en un sistema electrónico de adquisición de datos ubicados en dos puntos referenciales, (Norte /Sur) los cuáles son obtenidos a través de un sistema de adquisición de datos los mismos que son identificados en una etapa de conversión como índices UV (IUV) según indica la OMS Organización Mundial de la Salud (OMS) en colaboración con el Programa de las Naciones Unidas para el Medio Ambiente (PNUMA), la Organización Meteorológica Mundial (OMM), la Comisión Internacional sobre Protección contra la Radiación no Ionizante (ICNIRP) y la Oficina Federal Alemana para la Protección contra la Radiación (Bundesamt für Strahlenschutz, BfS).*

La exactitud de la predicción basado en valores tomados del sistema electrónico, La evaluación del modelo predictivo entregará la posibilidad de retrasar el valor de la radiación con el fin de establecer un mayor número de entradas, La determinación de los mejores estudios realizados anteriormente y basados en técnicas IA como las redes neuronales, sistemas expertos.

Con el fin de profundizar en estos avances e implementar nuevos modelos y especificaciones a partir de las necesidades expuestas definirán un modelo robusto.

Es importante notar que los valores obtenidos de las mediciones de la radiación solar se tomaron todos los días las 24 horas los 365 días del año por 3 años [13] (Datos históricos 2012-2014 obtenidos con antelación al desarrollo de la tesis) en los meses comprendidos entre julio 2012 a julio 2014, teniendo en cuenta que los valores obtenidos son no lineales.

Factores como la ubicación geográfica en los dos hemisferios sus zonas meridionales, la línea ecuatorial, la altitud superior a los 2000 metros sobre el nivel del mar, el daño de la capa de ozono, son algunos de los eventos que tiene influencia en el alto potencial de radiación solar según indica el Laboratorio Nacional de Investigación de Energía en su artículo publicado en la revista Scielo Laboratorio Nacional de Investigación De Energía[14] indica

que la variación del sol durante el día es más inminente cada vez, por tal razón los valores medidos son diversos y no presentan el mismo patrón, es decir los valores son totalmente diversos entre días meses y años.

Sin embargo, Importantes investigaciones se han destacado en el uso de las técnicas IA para predicciones en series no lineales, como son: los sistemas expertos, y las redes neuronales.

2.2 Aplicación a predicción de series temporales no lineales usando técnicas de inteligencia artificial

En el presente estudio se realiza una revisión de varias predicciones usando la técnica redes neuronales: Así por Ej. :

Juan Carlos Peña científico investigador de la Universidad Complutense de Madrid usa la red neuronal *Madaline* y una *perceptrón multicapa* para la predicción de la radiación solar a partir de redes neuronales artificiales.

Que consistió en obtener un valor aproximado de la radiación solar que incide en los vehículos solares de la World Solar Challenge 2013 en cinco puntos de referencia, en los alrededores de la carretera donde se produce la competencia; Para lo cual analiza y desarrolla un modelo basado en redes neuronales, se determinó el valor de la radiación solar en los puntos intermedios de referencia.

El autor determina el número de datos para el entrenamiento de la red, Elabora la base de datos que contendrá las entradas de la red neuronal para la predicción de la radiación solar. Diseña un modelo de predicción el cual consiste en una red neuronal MADALINE y una perceptrón multicapa para el aprendizaje de las variables climáticas y la predicción de la radiación solar en cada uno de los puntos. Evalúa la posibilidad del uso de una red neuronal base radial para la aproximación de valores de radiación en puntos intermedios, Concluye que las redes neuronales permiten incluir otro tipo

de información como entrada para dar robustez y exactitud al modelo predictivo.

Mellit y Kalogirou usa ANN para aplicaciones fotovoltaicas. En su estudio "Methodology for predicting sequences of mean monthly clearness index and daily solar radiation data in remote areas.: Application for sizing a stand-alone PV system" Usa Wavelet-redes son redes de alimentación directa utilizando ondas como funciones de activación. Wavelet-redes se han utilizado con éxito en diversas aplicaciones de ingeniería, tales como problemas de clasificación, identificación y control. En este trabajo, se investigó el uso de la arquitectura de onda de la red de adaptación en la búsqueda de un modelo de previsión adecuada para la predicción de la radiación solar total diario.

Entrena la **red neuronal**, los resultados indican que el modelo predice valores totales de radiación solar diaria con una buena exactitud de aproximadamente el 97% y el porcentaje de error medio no más de 6%.

Compara el modelo con diferentes estructuras de red neural y modelos clásicos. Algoritmos de entrenamiento para wavelet-redes requieren un menor número de iteraciones en comparación con otras redes neuronales.

Concluye que el modelo propuesto se puede generalizar y utilizar en diferentes lugares y para otros datos meteorológicos, tales como duración de la insolación y la temperatura ambiente. Por último, una aplicación que utiliza el modelo para el dimensionamiento de un sistema PV-poder se presenta con el fin de confirmar la validez de este modelo.

Al-Alawi y Al-Hinai, en su estudio Un enfoque basado en ANN para predecir la radiación global en lugares sin instrumentos de medición directa indica que utilizó una red neuronal artificial para desarrollar un modelo.

En el modelo se realiza el análisis de la relación entre la radiación global (GR) y variables climatológicas, y para predecir GR para lugares no

cubiertos por los datos de entrenamiento del modelo, Los valores de radiación globales previstas para los diferentes lugares (para diferentes meses) se compararon con los valores reales. Los resultados indican que el modelo predice los valores de radiación global con una buena precisión de aproximadamente el 93% y un error porcentual absoluto medio de 7,30.

El autor concluye que los valores mensuales pronosticados del modelo ANN en comparación con los valores reales de GR para Seeb producen una precisión de 95% y un porcentaje de error absoluto medio de 5,43, los datos para estos lugares no se incluyeron como parte de los datos de entrenamiento ANN. Por lo tanto, estos resultados demuestran la capacidad de generalización de este nuevo enfoque sobre los datos invisibles y su capacidad de producir estimaciones precisas, El modelo basado en ANN también se utilizó para predecir los valores de radiación global para Majees, una nueva ubicación en el norte de Omán.

(Chatterjee, Keyhani, 2007) define en su estudio Neural Network Estimation of Microgrid Maximum Solar Power el uso de una red neural (NN) para estimar el ángulo óptimo de inclinación en un lugar determinado y por lo tanto una estimación de la cantidad de energía disponible de la energía fotovoltaica en una microrred.

La integración de sistemas fotovoltaicos (PV) la generación de estaciones en las redes de energía eléctrica requiere la cantidad de energía disponible en el PV a estimar para sistemas de energía planean control anual de base y la operación en base diaria. Para determinar la estación PV potencia de salida máxima, los paneles fotovoltaicos deben ser colocados en un ángulo óptimo de inclinación para absorber la máxima energía del sol. Este ángulo de inclinación óptimo es una función no lineal de la localización, época del año, reflectividad del suelo y el índice de claridad de la atmósfera.

De manera similar (Khatib et Al, 2012) en su paper “Solar Energy Prediction for Malaysia Using Artificial Neural Networks” presenta un método de predicción de la energía solar mediante redes neuronales artificiales (RNA). Un ANN predice un índice de claridad que se utiliza para calcular global y difusa irradiaciones solares.

El modelo ANN se basa en el modelo de percepción multicapa de alimentación hacia adelante con cuatro entradas y una salida. Las entradas son latitud, longitud, número de días, y la relación de la luz del sol; la salida es el índice de claridad. Los datos de 28 estaciones meteorológicas se utilizaron en esta investigación, y 23 estaciones fueron utilizadas para entrenar a la red, mientras que 5 estaciones se utilizaron para probar la red.

Además, se utilizaron las irradiaciones solares medidos de los sitios de derivar una ecuación para calcular la irradiación solar difusa, una función de la irradiación solar global y el índice de claridad.

La ecuación propuesta ha reducido el porcentaje de error absoluto medio (MAPE) en la estimación de la radiación solar difusa en comparación con la ecuación convencional. Con base en los resultados, el promedio MAPE, error medio sesgo y error cuadrático medio para la irradiación solar global predicho son 5.92%, 1.46%, y 7.96%. La MAPE en la estimación de la radiación solar difusa es de 9,8%. Una comparación con el trabajo previo se hizo, y el enfoque propuesto se encontró que era más eficiente y preciso que los métodos anteriores.

(Rodríguez, 2013), en su paper “Comparison of Regression and Neural Networks Models to Estimate Solar Radiation” hace una comparación y regresión a través de modelos Rn para estimar la radiación solar, e indica que ‘para estimar esta variable, los modelos empíricos se han desarrollado utilizando varios parámetros y, recientemente, los modelos de pronóstico y predicción basado en

técnicas de inteligencia artificial, como las redes neuronales. Desarrolla un modelo lineales y redes neuronales, perceptrón multicapa, para estimar la radiación solar global diaria y comparar su eficiencia en su aplicación a una región de la Provincia de Salta, Argentina.

Utiliza los datos suministrados por la Estación Experimental Salta, Instituto Nacional de Tecnología Agropecuaria (INTA), Argentina. Por tanto, consideró los modelos neuronales redes y regresiones lineales, tres combinaciones alternativas de parámetros meteorológicos.

Obtuvo resultados con ambos métodos de predicción, con raíz media de los valores del error cuadrático medio (RMSE) entre 1,99 y 1,66 MJ m⁻² d⁻¹ para regresiones lineales y redes neuronales, y los coeficientes de correlación (r²) entre 0,88 y 0,92, respectivamente.

Concluye que a pesar de que las redes neuronales y modelos de regresión lineal pueden ser utilizados para predecir la radiación solar global diaria apropiada.

Las redes neuronales producen mejores estimaciones, si bien las redes neuronales produjeron mejores resultados. Los modelos fueron validados obteniendo resultados que bordean el 3, 4 y 5% de error respectivamente.

(Ming Qu et al, 2011) en su trabajo titulado Automatic Solar Flare Detection Using MLP, RBF, and SVM en la cual realiza la detección automática solar, hace una comparación de 3 técnicas perceptrón multicapa (MLP), de función de base radial (RBF), y la máquina de vectores de soporte (SVM) clasificadores. El paso de preprocesamiento es obtener nueve características principales de las llamaradas solares de los clasificadores. Los resultados experimentales muestran que mediante el uso de SVM.

La magistral investigación de (Luis Martín,... Luis F. Zarzalejo, Jesús Polo, Ana Navarro, Ruth Marchante, Marco Cony, 2009) en su artículo “Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning” presenta una comparación de los modelos estadísticos basados en series temporales aplicadas para predecir los valores media diarias de irradiancia solar global con un horizonte temporal de 3 días. La mitad de los valores diarios consisten acumulados por hora irradiancia solar global de aumento solar a mediodía solar y desde el mediodía hasta el amanecer de cada día. El conjunto de datos de la radiación solar en tierra utilizado pertenece a las estaciones del Servicio Meteorológico Nacional española (AEMET).

Los modelos probados son, redes neuronales y modelos autorregresivos de lógica difusa. Debido al hecho de que media diaria de series de tiempo de irradiancia solar es no estacionario, ha sido necesario para transformar a dos nuevas variables estacionarias (índice de claridad y de componentes perdido) que se utilizan como entrada de los modelos predictivos.

El mejor método para predecir valores media diarias de irradiancia solar es los modelos *de redes neuronales* con componente perdido como entrada, excepto la estación de Lérida, donde los modelos basados en el índice de claridad tienen menos incertidumbre porque esta magnitud tiene un comportamiento lineal y es más fácil para simular por modelos.

2.3 Argumentación de criterios para aplicación de técnicas ia en la predicción de radiación solar

De las investigaciones y aplicaciones antes mencionadas se destaca los siguientes argumentos

- Kalogirou[13], usa ANN para aplicaciones de sistemas de energías renovables.
- Mellit et al[14], estudia ANN para el dimensionado de sistemas fotovoltaicos.

- Mellit y Kalogirou, utiliza ANN para aplicaciones fotovoltaicas. Al-Alawi y Al-Hinai, aplica multicapa alimentación red delantero, nuevo algoritmo de entrenamiento de programación (BP) para predecir la radiación global para localizaciones Seeb.
- Mohades et al, utiliza ANN para el modelado de la radiación solar global en Arabia Saudita. El algoritmo (BP) se utiliza para la formación de las diferentes configuraciones de redes neuronales de feed-forward de múltiples capas.
- Rehman y Mohandes, utilizo cuatro combinaciones de parámetros de entrada (día, máxima temperatura, temperatura media del aire, humedad relativa) para estimar la radiación solar difusa para la ciudad de Abha en Arabia Saudita.
- Lazzús et al, utiliza un modelado ANN con la velocidad relativa del viento, humedad relativa, temperatura del aire y temperatura del suelo, para estimar la radiación solar global por hora para La Serena en Chile.

La técnica ANN es un método alternativo y de vanguardia frente a los métodos convencionales.

La técnica ANN es específica para aplicaciones de la energía solar, todos los modelos mencionados tienen los resultados deseados en virtud de valores obtenidos según la inclinación y parámetros de medición.

- Los errores aproximados muestran valores cuadráticos absolutos que bordean entre 0,88 y 1,66
- Los errores porcentuales absoluto medio de 7,30.
- Error Cuadrático global 5.92%, 1.46%, y 7.96%.
- Error Medio absoluto entre 5,43 y 6%.
- La Exactitud de predicción que bordea aproximadamente el 98% , con una precisión del 97%

Se concluye que:

- La selección de Parámetros de medición es relevantes para obtener el menor error de predicción.
- Se puede obtener la menor tasa de error en las predicciones en tiempo real utilizando técnicas avanzadas de reconocimiento de patrones o características.
- La técnica que permite configurar los parámetros radiación solares exactitud precisión en series no lineales es NN Backpropagation
- La predicción de la radiación solar requiere de variables geográficas y meteorológicas como parámetros de entrada a los modelos de ANN

De la revisión bibliográfica se puede destacar que según históricos las técnicas que se están utilizando actualmente para predecir radiación solar Las ANN Backpropagation y los SVM, y específicamente para predicción de variables metereológicas las ANN Backpropagation.

Al predecir eventos relacionados con afecciones con los seres humanos con lapsos de exposición largos de la radiación solar es evidente, que la precisión, la exactitud y el menor valor de error son característicos en las aplicaciones de las técnicas IA.

La información científica derivada del presente proyecto permitirá conocer con anticipación sobre el nivel de radiación, excesiva y/o acumulada de la fuente natural (SOL), la población de Riobamba, podrá conocer sobre los efectos dañinos a corto y largo plazo principalmente en la piel, la información obtenida servirá de base para los Ministerios y en especial para la Sociedad de lucha contra el cáncer.

CAPÍTULO III

ASIGNACIÓN DE TÉCNICAS IA PARA PREDICCIÓN DEL ÍNDICE DE RADIACIÓN SOLAR EN RIOBAMBA

3. Introducción

Las Redes neuronales Artificiales (RNA) son una rama de la inteligencia artificial que consiste en el aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso biológico, el objetivo de las RNA es conseguir respuestas similares a las que es capaz de dar el cerebro, su característica principal es la generalización y robustez en la predicción de series de tiempo dado permitiendo un modelo con más grados de libertad, La teoría de las RNA brindan alternativas a problemas en los que los métodos tradicionales no han entregado resultados efectivos.

Dentro de las aplicaciones más exitosas de las RNA son:

- Procesamiento de imágenes y de voz
- Reconocimiento de patrones
- Planeamiento
- Interfaces adaptivas para sistemas Hombre/máquina
- *Predicción*
- Control y optimización
- Filtrado de señales

Las redes neuronales han sido utilizadas exitosamente en muchos tipos de problemas de predicción, debido a que son capaces de modelar y predecir series de tiempo lineal y no lineal, con un alto grado de exactitud y precisión.

El análisis de las redes neuronales como técnica de inteligencia artificial obedece a la evaluación de la capacidad de las redes neuronales artificiales como herramienta de predicción a través de las distintas configuraciones así por Ej. NN Backpropagation SVM, algoritmos de aprendizaje Backpropagation y Resilient Propagation [15] serán fuente de análisis en este apartado.

3.1 Modelo biológico

La clave para el procesamiento de la información son las neuronas y las conexiones entre ellas (sinapsis). De alguna manera la neurona elabora una señal de salida a partir de las dendritas, que son la vía de entrada de las señales que se combinan en el cuerpo de la neurona, las envía al axón, que es el camino de salida de la señal generada por la neurona [16]. Las sinapsis, son las unidades funcionales y estructurales elementales que median entre las interacciones de las neuronas.

En las terminaciones de las sinapsis se encuentran unas vesículas que contienen unas sustancias químicas llamadas neurotransmisores, que ayudan a la propagación de las señales electroquímicas de una neurona a otra.

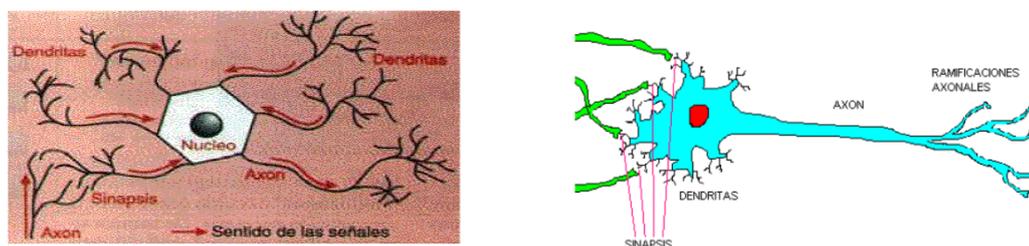


Figura 6 Neurona biológica y conexiones sinápticas

Fuente: (S.Y. Kung, 2012)

En una neurona biológica, la neurona es estimulada o excitada a través de sus entradas y cuando se alcanza un cierto umbral, la neurona se dispara o activa, pasando una señal hacia el axón.

3.1.1 Elementos de una red neuronal artificial

Las redes neuronales constan de dispositivos de proceso:

Las neuronas, que a partir de ellas, se pueden generar 4 representaciones específicas, de tal forma que un estado conjunto de ellas pueden significar una letra, un número u otro objeto.

- Las que reciben estímulos externos relacionados con el aparato sensorial, (Toman información de entrada), dicha información se transmite a ciertos elementos internos que se ocupan de su procesamiento.
- Es en las sinapsis y neuronas correspondientes a este segundo nivel, donde se genera cualquier tipo de representación interna de información. (Como no tienen relación directa con la información de entrada ni con la salida, estos elementos se denominan unidades ocultas).

Tabla 2

Comparativa entre redes biológicas y Redes artificiales

<i>Redes biológicas</i>	<i>Redes artificiales</i>
Neuronas	Unidades de proceso
Conexiones sinápticas	Conexiones ponderadas
Efectividad de la sinápsis	Peso de las conexiones
Efecto excitatorio de una conexión	Signo del peso de una conexión
Efecto combinado de la sinápsis	Función de propagación

- Al finalizar el período de procesamiento, la información llega a las unidades de salida, cuya misión es dar la respuesta al sistema.

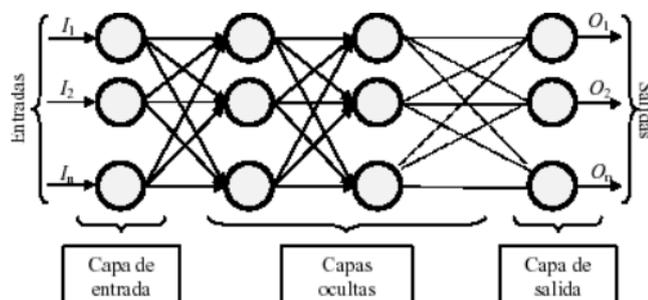


Figura 7 Esquema de una red neuronal 3 capas.

Fuente: (S.Y. Kung, 2012)

La Figura 8 muestra el esquema de neuronas interconectadas y arregladas en tres capas y se ilustra como los datos ingresan por la capa de entrada, pasan a través de la capa oculta y salen por la capa de salida⁴¹

Cabe mencionar que la capa oculta puede estar constituida por varias capas a su vez.

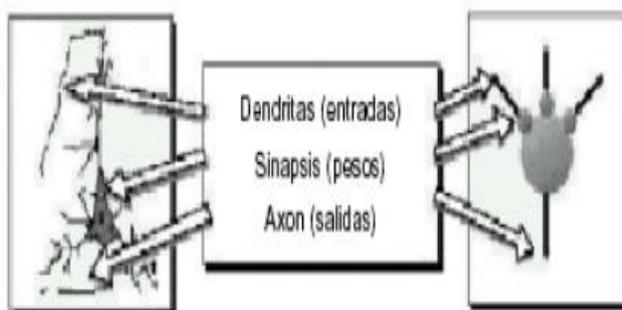


Figura 8 Comparación neurona biológica vs neurona artificial

Fuente: (S.Y. Kung)

De la comparación en la Figura 9 se obtiene que una neurona biológica con una neurona artificial, en la misma se observen similitudes como entradas, pesos y salidas.

Cada neurona i -ésima está caracterizada en cualquier instante por un valor numérico denominado valor o estado de activación $a_i(t)$ asociado a cada unidad, existe una función de salida f_i que transforma el estado actual de activación en una señal de salida. Dicha señal es enviada a través de los canales de comunicación unidireccionales a otras unidades de la red; en estos canales la señal se modifica de acuerdo con la sinápsis el peso w_{ji} asociada a cada uno de ellos según determinada regla. Las señales moduladas que han llegado a la unidad j -ésima se combinan entre ellas, generando así la entrada total Net_j .

$$Net_j = \sum_i y_i W_{ji} \quad (3.1)$$

Una función de activación, F , determina el nuevo estado de activación $a_j(t+1)$ de la neurona, teniendo en cuenta la entrada total calculada y el anterior estado de activación $a_j(t)$

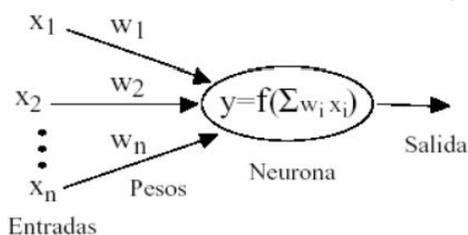


Figura 9 Representación función f neuronas interconectadas

Fuente: Bonifacio Martín y Alfredo Sanz,

3.2 Regla de aprendizaje

La red retro propagada trabaja bajo aprendizaje supervisado y por tanto necesita un conjunto de entrenamiento que le describa cada salida y su valor de salida,

$$(p_1, t_1), (p_2, t_2), \dots, (p_q, t_q) \quad (3.2)$$

Donde p_q es una entrada a la red y t_q es la correspondiente salida deseada para el patrón q -ésimo.

El algoritmo debe ajustar los parámetros de la red para minimizar el error cuadrático medio.

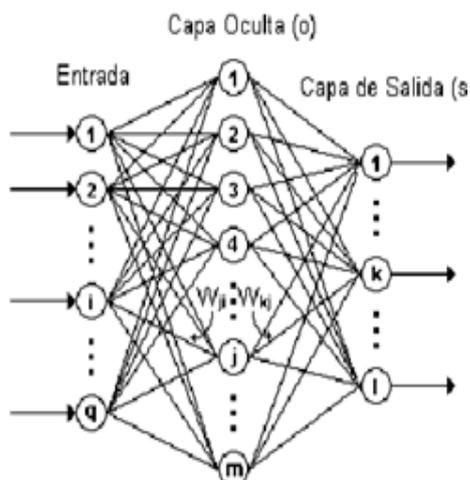


Figura 10 Red retropropagada bajo aprendizaje supervisado

Fuente: (URJC, 1994)

3.3 Modelo perceptrón

Un Perceptrón consiste en una neurona que es capaz de aprender una función discriminante lineal $fd(X)$, que permite dividir a dos conjuntos de entrenamiento linealmente separables.

- Su respuesta consiste una suma ponderada de sus entradas:

$$fd(X) = \sum_{i=1}^n w_i x_i + w_{n+1} \quad (3.3)$$

Que representa la ecuación de un hiperplano en el espacio patrón n -dimensional.

- La salida depende del signo de $fd(X)$. A la salida se aplica una función de activación (escalón, sigmoide...) que indica si se activa o no la neurona.

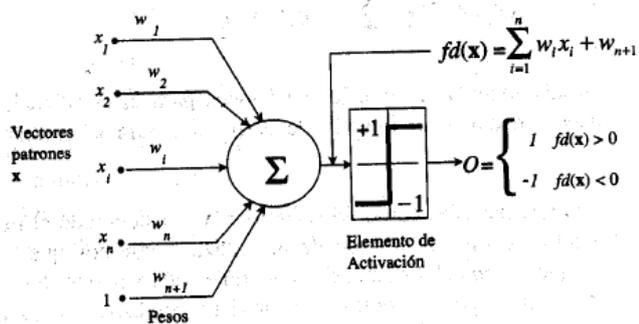


Figura 11 Modelo Perceptrón

Fuente: (URJC, 1994)

Separación de dos clases (regiones) con un Perceptrón:

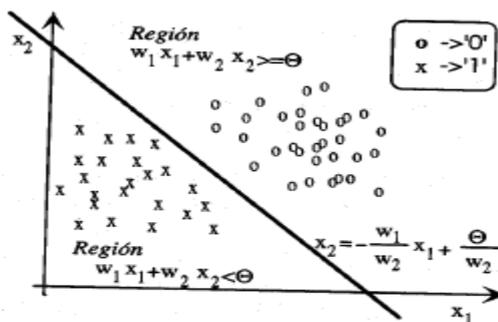


Figura 12 Modelo Perceptrón

Fuente: (URJC, 1994)

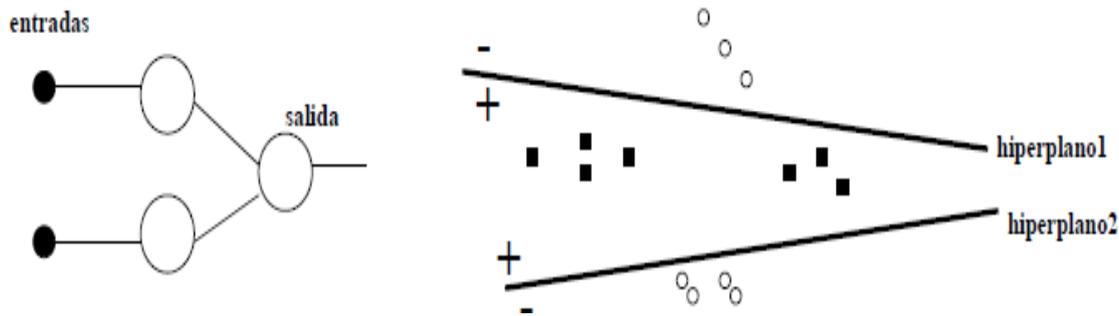


Figura 13 Frontera de decisión perceptrón de dos capas

Fuente: (URJC, 1994)

3.3.1 Perceptrón multicapa mpl (redes feedforward)

Las redes Neuronales de arquitectura MLP, típicamente son definidas con la especificación del número de capas, tipo de función de activación de cada unidad, y de los pesos de conexiones sinápticas entre las diferentes unidades (nodos) que componen la arquitectura neuronal. La Figura 14 ilustra la arquitectura MLP con múltiples entradas y salidas.

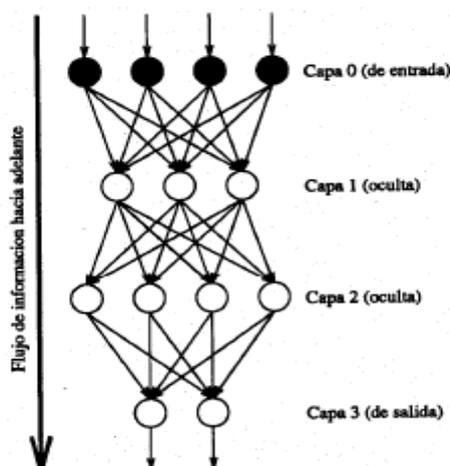


Figura 14 Estructura general perceptrón multicapa(de 3 capas)

Fuente: (URJC, 1994)

Network structure	Type of decision region	Solution to exclusive-OR problem	Classes with meshed regions	Most general decision surface shapes
Single layer 	Single hyperplane			
Two layers 	Open or closed convex regions			
Three layers 	Arbitrary (complexity limited by the number of nodes)			

Figura 15 Decisión reconocidas por distintos tipos de perceptrones

Fuente: (URJC, 1994)

Una vez entrenada la red, usando un algoritmo de aprendizaje (p.ej. quickprop) y un conjunto de patrones de entrenamiento, esta red será capaz de resolver el problema para diferentes patrones desconocidos.

3.4 Algoritmo NN Backpropagation

Rumelhart, Hinton y Williams en 1986, formalizan un método para que una red neuronal aprendiera la asociación que existe entre los patrones de entrada y las clases correspondientes, utilizando varios niveles de neuronas. La propagación hacia atrás de errores o retropropagación (Backpropagation) es un algoritmo de aprendizaje supervisado (técnica para deducir una función a partir de datos de entrenamiento) que se usa para entrenar redes neuronales artificiales.

El método Backpropagation basado en la generalización de la regla delta, a pesar de sus limitaciones, ha ampliado de forma considerable el rango de aplicaciones de las redes neuronales, *El funcionamiento de la red Backpropagation (BPN)* consiste en el aprendizaje de un conjunto predefinido de pares de entradas-salidas dados

Así por Ej. empleando un ciclo de programación-adaptación de dos fases

Primero se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red, *se va propagando* a través de todas las capas superiores hasta generar una salida, *se compara* el resultado en las neuronas de salida con la salida que se desea obtener y se calcula un valor de error para cada neurona de salida, *Posteriormente*, estos errores se transmiten hacia atrás, partiendo de la capa de salida hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida.

Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total.

Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presente el mismo patrón, la salida esté más cercana a la deseada.

La importancia de la red Backpropagation consiste en su *capacidad de autoadaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe ente un conjunto de patrones de entrada y sus salidas correspondientes.*

Es importante la capacidad de *generalización*, facilidad de dar salidas satisfactorias a entradas que el sistema no ha visto nunca en su fase de entrenamiento, La red debe encontrar una representación interna que le permita generar las salidas deseadas cuando se le dan entradas de entrenamiento, y que pueda aplicar, además, a entradas no presentadas durante la etapa de aprendizaje para clasificarlas.

3.4.1 Estructura y aprendizaje de la Red Backpropagation

En la red Backpropagation existe una capa de entrada con **n** neuronas y una capa de salida con **m** neuronas y al menos una capa oculta de neuronas internas.

- Cada neurona de una capa (excepto las de entrada) recibe entradas de todas las neuronas de la capa anterior y envía su salida a todas las neuronas de la capa posterior (excepto las de salida).
- No hay conexiones hacia atrás feedback ni laterales entre las neuronas de la misma capa.
- La aplicación del algoritmo tiene dos fases, una hacia delante y otra hacia atrás.
- *Durante la primera fase* el patrón de entrada es presentado a la red y propagado a través de las capas hasta llegar a la capa de salida.

- Obtenidos los valores de salida de la red, se inicia la segunda fase, comparándose éstos valores con la salida esperada para así obtener el error. Se ajustan los pesos de la última capa proporcionalmente al error.
- Se pasa a la capa anterior con una retropropagación del error, ajustando los pesos y continuando con este proceso hasta llegar a la primera capa. De esta manera se han modificado los pesos de las conexiones de la red para cada patrón de aprendizaje del problema, del que conocíamos su valor de entrada y la salida deseada que debería generar la red ante dicho patrón.

La técnica Backpropagation requiere el uso de neuronas cuya función de activación sea continua, y por lo tanto, diferenciable. Generalmente, la función utilizada será del tipo sigmoideal.

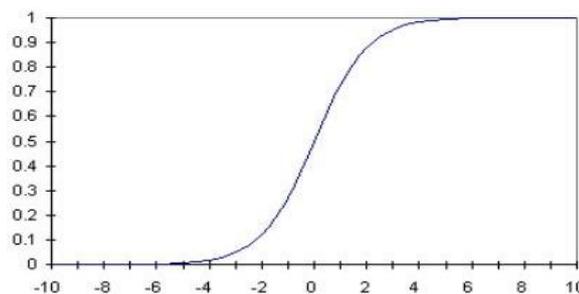


Figura 16 Función de transferencia, representación sigmoideal

Fuente: (D. Michie, D.J. Spiegelhalter, C.C. Taylor, 2007)

El algoritmo NN Backpropagation emplea un ciclo *propagación, adaptación* constituida por dos fases, una vez que se ha aplicado un patrón a la entrada de la red como estímulo, éste se propaga desde la primera capa a través de las capas superiores de la red, hasta generar una salida.

La señal de salida se compara con la salida deseada y se calcula una señal de error para cada una de las salidas. Las salidas de error se propagan hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa oculta que contribuyen directamente a la salida.

Sin embargo las neuronas de la capa oculta solo reciben una fracción de la señal total del error, basándose aproximadamente en la contribución relativa que haya aportado cada neurona a la salida original, este procedimiento se repite capa por capa, hasta que todas las neuronas de la red hayan recibido una señal de error.

A medida que se entrena la red, las neuronas de las capas intermedias se organizan a sí mismas de tal modo que las distintas neuronas aprenden a reconocer las características de las entradas. Después del entrenamiento, cuando se les presente un patrón arbitrario de entrada que contenga ruido o que esté incompleto, las neuronas de la capa oculta de la red responderán con una salida activa si la nueva entrada contiene un patrón que se asemeje a aquella característica que las neuronas individuales hayan aprendido a reconocer durante su entrenamiento.

Sin embargo, actualmente, cuando “atacamos” un problema de predicción, clasificación o de tendencias, en base a un histórico de entradas y salidas y una red neuronal, uno de los principales problemas que tenemos es determinar la estructura de la red neuronal.

Es decir, debemos hacer el trabajo que hace el cerebro ante un problema, sólo que al revés, primero ideamos qué o cuáles son las estructuras de redes neuronales que pueden, quizás, solucionarnos el problema, y luego pasamos a una fase de ensayo-error, hasta que ajustamos nuestra red.

De esta forma, si se decide que la red debe ser retroalimentada, con aprendizaje Backpropagation, es imprescindible “probar” con distintas capas intermedias, distintos tipos de activación, alternamos entre redes recurrentes, de base radial entre otras hasta que los resultados sean los óptimos.

3.4.2 Pasos para aplicar el algoritmo de entrenamiento nn Backpropagation

- Inicializar los pesos de la red con valores pequeños aleatorios.
- Presentar un patrón de entrada y especificar la salida deseada que debe generar la red. Paso
- Calcular la salida actual de la red. Para ello se presentan las entradas a la red y se calcula la salida de cada capa hasta llegar a la capa de salida, ésta será la salida de la red.

Los pasos son los siguientes: Se calculan las entradas netas para las neuronas ocultas procedentes de las neuronas de entrada. Para una neurona j oculta:

$$net_{pj}^h = \sum_{i=1}^N w_{ji}^h x_{pi} + \theta_j^h \quad (3.4)$$

En donde el índice h se refiere a magnitudes de la capa oculta; el subíndice p , al p -ésimo vector de entrenamiento, y j a la j -ésima neurona oculta. El término θ puede ser opcional, pues actúa como una entrada más. Se calculan las salidas de las neuronas ocultas:

$$y_{pj} = f_j^h(net_{pj}^h) \quad (3.5)$$

Se realizan los mismos cálculos para obtener las salidas de las neuronas de salida:

$$net_{pk}^o = \sum_{j=1}^L w_{kj}^o y_{pj} + \theta_k^o \quad (3.6)$$

$$y_{pk} = f_k^o(net_{pk}^o)$$

- Calcular los términos de error para todas las neuronas. Si la neurona k es una neurona de la capa de salida, el valor de la delta es:

$$\delta_{pk}^0 = (d_{pk} - y_{pk}) f_k^{o'}(net_{pk}^o) \quad (3.6.1)$$

La función f debe ser derivable. En general disponemos de dos formas de función de salida: La función lineal:

$$f_k(net_{jk}) = net_{jk} \quad (3.6.1)$$

La función sigmoideal:

$$f_k(net_{jk}) = \frac{1}{1 + e^{-net_{jk}}} \quad (3.7)$$

La selección de la función depende de la forma que se decida representar la salida: si se desea que las neuronas de salida sean binarias, se utiliza la función sigmoideal, en otros casos, la lineal. Para una función lineal, tenemos:

$$f_k^{o'} = 1 \quad (3.8)$$

Mientras que la derivada de una función sigmoideal es:

$$f_k^{o'} = f_k^o(1 - f_k^o) = y_{pk}(1 - y_{pk}) \quad (3.9)$$

Por lo que los términos de error para las neuronas de salida quedan:

$$\delta_{pk}^o = (d_{pk} - y_{pk}) \quad \text{Para la salida lineal} \quad (3.10)$$

$$\delta_{pk}^o = (d_{pk} - y_{pk}) y_{pk}(1 - y_{pk}) \quad \text{Para la salida sigmoideal} \quad (3.11)$$

Si la neurona j no es de salida, entonces la derivada parcial del error no puede ser evaluada directamente, por tanto se obtiene el desarrollo a partir de valores que son conocidos y otros que pueden ser evaluados.

La expresión obtenida en este caso es:

$$\delta_{pj}^h = f_j^{h'}(net_{pj}^h) \sum_k \delta_{pk}^0 w_{kj}^0 \quad (3.12)$$

Donde observamos que el error en las capas ocultas depende de todos los términos de error de la capa de salida. De aquí surge el término propagación hacia atrás.

- Actualización de los pesos: para ello utilizamos un algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atrás hasta llegar a la capa de entrada, ajustando los pesos de la siguiente forma:

Para los pesos de las neuronas de la capa de salida:

$$\begin{aligned}w_{kj}^o(t+1) &= w_{kj}^o(t) + \Delta w_{kj}^o(t+1) \\ \Delta w_{kj}^o(t+1) &= \alpha \delta_{pk}^o y_{pj}\end{aligned}\tag{3.13}$$

Para los pesos de las neuronas de la capa oculta:

$$\begin{aligned}w_{ji}^h(t+1) &= w_{ji}^h(t) + \Delta w_{ji}^h(t+1) \\ \Delta w_{ji}^h(t+1) &= \alpha \delta_{pj}^h x_{pi}\end{aligned}\tag{3.14}$$

En ambos casos, para acelerar el proceso de aprendizaje se puede añadir un término momento.

- El proceso se repite hasta que el término de error

$$E_p = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2\tag{3.15}$$

Resulta aceptablemente pequeño para cada uno de los patrones aprendidos

3.4.3 Número de capas ocultas

No se pueden dar reglas concretas para determinar el número de neuronas o número de capas de una red para resolver un problema concreto, respecto al número de capas de la red, en general tres capas son suficientes (entrada –oculta-salida). Sin embargo, hay veces que un problema es fácil de resolver con más de una capa oculta.

El tamaño de las capas, tanto de entrada como de salida, suele estar determinado por la naturaleza de la aplicación. En cambio, decidir cuántas neuronas debe tener una capa oculta no suele ser tan evidente. El número de neuronas ocultas interviene en la eficiencia de aprendizaje y de generalización de la red. No hay ninguna regla que indique el número óptimo, en cada problema

3.5 Aplicación de NN Backpropagation para predicción de variables meteorológicas.

Los modelos de predicción de series temporales proporcionan valores futuros de la serie, a partir de los valores presentes y pasados. En estos modelos se pretende establecer la dinámica de un proceso observado a través del conjunto de datos, que constituye la serie temporal.

Para ello el modelo ha de ser capaz de describir las relaciones subyacentes en las observaciones pasadas, para posteriormente extrapolar en el tiempo las series de datos. En la actualidad las redes neuronales pueden ser utilizados para llevar a cabo pronósticos en diversos grados de éxito, las ANN depende únicamente de datos históricos de entrenamiento en la expectativa de que dependencias ocultas se descubren que se puede utilizar para realizar predicciones futuras

En otras palabras, es ANN no representado como un modelo explícito. Más bien, es un cuadro negro que es capaz de aprender. La ventaja de utilizando ANN para la predicción es que se aprende de los Ej. s, y después de que aprende que es capaz de generalizar como que encuentra patrones no lineales ocultos, incluso cuando hay ruido en la colección de Ej. s de entrenamiento

Diversas han sido las investigaciones aplicando ANN para predicción en variables meteorológicas como las indicadas en el apartado 2.2 en la que

indica que se pueden encontrar metodologías diversas, en la que las RNAs se aplican a la predicción de series temporales climatológicas [17]

Las RNAs no son capaces de capturar variaciones estacionales y tendencias de forma efectiva, si no se realiza previamente un tratamiento de datos encaminado a lograr una serie estacionaria y desestacionalizada. Dorffner también propone la des estacionalización de las series en una fase previa a la aplicación del modelo [18]

(Labajo,A.L, Martín,2014) Docentes investigadores de la Universidad de Salamanca refieren el uso de una metodología experimental para configurar los parámetros de un modelo de predicción de variables climáticas.

Esta predicción usa una red neuronal artificial (RNA) de tipo (Multilayer Perceptron, MLP) Perceptrón multicapa, para la configuración del modelo emplean las series de datos pertenecientes a la red sinóptica y climatológica del Instituto Nacional de Meteorología (INM) de España.

Para optimizar la topología de la red neuronal y su rendimiento se realiza un tratamiento de datos previo a la aplicación del modelo. Debido a que el modelo resultante se aplica a las series de datos de la variable temperatura mínima media mensual observadas en las estaciones de la red sinóptica y climatológica del INM en la meseta central española, proporcionando un alto grado de ajuste entre series reales y simuladas, según indican los valores de los coeficientes de determinación (R^2) y error cuadrático medio (MSE) y las gráficas de dispersión y secuencia de las series reales y simuladas.

3.5.1 Metodología utilizada

La RNA de tipo MLP adaptado de tres capas se usó como base, como se indica en la Figura 17.

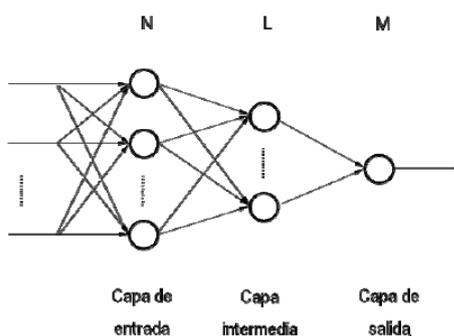


Figura 17 RNA MLP con tres capas de neuronas

Fuente: (Universidad de Salamanca, 2007)

Las funciones de activación tipo rampa en la capa de salida y tangente hiperbólica en las capas de entrada y oculta.

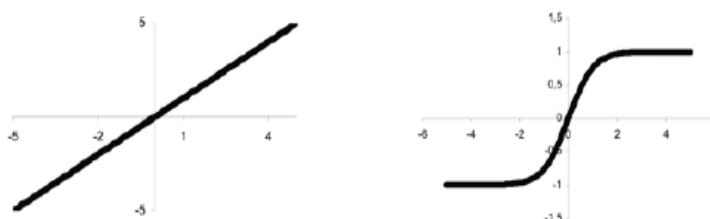


Figura 18 Funciones de activación rampa y sigmoide

Fuente: (Universidad de Salamanca, 2007)

El modelo empleado fue un modelo de predicción en un paso de tiempo (Figura 19), en el que la estimación en el instante de tiempo posterior se obtiene a partir de los valores anteriores de la serie de datos. Los patrones de entrenamiento y prueba fueron partes de la serie temporal, de la misma longitud y desfasados en un instante de tiempo.

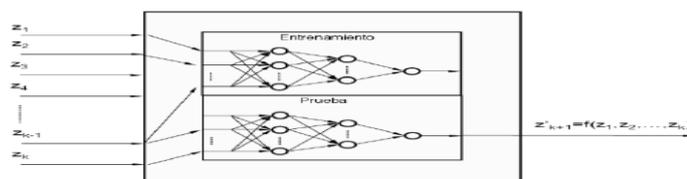


Figura 19 Esquema del RNA de predicción de series temporales

Fuente: (Labajo, A.L.; Martín, Q. y Labajo, 2010)

El modelo propuesto, entre otros, por Isasi y Galván indicado en la Figura 21 se obtuvo según se indica en la tabla 3. El modelo aplica a las fases entrenamiento y prueba, En la fase de entrenamiento la RNA se entrena con el algoritmo de retro propagación (Backpropagation, BP) y algunas de sus variantes. En la etapa de prueba se valida el modelo a partir de los valores de R2 y MSE entre series reales y simuladas y las correspondientes gráficas de secuencia y dispersión.

En la Figura 20 se muestra el esquema del proceso. El ajuste del modelo se realizó con las series observadas en la estación meteorológica de Ávila y se partió de la configuración inicial que se muestra en la tabla 3. Obteniendo el modelo de predicción como se muestra en la Figura 20.

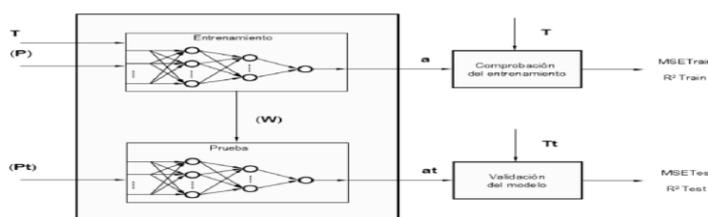


Figura 20 Modelo de pronóstico

Fuente: (Labajo, A.L.; Martín, Q. y Labajo, 2010)

Tabla 3

Configuración de los patrones de entrenamiento y prueba según el esquema de predicción de un paso de tiempo (N=no de datos por patrón)

Patrón	Entrada	Salida deseada
Entrenamiento 1	$t(1), t(2), \dots, t(N)$	$t(N+1)$
Entrenamiento 2	$t(2), t(3), \dots, t(N+1)$	$t(N+2)$
⋮	⋮	⋮
Entrenamiento NTr	$t(NTr), t(NTr+1), \dots, t(NTr+N-1)$	$t(NTr+N)$
Prueba 1	$t(NTr+1), t(NTr+2), \dots, t(NTr+N)$	$t(NTr+N+1)$
Prueba 2	$t(NTr+2), t(NTr+3), \dots, t(NTr+N+1)$	$t(NTr+N+2)$
⋮	⋮	⋮
Prueba NTe	$t(NTr+NTe), t(NTr+NTe+1), \dots, t(NTr+NTe+N-1)$	$t(NTr+NTe+N)$

Posterior a los experimentos el modelo fue alimentado por patrones de datos en los que la serie fue previamente procesada mediante diferenciación, desestacionalización, anomalías, normalización y estandarización.

El tratamiento previo de los datos obligó a realizar una etapa posterior de tratamiento inverso de datos antes de poder validar los resultados, según se muestra en el esquema de la Figura 21

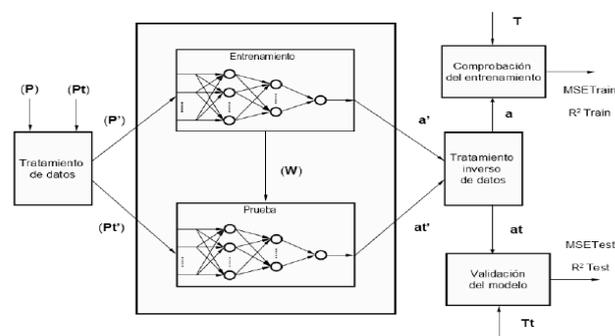


Figura 21 Modelo de pronóstico con preprocesamiento de las series de datos.

Fuente: (Labajo, A.L.; Martín, Q. y Labajo,2010)

3.5.2 Algoritmo de retro propagación predictivo

Corchado et al muestran el desarrollo del algoritmo con dos etapas en el aprendizaje.

En la primera los patrones de datos de entrenamiento se presentan a la capa de entrada a la red. (Estos datos se propagan desde la capa de entrada a la de salida, en la que se genera una salida de la RNA). En la segunda etapa, el resultado que proporciona la RNA se compara con el resultado esperado para cada uno de los patrones de entrenamiento. Si no coinciden, el error se utiliza para modificar los pesos de las capas intermedias.

La regla de actualización se expresa de la siguiente forma:

$$w_{ik}(t+1) = w_{ik}(t) + \Delta_p w_{ik}(t) + \mu(w_{ik}(t) - w_{ik}(t-1))$$

$$\Delta_p w_{ik} = -\eta \frac{\partial E^p}{\partial w_{ik}}, \quad E^p = \frac{1}{2} \sum_{k=1}^M (a_k^p - y_k^p)^2 \quad (3.16)$$

$$\vec{x}^p = (x_1^p, x_2^p, \dots, x_N^p)^T, \quad \vec{a}^p = (a_1^p, a_2^p, \dots, a_M^p)^T$$

Dónde:

w_{ik} : Es el peso asociado a la conexión entre la neurona i-ésima de una capa y la k-ésima de la siguiente,

E_p es la función de error

x_p el p-ésimo vector de entrada a la red, a_p , el p-ésimo vector de salida esperado, la tasa de entrenamiento y el momento.

3.5.3 Configuración del modelo predictivo

Los experimentos de configuración del modelo se agrupan según diversos parámetros que se pretenden ajustar, a cada experimento se le ha asignado una referencia por una mayor claridad en la exposición de los resultados. En las Figura 22 se muestran las gráficas de secuencias correspondientes a las etapas de validación del modelo.

En las gráficas de secuencia se ha representado la temperatura (°C) frente la fecha (mes y año). La línea continua representa la secuencia real y la discontinua la simulada por el modelo.

Tabla 4

Valores del coeficiente de determinación MSE, en el entrenamiento y la prueba, para el conjunto de experimentos P

Ref.	N	NTr	NTE	Entrenamiento		Prueba	
				R ²	MSE	R ²	MSE
P1	180	432	60	0,880	3,42	0,844	4,38
P2	144	468	60	0,592	11,56	0,645	9,80
P3	120	492	60	0,931	1,96	0,899	2,87
P4	96	516	60	0,919	2,31	0,884	3,26
P5	60	552	60	0,304	19,70	0,328	18,41

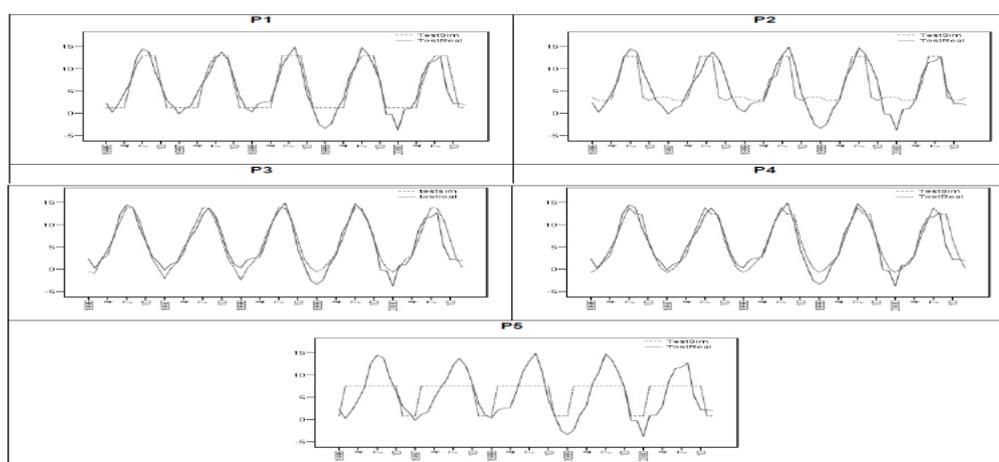


Figura 22 Evolución temporal de las series

Fuente: (Labajo, A.L.; Martín, Q. y Labajo,2010)

El conjunto de experimentos P realizados permiten definir la configuración de la longitud de los patrones (N), el número de patrones de entrenamiento (NTr) y el número de patrones de prueba (NTE).

Los experimentos A, cuyo objetivo es determinar el algoritmo de entrenamiento con el que el modelo proporciona mejores resultados, de entre el BP estándar y sus variantes. En los experimentos A1 a A5 el algoritmo se entrenó durante 20.000 iteraciones, valor que se consideró lo suficientemente grande como para conseguir un grado de ajuste muy elevado entre las series reales y simuladas en la fase de entrenamiento. Los resultados obtenidos se muestran en la tabla 54. En algunos casos se obtiene un entrenamiento casi perfecto, situación que no se mantiene en la fase de prueba. Por lo tanto se deja constancia de un problema de *sobre-entrenamiento*

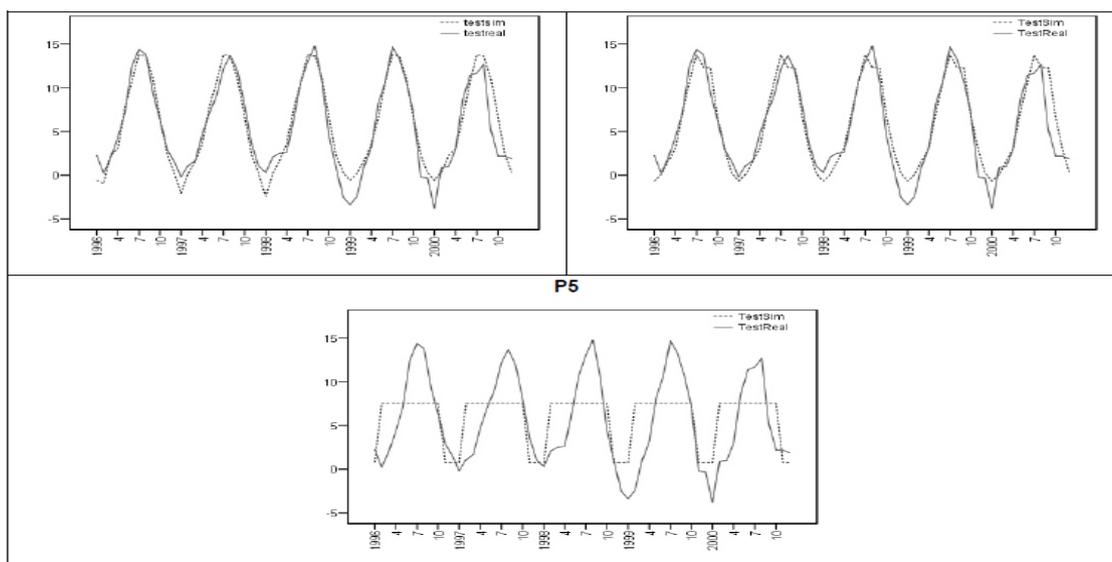


Figura 23 Evolución temporal de las series reales RNA

Fuente: (Labajo, A.L.; Martín, Q. y Labajo,2010)

A continuación se realiza el conjunto de experimentos A, cuyo objetivo es determinar el algoritmo de entrenamiento con el que el modelo proporciona mejores resultados, de entre el BP estándar y sus variantes, con el grado de ajuste muy elevado entre las series reales y simuladas en la fase de entrenamiento

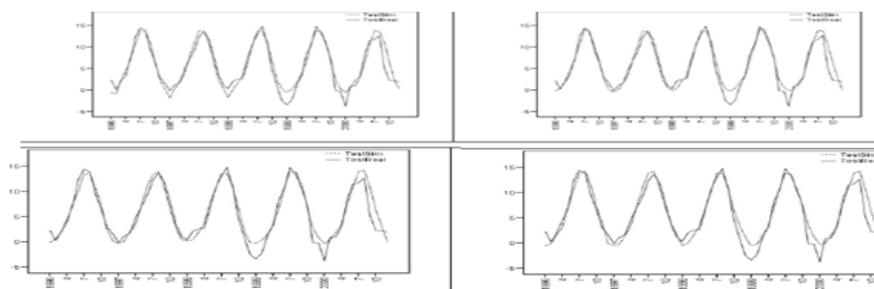


Figura 24 Evolución temporal de las series reales RNA para experimentos A

Fuente: (Labajo, A.L.; Martín, Q. y Labajo,2010)

3.5.4 Resultados del modelo predictivo

Para comprobar las capacidades del modelo de pronóstico desarrollado, se comprobó con las series observadas en estaciones de la meseta central española. El periodo de estudio se amplió hasta diciembre de 2014, gracias

a la disponibilidad de datos en el momento en que se realizó el citado estudio. Al igual que en los apartados anteriores se muestran las tablas de resultados y las gráficas de secuencia correspondientes para Castilla Ávila, Burgos y Salamanca [19] los resultados se muestran en la Figura 25

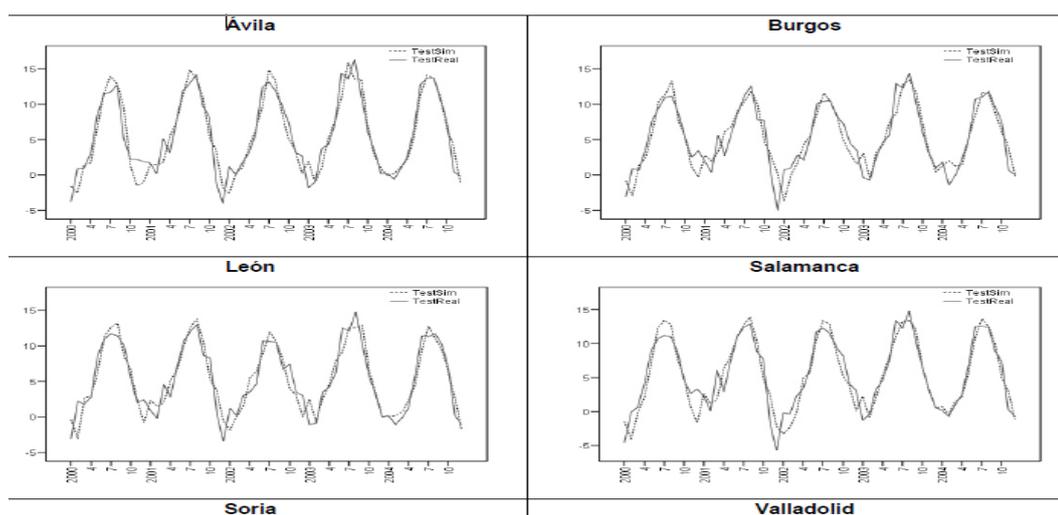


Figura 25 Gráficas de secuencias para Castilla Ávila Burgos

Fuente: (Labajo, A.L.; Martín, Q. y Labajo, 2010)

La aplicación del modelo a las series de datos observadas en Castilla la Mancha presenta resultados algo mejores, exceptuando también el caso de Castilla y León.

Las gráficas de secuencia, que se muestran en las Figura 26, dejan constancia de un alto grado de concordancia entre las series de valores reales y previstos por el modelo, prácticamente en todos los casos. Las series de valores reales y de valores simulados por el modelo presentan las mismas características generales de evolución y tendencia.

3.5.5 Conclusiones

La RNA MLP adaptada para este trabajo, de tres capas y con funciones de activación como las descritas en el apartado III, demuestra su capacidad para modelizar de forma no lineal los valores futuros de una serie temporal a

partir únicamente de valores pasados de la misma serie temporal, en relación a la presente investigación en donde la predicción de la radiación solar pretende basarse en históricos pasados de la misma serie temporal.

La aplicación del modelo RNA a los datos de las series de trabajo presenta un efecto de agrupación de los valores de salida del modelo RNA, efecto que se acentúa en los valores extremos y que se observa con facilidad en las gráficas de dispersión correspondientes. Este efecto se traduce en que los valores extremos de la serie simulada por el modelo RNA varían muy poco a lo largo del tiempo. Este efecto se modifica en función de los tratamientos previos de las series de trabajo [20]. Las redes neuronales biológicas refieren un cúmulo de neuronas físicamente interconectadas para ayudar a definir un circuito reconocible en el sistema nervioso, y las redes artificiales, son modelos matemáticos y computacionales de una red neuronal empleados en estadística cognitiva e inteligencia artificial.

3.6 Máquinas de soporte vectorial (SVM)

Las máquinas de soporte vectorial (SVM) son un conjunto de algoritmos de aprendizaje supervisado desarrollados por Vladimir Vapnik y su equipo en los laboratorios AT&T. El método se basa en el uso de programación matemática, formulada de forma que la interpretación estadística del modelo resulta particularmente apropiada y sustentada por las teorías estadísticas de aprendizaje propuestas. Estos métodos están propiamente relacionados con problemas de *clasificación y regresión*. Dado un conjunto de Ej. s de entrenamiento (de muestras) se puede etiquetar las clases y entrenar una SVM para construir un modelo que prediga la clase de una nueva muestra[11]

Una SVM construye un hiperplano o conjunto de hiperplanos en un espacio de dimensionalidad muy alta que puede ser utilizado en problemas de clasificación o regresión. Una buena separación entre las clases permitirá una clasificación correcta.

Las SVM (o Máquinas de Vectores de Soporte) son máquinas de Aprendizaje que necesitan primero entrenarse con situaciones en las que se les dice la respuesta correcta sobre muchos Ej. Entrenamiento, y una vez entrenados, entra la fase de uso, y se convierte en una caja que devuelve la respuesta ante un nuevo caso. (Es decir es un método de aprendizaje supervisado) [12]. Su importancia radica en la superación a los algoritmos antecesores regresión y clasificación, de tal manera que las SVMs no han sido superadas sino por sus propios algoritmos su aplicación fundamentalmente radica en la predicción de datos una vez entrenada la maquina siendo una predicción de clasificación binaria, multi-categoría y/o de regresión general

3.6.1 La idea básica de las máquinas de soporte vectorial

En la literatura de los SVMs, se llama *atributo* a la variable predictora y *característica* a un atributo transformado que es usado para definir el hiperplano. La elección de la representación más adecuada del universo estudiado, se realiza mediante un proceso denominado selección de características.

Los modelos basados en SVMs están estrechamente relacionados con las redes neuronales. Usando una función kernel, resultan un método de entrenamiento alternativo para clasificadores polinomiales.

Funciones de base radial y perceptrón multicapa El método de aprendizaje en el que se basan las Support Vector Machines (SVM) no es más que la kernelización de un clasificador lineal de margen máximo con holgura. Se aplica un algoritmo para aprender un clasificador lineal de margen máximo con holgura de forma implícita sobre el espacio de características F inducido por la función kernel empleada (en lugar de sobre el espacio de entrada original X)

Así por Ej. Dado un conjunto de puntos, subconjunto de un conjunto mayor (espacio), en el que cada uno de ellos pertenece a una de dos

posibles categorías, un algoritmo basado en SVM construye un modelo capaz de predecir si un punto nuevo (cuya categoría desconocemos) pertenece a una categoría o a la otra.

Dado un conjunto de Ej. s de entrenamiento previamente clasificados

$$L = \{(\vec{x}_1, y_1), (\vec{x}_2, y_2), \dots, (\vec{x}_l, y_l)\} \quad (3.17)$$

Una cota máxima de pérdidas C y una función kernel $K(x,y)$ encontrar los valores

$$\alpha_1, \alpha_2, \dots, \alpha_l \quad (3.18)$$

Que maximicen

$$\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \underline{\underline{k(\vec{x}_i, \vec{x}_j)}} \quad (3.19)$$

Sujeto a:

$$\sum_{i=1}^l \alpha_i y_i = 0 \quad (3.20)$$

$k(\vec{x}_i, \vec{x}_j)$ equivale al producto escalar, $\Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j)$, en \mathcal{F} .

Es decir, la optimización $\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j k(\vec{x}_i, \vec{x}_j)$ se aplica realmente de forma implícita sobre $\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j)$

$$0 \leq \alpha_i \leq C \quad \forall i \in \{1, \dots, l\}$$

Como en la mayoría de los métodos de clasificación supervisada, los datos de entrada (los puntos) son vistos como un vector p -dimensional (una lista de p números), La SVM busca un hiperplano que separe de forma óptima a los puntos de una clase de la de otra, que eventualmente han podido ser previamente proyectados a un espacio de dimensionalidad superior

En ese concepto de "separación óptima" es donde reside la característica fundamental de las SVM: este tipo de algoritmos buscan el hiperplano que tenga la máxima distancia (margen) con los puntos que estén más cerca de él mismo. Por eso también a veces se les conoce a las SVM como *clasificadores de margen máximo*. De esta forma, los puntos del vector

que son etiquetados con una categoría estarán a un lado del hiperplano y los casos que se encuentren en la otra categoría estarán al otro lado.

Los algoritmos SVM pertenecen a la familia de los clasificadores lineales. También pueden ser considerados un caso especial de la regularización de Tikhonov.

3.6.2 Método kernel

Los métodos kernel son algoritmos para el análisis de patrones cuyo miembro más conocido es la máquina de soporte vectorial (SVM), la tarea general de análisis de patrones es encontrar y estudiar los tipos generales de las relaciones (Por Ej., las agrupaciones, las clasificaciones, componentes)

Para muchos algoritmos que resuelven estas tareas, los datos de representación en bruto tienen que ser transformados de forma explícita en representaciones característica vector a través de un mapa de características especificado por el usuario: por el contrario, núcleo métodos requieren sólo un kernel especificado por el usuario, es decir, una función de similitud sobre pares de puntos de datos en la representación cruda, Varios de los algoritmos del núcleo se basan en la optimización convexa y son estadísticamente bien fundados, sus propiedades estadísticas se analizan usando la teoría del aprendizaje estadístico[13]

La manera más simple de realizar la separación es mediante una línea recta, un plano recto o un hiperplano N-dimensional. Un algoritmo SVM debe tratar con

- Más de dos variables predictoras
- Curvas no lineales de separación
- Casos donde los conjuntos de datos no pueden ser completamente separados

Perceptrón

$$K(x_i, x_j) = \|x_i - x_j\| \quad (3.23)$$

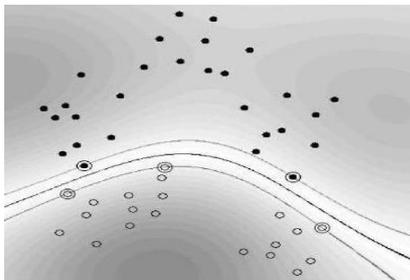


Figura 27 Función Kernel: Polinomial-homogénea

Fuente: Labajo, A.L.; Martín, Q. y Labajo⁶⁹

- Función de base radial Gaussiana: separado por un hiperplano en el espacio transformado.

$$K(x_i, x_j) = \exp(-(x_i - x_j)^2 / 2(\sigma)^2) \quad (3.24)$$

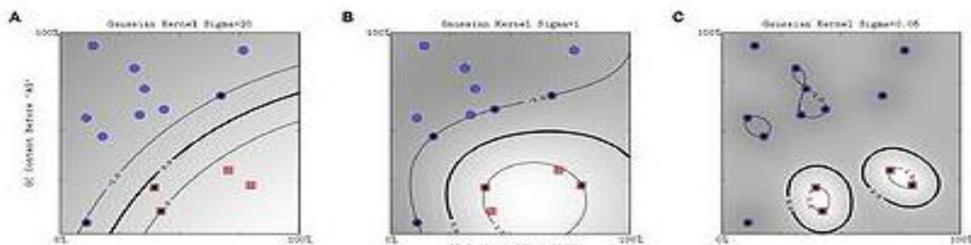


Figura 28 Función de base radial Gaussiana

Fuente: Labajo, A.L.; Martín, Q. y Labajo⁶⁹

3.8 Aplicaciones de las funciones kernel

De entre las múltiples técnicas de predicción existentes Las SVM se han introducido hace poco como técnica de aprendizaje automático por su amplia aplicación estadística, y precisamente por ello es una técnica más robusta en diferentes ámbitos especialmente porque incluye predicciones, aspectos y técnicas de aprendizaje automático, estadística y análisis matemático.

Existen diversos trabajos para predicciones en series no lineales es así que para una resolución temporal horaria y diaria y en localidades de España.se puede ver una comparación de los resultados de diferentes modelos MVS En la presente investigación se analiza la técnica IA, SVM para predicción de la radiación solar.

En el trabajo realizado en la Escuela Técnica Superior De Ingeniería Informática de Rafael Mena Yedra Madrid 2014 realiza el estudio de Modelo de predicción de la radiación solar con SVM como técnica IA.[14]. El objetivo principal del trabajo fue diseñar e implementar un modelo para la predicción a corto plazo (1-12 horas) y a medio plazo (1-14 días) de la radiación solar e indica que las SVMs se han aplicado con éxito en tareas de clasificación [15] regresión y predicción [16], siendo capaces de modelar sistemas multivariable y no lineales como es el caso de la radiación solar.

Al tratarse de un problema de predicción basado en series temporales, es difícil conseguir continuidad a lo largo del tiempo en los datos debido a las interrupciones en los sistemas de adquisición de datos.

Por ello, se plantea una solución para recuperar información perdida a través de la interpolación [17].

(3.24)

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0}$$

De la que despejando y , que es el valor desconocido en el punto x , se obtiene

(3.25)

$$y = y_0 + (x - x_0) \frac{y_1 - y_0}{x_1 - x_0} = y_0 + \frac{(x - x_0)y_1 - (x - x_0)y_0}{x_1 - x_0}$$

Que es la fórmula para la interpolación lineal en el intervalo.

. task 36: Solar resource knowledge management,” tech. rep., IEA. IEA SHC Task 36 Solar Resource Knowledge Management., 2009.

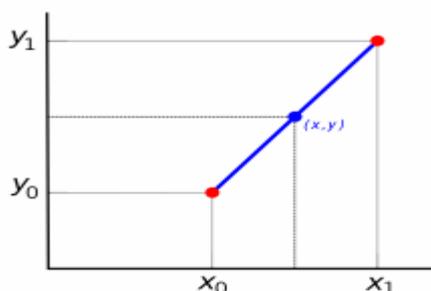


Figura 29 Interpolación Lineal en (x0, x1)

Fuente: (Renné, Semi-annual status report, 2010)

Punto seguido la normalización para escalar los datos en los intervalos y reducir el error del modelo. La fórmula utilizada es:

$$y_i = \frac{(y_{max} - y_{min})(x_i - x_{min})}{(x_{max} - x_{min})} + y_{min} \quad (3.26)$$

Dónde:

$[y_{min}, y_{max}] = [0, 1]$; Intervalo de normalización

$[x_{min}, x_{max}]$; Mínimo y máximo de los datos

x_i ; Valor a normalizar

(3.27)

3.8.1 Procedimiento clasificación

El objetivo preliminar de clasificación SVM es establecer límites de decisión en el espacio característico para separar los puntos de datos en diferentes clases excluyentes. Su intención es crear un hiperplano de separación óptima entre dos clases para minimizar el error de generalización y con ello maximizar el margen.

De esta manera, los objetos de la clase + se encuentran tras el hiperplano H2 mientras que los de la clase o tras el hiperplano H1, y los objetos de cada clase que se sitúan exactamente sobre estos hiperplanos H1 y H2 se denominan vectores de soporte (support vectors), mientras que la distancia entre estos hiperplanos H1 y H2 (δ) es el máximo margen obtenido. La mayoría de los datos de entrenamiento importantes son vectores de soporte puesto que definen el hiperplano y tienen influencia directa en la ubicación óptima de la superficie de decisión.

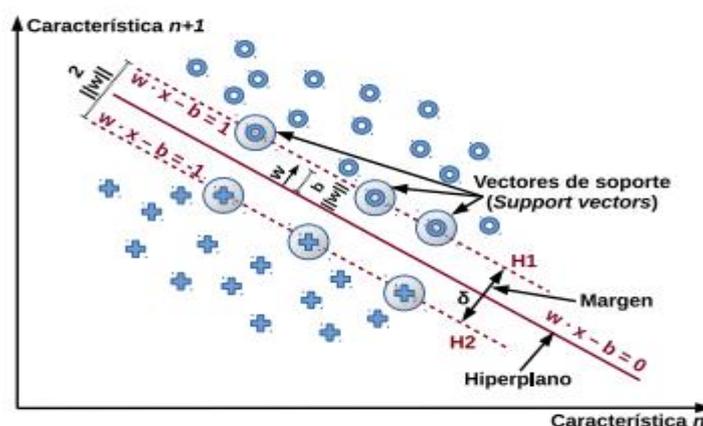


Figura 30 Hiperplano de máxima separación

Fuente: Renné, Semi-annual status report[14]

Más formalmente, dado un conjunto de datos de entrenamiento T formado por n puntos de la forma:

$$T = \{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{-1, 1\}\}_{i=1}^n \quad (3.28)$$

Dónde:

y_i indica la clase ($\{-1, 1\}$) a la que pertenece cada punto x_i y cada punto x_i es un vector p -dimensional en \mathbb{R} . En este espacio, cualquier hiperplano se puede expresar como un conjunto de puntos x que satisfacen

$$w \cdot x - b = 0 \quad (3.29)$$

Si los datos son linealmente separables, se pueden seleccionar dos hiperplanos de forma que separen los puntos de ambas clases sin tener puntos entre ellos y al mismo tiempo maximizando el margen. Estos hiperplanos óptimos se pueden describir por las ecuaciones:

$$\begin{aligned} w \cdot x - b &= 1 \\ w \cdot x - b &= -1 \end{aligned} \quad (3.30)$$

Finalmente, lo que se tiene es un problema de optimización que se puede escribir de la siguiente forma:

$$\begin{aligned} &\text{Minimizar (en } w, b): \\ &\|w\| \\ &\text{sujeto a (para cualquier } i = 1, \dots, n): \\ &y_i(w \cdot x_i - b) \geq 1 \end{aligned} \quad (3.31)$$

Así, el problema puede resolverse mediante técnicas de programación cuadrática estándar. La condición “estacionaria” de Karush-Kuhn-Tucker [71] implica que la solución puede expresarse como una combinación lineal de los vectores de entrenamiento:

$$w = \sum_{i=1}^n \alpha_i y_i x_i \quad (3.32)$$

De aquí se puede derivar que los vectores de soporte también satisfacen:

$$w \cdot x_i - b = \frac{1}{y_i} = y_i \iff b = w \cdot x_i - y_i \quad (3.33)$$

Lo que permite definir el desplazamiento b . En la práctica, resulta más robusto promediar todos los vectores de soporte N_{SV} :

$$b = \frac{1}{N_{SV}} \sum_{i=1}^{N_{SV}} (w \cdot x_i - y_i) \quad (3.34)$$

Por último, expresar el problema de optimización en su forma dual sin restricciones revela que el hiperplano que maximiza el margen, es decir la tarea de clasificación, es sólo una función de los vectores de soporte, esto es el subconjunto de los datos de entrenamiento que se sitúan en el margen.

Maximizar (en α_i):

$$\begin{aligned} \tilde{L}(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum \alpha_i \alpha_j y_i y_j x_i^T x_j \\ &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum \alpha_i \alpha_j y_i y_j k(x_i x_j) \end{aligned} \quad (3.35)$$

sujeto a (para cualquier $i = 1, \dots, n$):

$$\begin{cases} \alpha_i \geq 0, \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases}$$

Así se define el kernel por

$$k(x_i x_j) = x_i \cdot x_j \quad (3.36)$$

y w se puede calcular gracias a los términos de

$$\alpha \cdot w = \sum_i \alpha_i y_i x_i \quad (3.37)$$

3.8.2 Método del kernel

En los problemas de clasificación reales, lo habitual es que una máquina de aprendizaje lineal no pueda determinar un hiperplano con una separación exacta de los datos por las limitaciones.

Entre las limitaciones: existen más de dos variables de entrada, hay curvas no lineales de separación, casos en los que los datos no puedan ser completamente separados o clasificaciones en más de dos categorías.

Para hacer frente a estos casos, el espacio de entradas original puede ser mapeado a un nuevo espacio característico de mayor dimensión (espacio de Hilbert) usando funciones no-lineales denominadas funciones de características ϕ (Figura 31). A pesar de que el espacio característico es de dimensionalidad alta, no sería factible en la práctica realizar directamente la clasificación. Así que en estos casos, el mapeado no lineal inducido por ϕ se utiliza para el cálculo usando funciones no lineales especiales llamadas kernels las más habituales se pueden ver en Ecuación 3.38.

El algoritmo resultante es similar, pero el producto escalar \cdot se reemplaza por la función de kernel no-lineal. Esto permite al algoritmo obtener un hiperplano de máximo margen en el nuevo espacio característico y de esta forma aunque el clasificador sea un hiperplano en el espacio característico de alta dimensionalidad, puede ser no-lineal en el espacio de entrada original Figura 32. La función de kernel K está relacionada con la función ϕ por la ecuación

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j). \quad (3.38)$$

El valor w también está en el nuevo espacio

$$w = \sum_i \alpha_i y_i \phi(x_i). \quad (3.39)$$

Los productos escalares con w para clasificación se pueden calcular mediante el método del kernel

$$w \cdot \phi(x) = \sum_i \alpha_i y_i k(x_i, x). \quad (3.40)$$

$$K(x_i, x_j) = \begin{cases} (x_i \cdot x_j) & \text{Lineal} \\ (x_i \cdot x_j)^d & \text{Polinomial (homogéneo)} \\ (x_i \cdot x_j + 1)^d & \text{Polinomial (no homogéneo)} \\ \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 & \text{Función de base radial (RBF)} \\ \tanh(\kappa x_i \cdot x_j + c), \kappa > 0; c > 0 & \text{Sigmoidea} \end{cases} \quad (3.41)$$

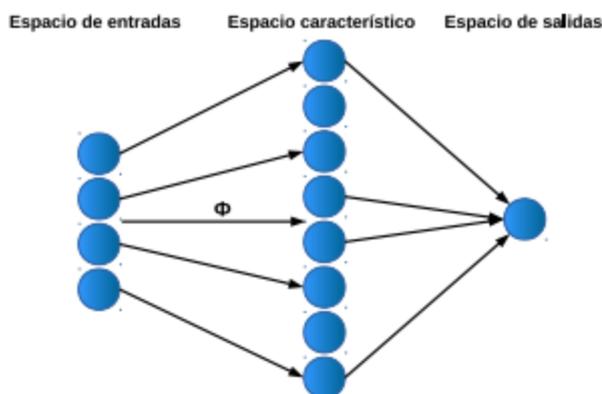


Figura 31 Transformación del espacio de entrada al espacio bcaracterístico de mayor dimensionalidad mediante ϕ

Fuente: (E. Lorenz et. Al, 2010)

3.9 Support vector regression (SVR)

Las SVMs no solo se pueden aplicar a problemas de clasificación, sino también para regresión y estimación de funciones [14].

De la misma manera que en el enfoque visto anteriormente para clasificación, también hay un interés en buscar y optimizar los límites de generalización dados para regresión, estos se basan en la definición de una función de pérdida que ignora los errores que están situados dentro de cierto umbral del verdadero valor.

Este tipo de enfoque se conoce como ϵ -SVR, en la Figura 33 se puede ver un Ej. de función de regresión con la banda ϵ , donde las variables ξ_i miden el coste de los errores en los puntos de entrenamiento siendo cero para aquellos que caen dentro de la banda. En SVR, el espacio de entrada primero se mapea en un espacio característico de mayor dimensión utilizando algún mapeado no-lineal fijado Φ y entonces se construye un modelo lineal en ese espacio característico. La expresión es:

$$f(x) = \sum_{i=1}^n w_i \Phi(x_i) + b \quad (3.42)$$

La SVR estándar utiliza una función de pérdida

$$L_{\epsilon}(y, f(x)) \quad (3.43)$$

La expresión describe la desviación de la función estimada y la original, y se define como:

$$L_{\epsilon}(y, f(x)) = \begin{cases} 0, & \text{si } |y - f(x)| \leq \epsilon \\ |y - f(x)| - \epsilon, & \text{en otro caso} \end{cases} \quad (3.44)$$

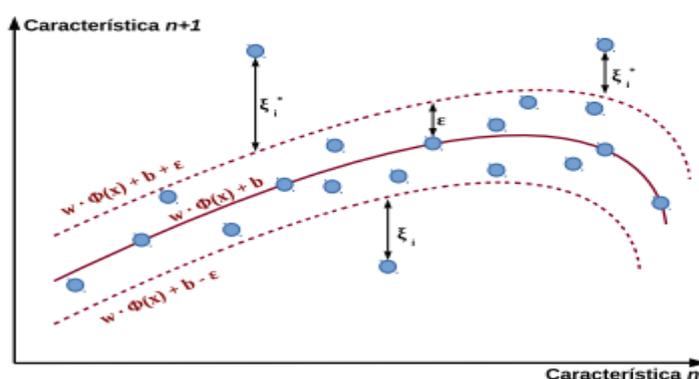


Figura 32 Función de regresión no-lineal con la banda ϵ

Fuente: (E. Lorenz et. Al, 2010)

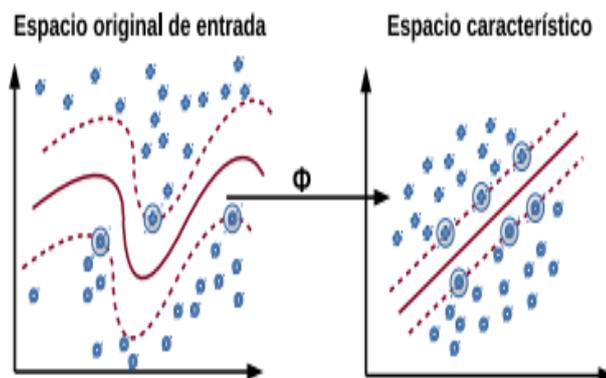


Figura 33 Hiperplano en el espacio tras la transformación mediante ϕ

Fuente: (E. Lorenz et. Al, 2010)

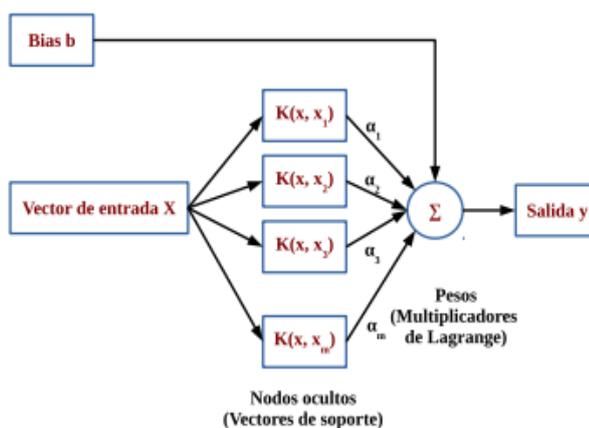


Figura 34 Arquitectura de una SVM

Fuente: (E. Lorenz et. Al, 2010)

Aplicando el método del kernel se obtiene la función de decisión del SVR no-lineal, y cuya arquitectura se puede ver representada en Figura 34:

$$\sum_{i=1}^{N_{SV}} (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \quad (3.45)$$

Los resultados obtenidos de la predicción de la radiación solar. Para evaluar la bonanza del modelo de estimación se ha utilizado la medida estadística del Coeficiente de Variación de la Raíz Cuadrada del Error

Cuadrático Medio (CVRMSE), que es una medida que no depende de la escala y cuya fórmula es:

$$CV - RMSE = \frac{RMSE}{\bar{Y}} \quad (3.46)$$

Dónde RMSE es la raíz cuadrada del error cuadrático medio:

$$RMSE = \sqrt{MSE} \quad (3.47)$$

Y el error cuadrático medio (MSE) se calcula de la siguiente manera:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \quad (3.48)$$

Siendo \hat{Y} el vector con las n predicciones y Y el vector con los valores reales. A modo de complementar y para conocer la bondad de los resultados en su escala, también se ha utilizado el (MAE) Error Absoluto Medio.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i| \quad (3.49)$$

Test	Día	Estación	Perfil de radiación
A	20/01/2014	Invierno	Soleado
B	10/01/2014	Invierno	Nublado
C	12/06/2014	Verano	Soleado
D	10/06/2014	Verano	Nublado

Figura 35 Datos de prueba para el modelo de predicción a corto plazo

Fuente: (E. Lorenz et. Al, 2010)

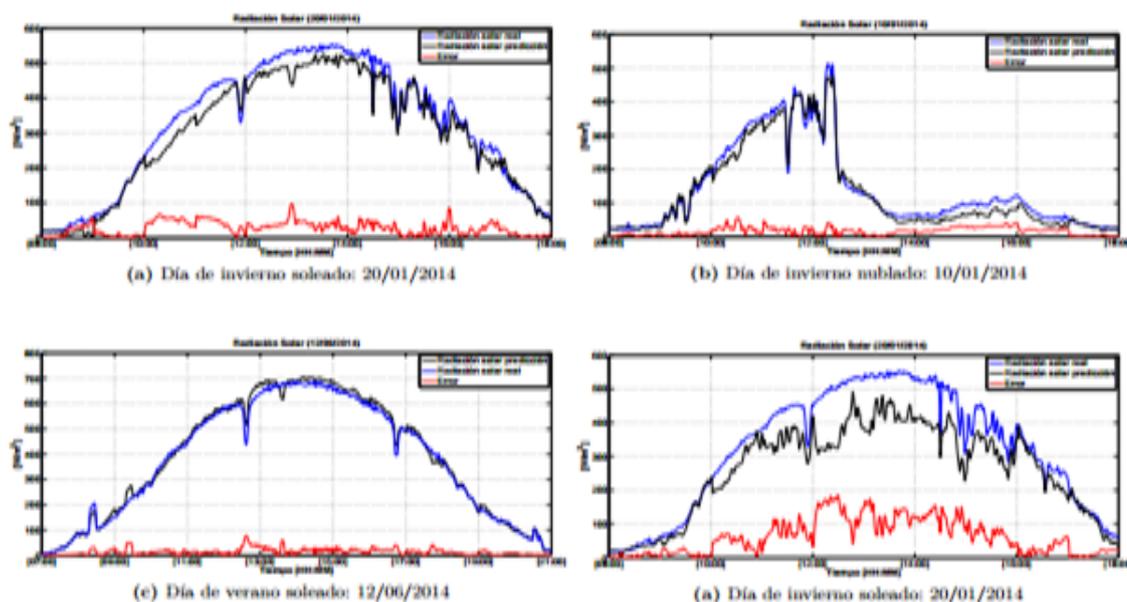


Figura 36 Gráficos de los datos de prueba en invierno y verano

Fuente: (E. Lorenz et. Al, 2010)

A partir de una base de datos recolectada de una planta real de fotobiorreactores, se plantea una combinación de estrategias basadas en el análisis de datos y el aprendizaje automático para el desarrollo de modelos predictivos a corto y medio plazo de la radiación solar, en concreto los modelos se basan en las máquinas de vectores de soporte (SVM) para la tarea de aprendizaje automático y predicción, así como en la teoría de inmersión para el análisis de series temporales de datos no lineales y la reconstrucción del espacio de estado utilizando optimización global mediante algoritmos genéticos (GA) para el ajuste de parámetros.

Así mismo, se ha planteado la utilización de algoritmos de clustering para la generación de perfiles diarios de radiación solar como base del planteamiento de predicción a medio plazo. Al tratarse de modelos basados en datos y aprendizaje automático, también se ha empleado otra variable que juega un papel directo en la radiación solar terrestre: la nubosidad, para ello se ha planteado una estrategia para caracterizar el conjunto de datos histórico así como crear el soporte para poder emplear predicciones externas de nubosidad de fuentes externas cuando se utilice el modelo en tiempo real.

Los resultados obtenidos en el modelo de predicción a corto plazo son muy favorables. En concreto, para la predicción a una hora los resultados oscilan entre el 6 % y el 24 % de CV-RMSE para el peor caso teniendo en cuenta la diversidad de condiciones y que son días no presentados en el entrenamiento del modelo. Además el modelo reconoce con bastante buen criterio la dinámica de la radiación solar, lo cual es muy útil en tareas de control. Para el caso de la predicción a corto plazo de 12 horas se han obtenido resultados con un ligero incremento en el error, pero similar en el sentido de que captan bien la dinámica de la evolución de la radiación solar. El error (CV-RMSE) para las pruebas de predicción a 12 horas oscila entre el 11 % y el 35 %.

3.10 Conclusión

En los SVM Las funciones kernel adecuada a cualquier algoritmo cualquier algoritmo que pueda expresarse en función de productos escalares sobre su espacio de entrada puede ser kernelizado en el algoritmo. El truco kernel permite que algoritmos “lineales” se apliquen sobre problemas “no lineales” La aplicación del modelo SVM es aplicable al uso de programación matemática formulada de forma que la interpretación estadística del modelo resulte particularmente apropiada y sustentada por teorías estadísticas de aprendizaje propuestas

Los métodos en las SVM están relacionados con problemas de *clasificación y regresión*. Los datos no lineales plantean la utilización de una combinación de estrategias basadas en el análisis de datos y el aprendizaje automático, la ventaja de la dinámica en el aprendizaje permite desarrollar modelos predictivos a corto y medio plazo en variables meteorológicas, los resultados con un ligero incremento en el error (CV-RMSE) para las pruebas de predicción a 12 horas oscila entre el 11 % y el 35 %.

3.11 Comparación ANN y SVM

Tabla 5

Comparación Técnica ANN y Técnica SVM

ANN	SVM
Capas ocultas transforman a espacios de cualquier dimensión	Kernel transforman a espacios de dimensión muy superior
El espacio de búsqueda tiene múltiples mínimos locales	El espacio de búsqueda tiene solo un mínimo global
El entrenamiento es muy costoso	El entrenamiento es muy eficiente
La clasificación es muy eficiente	La clasificación es muy eficiente
Se diseña el número de capas ocultas y nodos	Se diseña la función kernel y el parámetro de coste C
Buen funcionamiento en problemas típicos	Extremadamente robusto para generalización, menos necesidad heurística para entrenamiento
Su especificación se basa en el principio de minimización del riesgo estructural	de minimización del riesgo estructural que equivale a minimizar el límite superior del error de generalización del modelo

3.12 Comparación de trabajos de predicción de radiación solar

Tabla 6

Comparación de trabajos de predicción de radiación solar

AUTOR/ES	AÑO	PAÍS	TIPO DE MODELO	HORIZONTE DE PREDICCIÓN	VARIABLES DE ENTRADA	VARIABLES DE SALIDA	RESULTADOS
Sfetsos y Coonick [18]	1999	Grecia	Inteligencia Artificial ANN	Por hora	Radiación solar DIRECCIÓN DEL VIENTO	Radiación solar	CV- RMSE=31%
Kemmoku y Nakagawa [19]	1999	Japón	Inteligencia Artificial ANN	1 día	Datos de presión atmosférica	Radiación solar	MAPE=20%
Cao y Lin [20]	2008	Shangai China	Inteligencia artificial ANN y WAVELETS	1 Hora	Radiación solar pasada, modelo Ashare	Radiación Solar	RMSE=4.76%
Crispim <i>et ál.</i> [21]	2011	Portugal	Inteligencia artificial ANN	Hasta 30 minutos	Radiación solar pasada y nubosidad	Radiación solar	CV=RMSE=42,35%
]	2014	Turquia	Inteligencia artificial SVM	1 día	Temperatura, radiación solar	Radiación solar acumulada	CV-RSME=9,46%
Ramedani <i>et ál.</i> [24]	2014	Irán	Inteligencia artificial SVM	1 día	Horas del sol, Día del año	Radiación solar acumulada	RMSE=3.3 R ² =89,9%
				1 Hora	Radiación solar pasada y nubosidad	Radiación solar	CV=RMSE=25%
Mukherjee, Osuna y Giorsi [25]	1997	Shangai China	SVM Polinomios Redes RBF MLP	Hasta 30 minutos	Radiación solar pasada e índice de nubosidad	Radiación solar	CV-RSME=9,42%
Mulle, Smola <i>et al</i> [26]	2000	Pennsylvania	SVM RBF	1 día	Temperatura, radiación solar	Radiación solar acumulada	CV-RSME=9,42%
Pai y Hong[27]	2007	Thailandia	SVM Recurrente ANN Regression	1 día	Horas del sol, Día del año	Radiación solar acumulada	RP =76% PCM=24,46% EPOCHS=56% LM=51%

La aplicación de diversas técnicas IA en diversos escenarios como se indica en la tabla 7 de la comparación de trabajos de predicción de radiación solar, La predicción de radiación solar define a los SVM como una técnica robusta, su especificación se basa en el principio de minimización del riesgo estructural lo que equivale a minimizar el límite superior del error de generalización de un modelo determinado. Tomando en cuenta que el modelo software de predicción incide directamente en la salud de los seres humanos, se debe considerar un indicador importante como es el mínimo error en el modelo de predicción, ya que cualquier información errónea podría causar daños en la piel de las personas más vulnerables como primer objetivo

El modelo software a través de la selección de la técnica IA controlan el desempeño del modelo de predicción de series temporales, el pronóstico presenta diferentes características estructurales usando el modelo software para la posterior comparación de los resultados obtenidos.

CAPÍTULO IV

DISEÑO DEL MODELO SOFTWARE PARA PREDICCIÓN DEL INDICE DE RADIACIÓN SOLAR.

4. Introducción

En el presente apartado se presenta la metodología empleada para el diseño e implementación del modelo software para la predicción de la radiación solar en Riobamba-Ecuador. Cabe destacar que la metodología empleada se basa en el paradigma orientado a objetos, haciendo uso del Proceso Unificado de Desarrollo Software (PUD) y el lenguaje UML (Lenguaje Unificado de Modelado).

4.1 Lenguaje unificado de modelado (UML)

El Lenguaje Unificado de Modelado (UML), es un lenguaje estándar para escribir planos de software que puede utilizarse para visualizar, especificar, construir y documentar los artefactos de un sistema que involucra una gran cantidad de software

UML es un lenguaje de modelado que permite la construcción de distintos modelos: diagramas de clases, diagramas de casos de usos, es auto descriptivo ya que puede especificarse por medio de diagramas de clase de UML. Por todo ello, UML es independiente del proceso de desarrollo seguido, aunque para su uso de manera óptima se ha de utilizar en un proceso que sea dirigido por casos de uso, centrado en la arquitectura y que sea iterativo e incremental.

El modelado conceptual de UML se basa en elementos, las relaciones y los diagramas.

- Elementos: son abstracciones en un modelo. Los bloques básicos de construcción orientados a objetos de UML.
- Relaciones: enlaces entre elementos de un modelo. Las más importantes son:

- Dependencias: surge cuando un elemento depende de otro, de forma que un cambio en el elemento independiente afecta al elemento dependiente.
- Generalizaciones: relación de herencia entre el elemento más general (superclase) y otro más específico de ese mismo elemento general (subclase).
- Asociaciones: para tipos de relación estructural que especifica que tipos de objetos están conectados con los objetos de otro.
- Diagramas: representación gráfica de un conjunto de elementos visualizado la mayoría de las veces como un grafo conexo de nodos (elementos) y arcos (relaciones). Los diagramas agrupan colecciones de elementos ligados aportando un significado adicional, y ofrecen distintas perspectivas de una abstracción de la realidad. Un mismo elemento puede aparecer en distintos diagramas.

UML define un número de reglas que especifican a qué debe parecerse un modelo bien formado. Además, también se puede especificar la arquitectura de un sistema, mediante un conjunto de decisiones significativas sobre su organización, la selección de elementos estructurales y sus interfaces, su comportamiento, la composición de esos elementos estructurales y de comportamiento en subsistemas y el estilo arquitectónico.

4.2 Proceso unificado de desarrollo (Pud)

El Proceso Unificado de Desarrollo PUD contiene un conjunto de actividades necesarias para transformar los requisitos de usuario en un sistema software. Su principal objetivo es garantizar la producción de software con calidad dentro de plazos y presupuestos predecibles. El PUD usa UML para preparar los esquemas de los sistemas software a lo largo de sus iteraciones y ciclos.

4.2.1 Ciclo de vida del proceso unificado

En el Proceso de desarrollo Unificado Iterativo, cada uno de los ciclos de iteración forma el sistema software. Se repite a lo largo de una serie de ciclos que constituyentes la vida de un sistema.

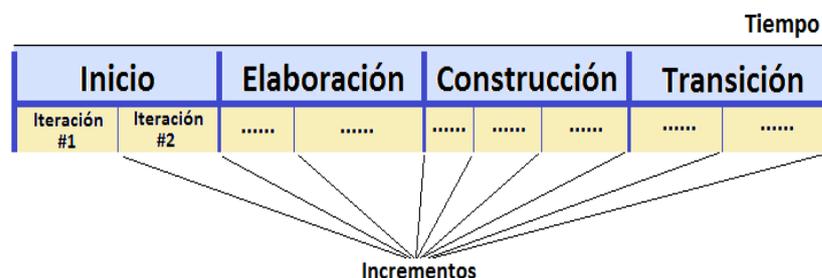


Figura 37 Ciclo de vida del Proceso Unificado de Desarrollo

Con la finalización de cada ciclo iterativo se concluye una versión del producto para los clientes. Cada ciclo consta de cuatro fases: inicio, elaboración, construcción y transición, además cada fase termina con un hito (artefactos), donde se deben tomar decisiones como se observa en la figura 38

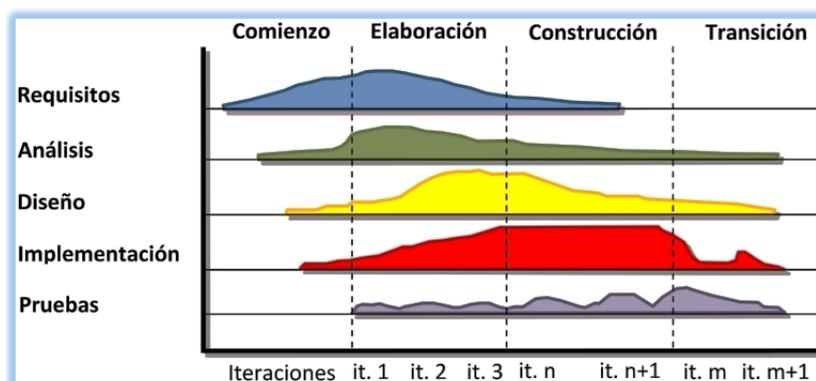


Figura 38 Fases del PUD

Cada ciclo posee una nueva versión del sistema, y cada versión es un producto preparado para su entrega. El producto software terminado en cada ciclo, incluye los requisitos, casos de uso, especificaciones no funcionales y casos de prueba de esa porción.

4.2.2 Flujos de trabajo

Los flujos de trabajo representan la realización de un caso de uso de negocio o parte de él. Se pueden escribir con diagramas de actividad en los

que se muestran qué trabajadores (roles) ejecutan qué actividades y qué artefactos producen.

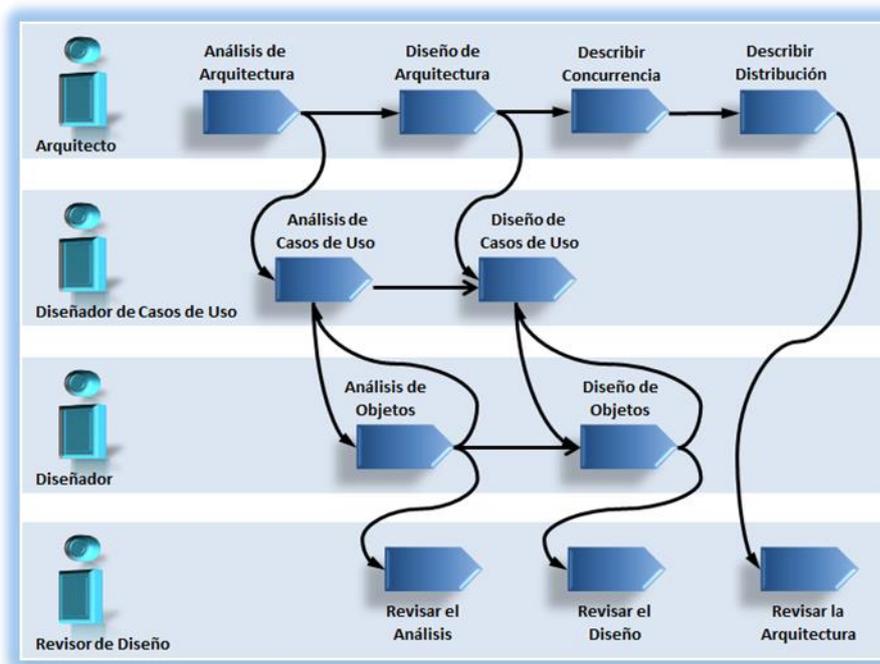


Figura 39 Flujos de Trabajo

Los flujos de trabajo fundamentales se describen a continuación:

- **Análisis:** los requisitos descritos en las mediciones son analizados, refinados y estructurados.
- **Diseño:** el sistema es modelado en base a los requisitos y una de las entradas fundamentales para este proceso es la salida producida de la fase de análisis.
- **Implementación:** a partir del diseño se implementa el sistema en términos de componentes.
- **Pruebas:** el resultado de la implementación es verificado probando cada construcción.

4.2.3 Fases

Inicio: Se desarrolla el análisis de las mediciones en el software hasta justificar la puesta en marcha de la primera versión del modelo software, primera versión del modelo de casos de uso, del modelo de análisis y diseño, primer esquema de la descripción de una arquitectura candidata, prototipo que muestra el uso del nuevo sistema.

Elaboración: se recopila la mayor parte de los requisitos pendientes, se establece una arquitectura, se realiza un esquema de de los modelos que representa el primer modelo de casos de uso, el modelo de análisis y el de diseño, y un primer esquema de la descripción de una arquitectura que perfila las vistas de los modelos de casos de uso, análisis, diseño e implementación.

Construcción: se prepara el software en su versión operativa inicial, con la calidad adecuada para su aplicación y que cumpla los requisitos. Se parte de la línea base de la arquitectura ejecutable, se detallan los casos de uso y escenarios resultantes, se cierran los métodos de análisis/diseño/implementación, se integran y se prueban los subsistemas y el sistema completo.

Transición: se deben cumplir los objetivos de las fases anteriores hasta la satisfacción del usuario final Esta fase concluye con el lanzamiento del producto. Los resultados obtenidos en esta fase son: el software ejecutable, documentos legales, versión completa y corregida de la línea base del producto, descripción completa y actualizada de la arquitectura.

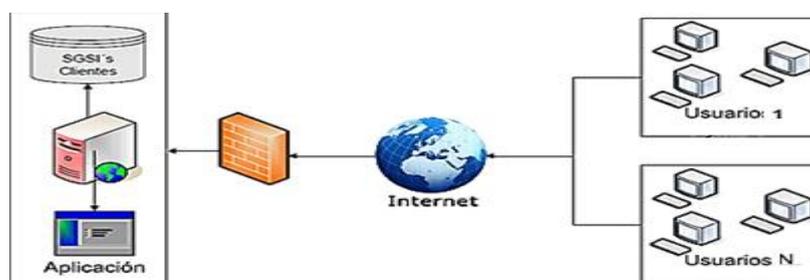


Figura 40 Arquitectura de Capas utilizada para el sistema

4.2.4. La plataforma de desarrollo .Net

La plataforma .NET es un entorno de desarrollo de software creada por Microsoft que permite un rápido proceso de aplicaciones, Ofrece soluciones a los problemas de programación actuales, como son la administración de código o la programación para Internet Mediante esta herramienta se ofrece un entorno de ejecución que permite crear aplicaciones robustas y escalables.

Entre sus características tenemos:

- Permite la utilización de diferentes lenguajes de programación, existiendo más de 30 adaptados a .Net: Visual Basic, C#, C++, J#, Python.
- Permite su integración con dispositivos distintos a un PC como teléfonos móviles o consolas.
- Facilita la creación, implementación y administración de aplicaciones seguras, sólidas y de gran rendimiento.
- Gestiona la mayor parte de la estructura necesaria para la generación de software permitiendo a los programadores centrarse en la lógica de negocio.



Figura 41 Arquitectura del .Net Framework

4.2.5 Desarrollo e implementación con Visual Studio

SE utiliza Visual Studio versión 2010 como entorno de desarrollo por ser una herramienta de fácil uso y soporta C#.

4.2.6 El lenguaje de programación C#

Es un lenguaje de propósito general orientado a objetos, que incluye soporte para programación estructurada y basada en componentes. Desarrollado y estandarizado por Microsoft como parte de su plataforma .NET, y posteriormente aprobado como estándar por la ECMA e ISO. Su sintaxis básica procede de C y C++, y utiliza el modelo de objetos de la plataforma .Net, similar al de Java pero con mejoras derivadas de otros lenguajes.

4.2.7 Sql server

El motor de base de datos utilizado para la herramienta web ha sido SQL Server el cual, es un sistema para la gestión de base de datos basado en el modelo relacional y generado por Microsoft. Entre algunas de sus características destacan

4.3 Iteraciones y evolución del proyecto

En base a la metodología PUD se muestra las 4 iteraciones en las que se ha dividido la creación de la herramienta infosolar que mide el índice de radiación Solar.

4.3.1. Iteración 1: planificación del proyecto

En esta primera iteración se realiza un estudio sobre la necesidad del proyecto teniendo en cuenta todos sus aspectos: viabilidad, alcance, recursos,... Tras decidir su construcción, se efectúa la captación de requisitos que, en este caso, proceden tanto de personas que trabajan en SOLCA, como de los estudiantes de la Facultad de Informática y Electrónica.

A continuación, se muestra un baremo (Tabla 4.1) de la primera iteración:

Tabla 7

Resumen de la Iteración 1 (Planificación del Proyecto)

Fase del desarrollo PUD implicada	Inicio
Intensivo en los flujos de trabajo	Requisitos
Objetivos	<ul style="list-style-type: none"> ✓ Estudiar la viabilidad del proyecto y sus aplicaciones. ✓ Automatizar mayor número de tareas.
Productos	<ul style="list-style-type: none"> ✓ Ámbito y límites del proyecto. ✓ Lista de requisitos. ✓ Análisis de actores implicados. ✓ Lista de funcionalidades a desarrollar. ✓ Esquemas iniciales para el análisis de casos de uso. ✓ Estimación de costes en recursos y tiempo. ✓ Estimación de riesgos.

4.3.2. Iteración 2: diseño y creación de la base de datos

Se lleva a cabo las tareas de diseño y creación de la base de datos. El Sistema Gestor de Base de Datos (SGBD) que se utiliza es MS SQL Server,

Se diseñan diagramas Entidad-Relación para el modelado de la Base de Datos, que describirán todas las relaciones existentes entre las entidades.

Tabla 8

Resumen de la Iteración 2 (Diseño y creación de la Base de Datos)

Fase del desarrollo PUD implicada	Elaboración
Intensivo en los flujos de trabajo	Análisis, diseño e implementación
Objetivos	<ul style="list-style-type: none"> ✓ Captar requisitos no recogidos con anterioridad. ✓ Obtener una Base de Datos que dé soporte de persistencia al sistema creado.
Productos	<ul style="list-style-type: none"> ✓ Lista actualizada de requisitos. ✓ Base de Datos operativa.

4.3.3. Iteración 3: construcción de la herramienta

En esta tercera iteración se construye la herramienta Web INFOSOLAR. Uno de los requisitos fundamentales del sistema es el acceso web, para el desarrollo de la herramienta web con los objetivos. En esta iteración, se produce un alto grado de comunicación con los usuarios.

Tabla 9

Resumen de la Iteración 3 (Construcción de la Herramienta)

Fase del desarrollo PUD implicada	Construcción
Intensivo en los flujos de trabajo	Análisis, diseño, implementación y pruebas
Objetivos	<ul style="list-style-type: none"> ✓ Dar respuesta a los requisitos obtenidos de los usuarios. ✓ Obtener un sistema que permita difundir el índice de radiación solar
Productos	<ul style="list-style-type: none"> ✓ Herramienta web para administrar los valores obtenidos desde el SAD. ✓ Resultados obtenidos a partir del uso de la herramienta.

4.3.4. Iteración 4: entrega

En esta última iteración se llevan a cabo todas las operaciones necesarias para la puesta en marcha de la herramienta: últimas pruebas, pequeñas mejoras detectadas, y la subida de la herramienta al servidor para que pueda accederse a ella a través de Internet.

Tabla 10

Resumen de la Iteración 4 (Entrega)

Fase del desarrollo PUD implicada	Transición
Intensivo en los flujos de trabajo	Pruebas
Objetivos	<ul style="list-style-type: none"> ✓ Versión final de la herramienta para su uso. ✓ Revisión y recolección de artefactos. ✓ Desarrollo de entregables finales.
Productos	<ul style="list-style-type: none"> ✓ Memoria del Proyecto. ✓ Presentación del Proyecto. ✓ Herramienta Web accesible desde Internet.

4.4. Resultados – Construcción de la herramienta web infosolar

Luego de haber indicado la metodología con las iteraciones es necesario indicar la especificación de requisitos, el análisis, diseño implementación y pruebas.

4.4.1. Especificación de Requisitos

A lo largo de esta fase se define el alcance, se analizan requisitos que debe cumplir el sistema, se produce la identificación de actores y casos de uso,

4.4.2. Ámbito y límites del proyecto

El sistema a desarrollar, es una herramienta web que permita observar las mediciones de la radiación dólara en tiempo real, estas mediciones son tomadas de un sistema de adquisición de datos en horario de 6am a 6pm cada 5 minutos

El sistema permitirá tener un monitoreo de las radiaciones solares en tiempo real, necesarias para llevar un control y afectación a las personas con cáncer de piel en la zona de Riobamba-Chimborazo

4.4.3. Análisis de los actores implicados

Actor Usuario que representa el rol más básico, que corresponderá al usuario final es decir quien consulta las mediciones realizadas a través del Software y que habitualmente para tareas de control o seguimiento, monitorización y mantenimiento de los niveles de cumplimiento.

Actor Administrador, que representa el rol superior, pensado para la persona que gestionará los datos recibidos desde el sistema de adquisición de datos

4.4.4. Análisis de objetivos o requisitos que debe cumplir el software

Los objetivos o requisitos pueden ser clasificados en los que se describe la funcionalidad del sistema y los que se refieren a aspectos del desarrollo en sí mismo.

Se realizan entrevistas a SOLCA se realiza una evaluación de la viabilidad de requisitos y son:

- La herramienta deberá permitir dar de baja a un usuario registrado en una Organización.

- La herramienta debe permitir el cambio de contraseña (y si existiera, de palabras secretas).
- La herramienta deberá indicar a los usuarios de alguna manera visual (p.ej. con un código de colores) si una medición se encuentra actualizada y revisada.
- La herramienta, sería interesante que además de los datos actuales, facilitara algún tipo de historial para los valores globales de las mediciones.
- La herramienta podrá permitir consultar las propiedades de las mediciones, indicadores.
- La herramienta deberá contener una parte o sección en la que los usuarios puedan observar un informe-resumen de los valores medidos versus los valores de la OMS y EXA
- La herramienta deberá facilitar información sobre la utilidad o significado de las tabas y mediciones
- La herramienta debe hacer un avance cronológico de las mediciones en tiempo real.

En cuanto a los requisitos no funcionales, los identificados finalmente han sido:

- La interfaz de usuario deberá resultar amigable y sencilla en su utilización de manera directa, clara y sin problemas.
- Las interfaces de monitorización de las mediciones deberá aportar una idea clara e inmediata sobre el estado del sol en el día.
- La herramienta deberá manejar algún tipo de seguridad en cuanto a las contraseñas

4.5 Costos en recursos y tiempos

Para todo proceso de desarrollo software debe realizarse una estimación en cuanto a recursos, tiempo y coste económico con la finalidad de que el proyecto pueda ser terminado sin problemas dentro de los recursos, presupuesto y fechas planificados.

En esta ocasión, no existe ni un presupuesto inicial ni una planificación de fechas aunque es importante que, como mucho, las tareas implicadas en la construcción del Software Infosolar se terminen a mediados de diciembre de 2015. Además, cabe señalar que en el proceso de desarrollo de la herramienta de este PFC, al participar una única persona, el cálculo de costes se verá muy simplificado.

Costos asociados:

- ✓ Una persona que se encarga de la administración de la página: 400 dólares al mes
- ✓ Una computadora portátil \$1200
- ✓ Entorno de desarrollo y sistema de adquisición de datos \$1000
- ✓ SQL Server: GPL.
- ✓ SQL Server Management Studio: GPL
- ✓ Microsoft Office: \$150
- ✓ Visual Paradigm for UML: GPL
- ✓ Otro material: \$50

Total: \$2400

4.6 Estimar riesgos e incertidumbres

Un riesgo en un proyecto “es un evento incierto o condición incierta que si ocurre, tiene un efecto positivo o negativo sobre el proyecto”, gestionándose normalmente los que representan una amenaza para el éxito del proyecto, debiendo ser identificados, analizados y “solucionados”

En el caso del nivel de riesgo, se define el nivel límite aceptable el como “Poco” debiendo ser tratados todos los situados en niveles superiores, mientras que los que se encuentren en los inferiores, podrán ser asumidos

sin ningún tipo de problema y no supondrán peligro alguno para el buen desarrollo del proyecto.

4.7 Análisis

En este apartado se ofrece el análisis de las distintas partes en las que puede descomponerse la herramienta web INFOSOLAR mediante los diagramas de caso de uso y las descripciones textuales de algunos de estos casos de uso.

A continuación, se muestra el diagrama de casos de uso general del sistema en el que se indica cada una de las funcionalidades a nivel global de la herramienta y los actores que intervienen (Fig. 42).

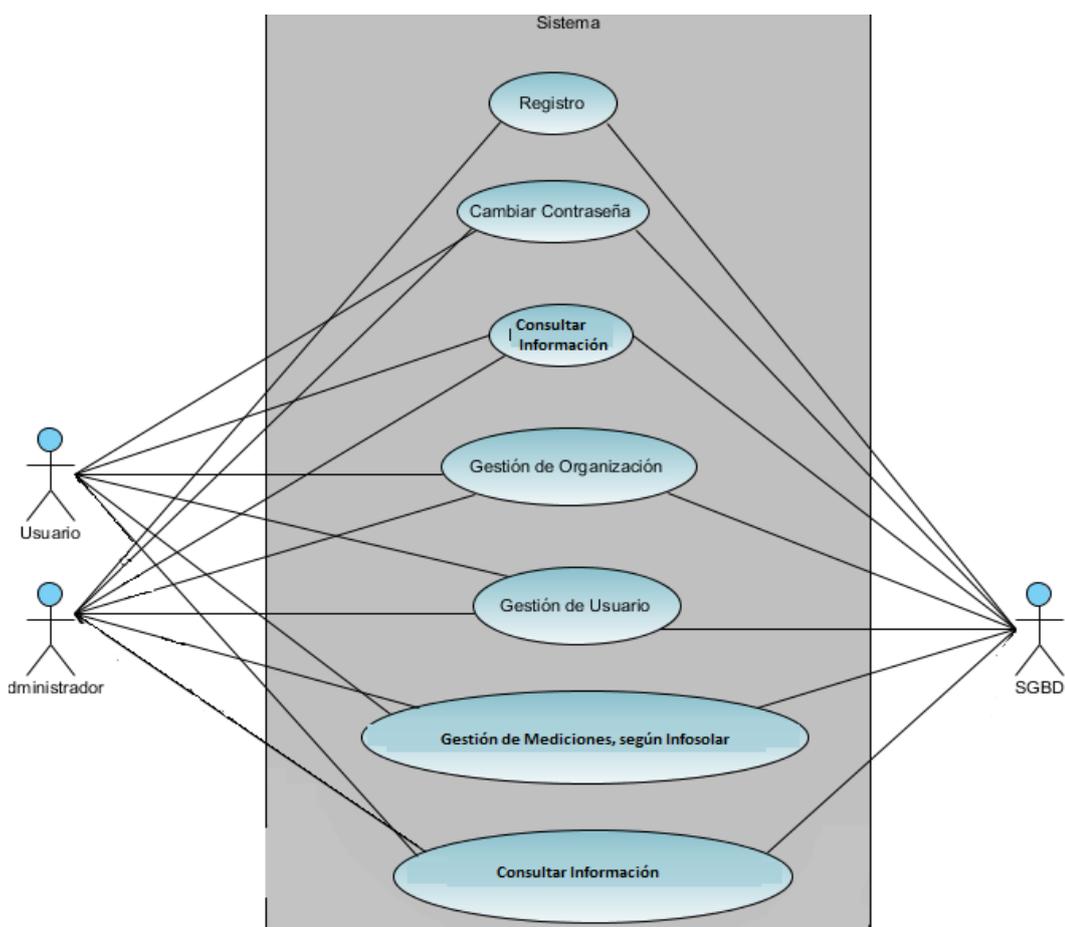


Figura 42 Diagrama general de casos de uso del sistema

A continuación, se muestra la *descripción textual* de **Gestión de Usuario**:

Tabla 11

Descripción textual de Gestión de Usuario

Nombre: Gestión de Usuario
Actores involucrados: Administrador o Usuario, y SGBD.
Flujo Normal: 1.- El usuario “Administrador” o “Usuario” pulsa, en la barra de botones general, el correspondiente a la <i>Gestión de Usuario</i> . 2.- Redirección a la pantalla correspondiente a las tareas de <i>Datos Personales</i> 3.- Se consulta en la base de datos los datos del “Usuario”. 4.- La consulta devuelve los datos del “Usuario”. 5.- Los datos del usuario se muestran en la pantalla.
Descripción: Caso de uso correspondiente al acceso a las funcionalidades de <i>Gestión de Usuario</i> ,

4.7.1 Arquitectura

En el presente apartado se tratará la elaboración en cuanto a análisis y diseño del sistema en cuánto al análisis se ofrece la elaboración de los casos de uso en relación a los requisitos, tomando en cuenta que este análisis es fuerte en la primera iteración del proyecto, luego de aquello el diseño basado en un diagrama de secuencia y diagrama de clase, esta fase es intensiva en las iteración dos y tres.

4.7.2 Registro

Módulo mediante el cual un usuario podrá realizar el registro en el sistema de una nueva Organización

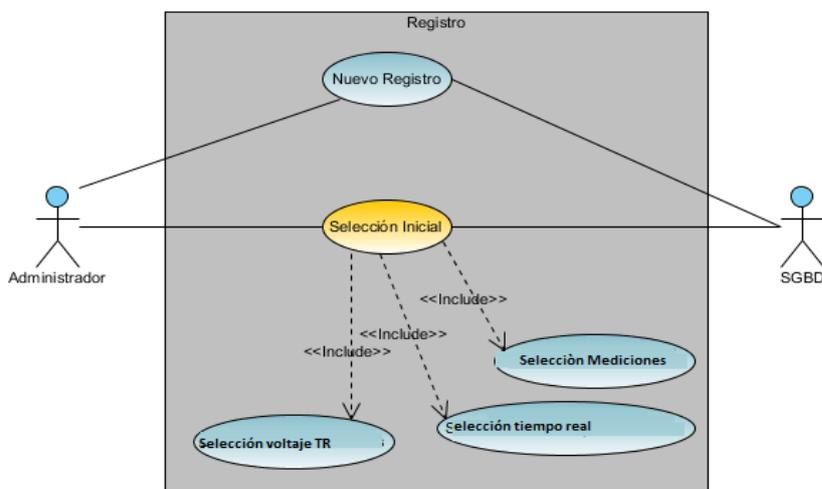


Figura 43 Diagrama de casos de uso de Registro

Tabla 12
Descripción textual de Nuevo Registro

Nombre: Nuevo Registro
Actores involucrados: Administrador y SGBD.
Precondiciones: Un usuario ha debido ejecutar el caso de uso <i>Registro</i> (consistente básicamente en pulsar el botón “Registrarse” de la pantalla de inicio de sesión).
Postcondiciones: El nuevo “Administrador” tendrá acceso al caso de uso Selección Inicial.
Flujo Normal: 1.- El usuario pulsa el botón “Registrarse”. 2.- Se muestra la pantalla para introducir los datos de la Organización y los datos personales del nuevo “Administrador” para poder registrar la nueva Organización y el nuevo “Administrador” asociado a ella. 3.- El usuario introduce los datos de la Organización y los suyos propios.
Flujo Alternativo 1: 1.- 6.- Mismos pasos que para el Flujo Normal. 7.- El resultado de la consulta que ya existe 8.- Se consulta en la base de datos sobre la existencia de alguien registrado en el sistema con el usuario introducido. 9.- El resultado de la consulta es que no existe nadie ya registrado con ese usuario. 10.- Se muestra un mensaje advirtiendo del problema.
Flujo Alternativo 2: 1.- 8.- Mismos pasos que para el Flujo Normal. 9.- El resultado de la consulta es que ya existe una persona registrada con ese usuario. 10.- Se muestra un mensaje advirtiendo del problema.
Flujo Alternativo 3: 1.- 2. Mismos pasos que para el Flujo Normal. 3.- El usuario introduce los datos de la Organización y los suyos propios, dejando vacío por error alguno de los campos. 4.- Una vez introducidos los datos requeridos. 5.- Faltan datos por introducir necesarios para los registros de la Organización o del “Administrador”. 6.- Se muestra un mensaje advirtiendo de que faltan datos por introducir.
Flujo Alternativo 4: 1.- 2. Mismos pasos que para el Flujo Normal. 3.- El usuario introduce los datos de la Organización y los suyos propios, utilizando por error un carácter distinto a un número en los campos para el teléfono o para el código postal. 4.- Existen datos que no son válidos.
Flujo Alternativo 5: 1.- 2. Mismos pasos que para el Flujo Normal. 3.- El usuario introduce los datos de la Organización y los suyos propios, no poniendo por error en el campo de confirmación de la contraseña lo mismo que puso para la contraseña. 4.- La contraseña y su confirmación no coinciden. 5.- Se muestra un mensaje advirtiendo del problema.
Flujo Alternativo 6: 1.- 2. Mismos pasos que para el Flujo Normal. 3.- El usuario introduce los datos de la Organización y los suyos propios, no poniendo por error en el campo de confirmación de la palabra clave lo mismo que puso para la palabra clave.

4.7.3 Cambiar contraseña

Módulo mediante el cual, un “Administrador” o un “Usuario” ya registrado en el sistema, podrá, sin iniciar sesión, cambiar su contraseña utilizando para ello su palabra clave secreta, pudiendo así seguir accediendo al sistema tras haber olvidado o perdido su contraseña de usuario de acceso al sistema.

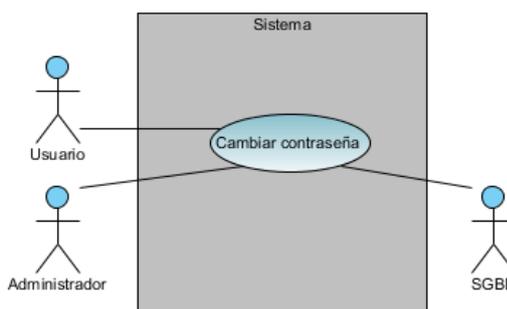


Figura 44 Diagrama de casos de uso de Cambiar contraseña

A continuación se ofrece la *descripción textual* correspondiente:

Tabla 13

Descripción textual de Cambiar contraseña

Nombre: Cambiar contraseña

Actores involucrados: Administrador o Usuario, y SGBD.

Flujo Normal:

- 1.- El “Administrador” o “Usuario” selecciona el enlace *Olvidé mi contraseña* en la pantalla de inicio de sesión.
- 2.- Se muestra la pantalla para introducir el usuario, palabra clave, y seleccionar el rol del “Administrador” o “Usuario”. Además, debe introducir la nueva contraseña así como confirmación de la misma.
- 3.- El “Administrador” o “Usuario” introduce sus datos de usuario, la nueva contraseña y su confirmación.
- 4.- A continuación, el “Administrador” o “Usuario” pulsa el botón *Cambiar Contraseña*.
- 5.- Los datos han sido introducidos bien.
- 6.- Se consulta en la base de datos sobre la existencia de un usuario al que corresponden el usuario, palabra clave y rol introducidos.
- 7.- El resultado de la consulta es que sí existe un usuario con esos datos.
- 8.- Así, la verificación resulta correcta y se cambia en la base de datos la contraseña del usuario “Administrador” o “Usuario” almacenada por la solicitada en este momento.
- 9.- Se muestra un mensaje de texto informando de que la nueva contraseña ha sido almacenada.

4.7.4 identificación

Módulo mediante el cual un “Administrador” o “Usuario” podrá identificarse en el sistema mediante su usuario y su contraseña pudiendo acceder a sus determinadas funciones.

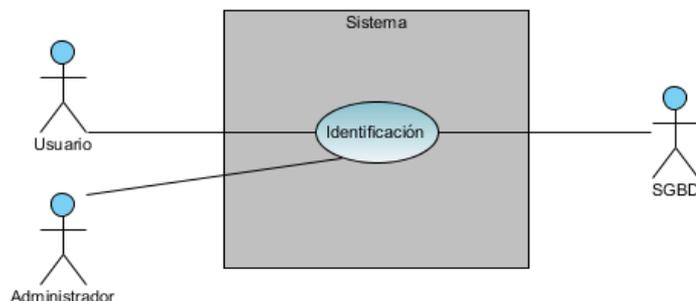


Figura 45 Diagrama de casos de uso de Identificación

La *descripción textual* para este caso de uso se corresponde con la siguiente:

Tabla 14

Descripción textual de Identificación

Nombre: Identificación
Actores involucrados: Administrador o Usuario, y SGBD.
Flujo Normal: 1.- Un usuario “Administrador” o “Usuario” introduce su usuario, contraseña y selecciona su rol en la pantalla de inicio de sesión, y pulsa el botón <i>Acceder</i> . 2.- Se consulta en la base de datos si existe un usuario con ese usuario, contraseña y rol. 3.- El resultado de la consulta es que sí existe un usuario con esos datos de identificación.
Flujo alternativo 1: 1.- 2.- Mismos pasos que para el Flujo Normal. 3.- El resultado de la consulta es que no existe un usuario al que correspondan esos datos de identificación. 4.- Así, la verificación resulta incorrecta ya que no existe en la base de datos ningún usuario que se corresponda con los datos introducidos.
Flujo alternativo 2: 1.- Un usuario “Administrador” o “Usuario” al introducir su usuario, contraseña y seleccionar su rol en la pantalla de inicio de sesión, deja en blanco por error el campo del usuario o de la contraseña, y pulsa el botón <i>Acceder</i> . 2.- Se muestra un mensaje de texto avisando al usuario de que los datos introducidos son incorrectos.
Descripción: Caso de uso correspondiente a la identificación de un usuario “Administrador” o “Usuario” en el sistema mediante su usuario, contraseña y rol.

4.7.5 Gestión de organización

Módulo mediante el cual un “Administrador” podrá consultar los valores correspondientes a la Organización, y además, podrá dejar constancia en el sistema de que la Organización ha sido sometida a una nueva revisión para la ISO/IEC 27001. Por su parte, el “Usuario” sólo podrá consultar los valores globales correspondientes a la Organización.

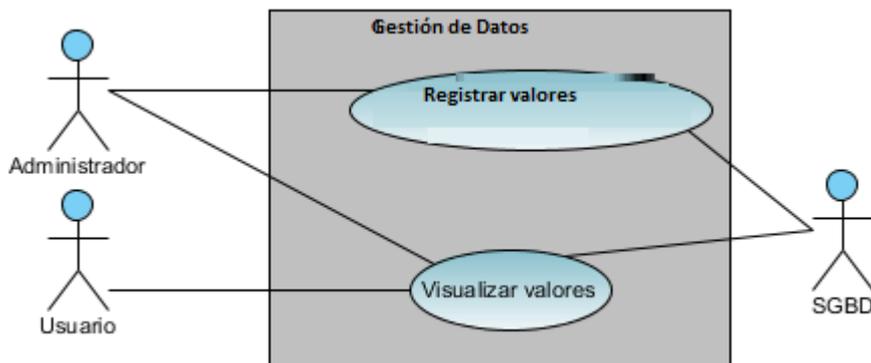


Figura 46 Diagrama de casos de uso de Gestión de Organización

4.7.6. Gestión de usuario

Módulo mediante el cual un “Administrador” podrá consultar sus propios datos, y además, podrá registrar a nuevos usuarios en la Organización, o eliminarlos (darlos de baja). El “Usuario”, podrá consultar sus datos personales.

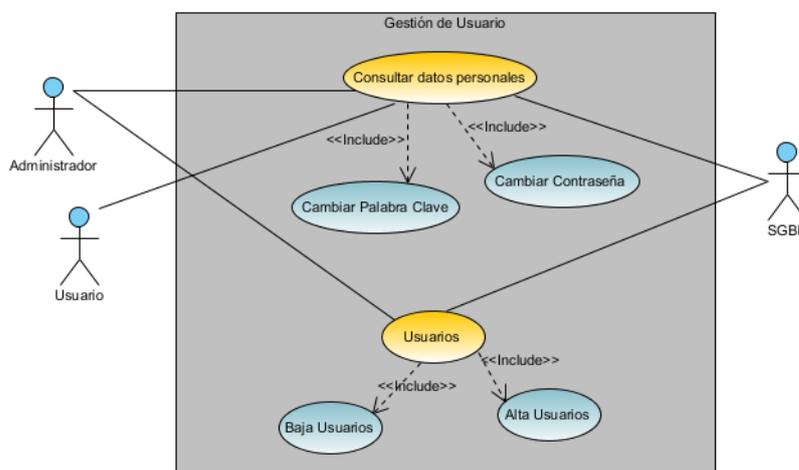


Figura 47 Diagrama de casos de uso de Gestión de Usuario

Con sólo ejecutar el caso de uso *Consultar datos personales*, al usuario ya se le ofrecerán sus datos personales y, la posibilidad de ejecutar los casos de uso *Cambiar Palabra Clave* y *Cambiar Contraseña*, cuya idea es similar al cambio de contraseña sin iniciar sesión. A continuación, se ofrece la *descripción textual* para **Usuarios**:

Tabla 15
Descripción textual de Usuarios

Nombre: Usuarios
Actores involucrados: Administrador y SGBD.
Precondiciones: El usuario "Administrador" ha debido ejecutar el caso de uso <i>Gestión de Usuario</i> .
Postcondiciones: El "Administrador" tendrá acceso a ejecutar los casos de uso <i>Alta Usuarios</i> y <i>Baja Usuario</i> .
Flujo Normal: 1.- El "Administrador" selecciona, en el menú superior de la pantalla de <i>Gestión de Usuario</i> , la opción <i>Usuarios</i> . 2.- Redirección a la pantalla correspondiente a las tareas de <i>Usuarios</i> . 3.- Se consulta en la base de datos los usuarios de la Organización. 4.- La consulta devuelve los usuarios de la Organización. 5.- Se muestra la pantalla correspondiente a las tareas de <i>Usuarios</i> con los usuarios de la Organización.
Descripción: Caso de uso correspondiente a la acción del "Administrador" de acceder a la opción <i>Usuarios</i> de la gestión de usuario.

La descripción textual para Alta Usuario seguirá una filosofía similar a la del registro de un "Administrador" inicial junto a la Organización pudiendo en esta ocasión, elegir si el nuevo usuario será "Administrador" o "Usuario" pudiendo ser ejecutada esta operación únicamente por un "Administrador" de la Organización.

Tabla 16
Descripción textual de Baja Usuarios

Nombre: Baja Usuarios
Actores involucrados: Administrador y SGBD.
Precondiciones: El usuario "Administrador" ha debido ejecutar el caso de uso <i>Usuarios</i> .
Postcondiciones:
Flujo Normal: 1.- El "Administrador" selecciona, en la zona correspondiente de la pantalla de gestión de los usuarios de la Organización, el usuario de la persona a eliminar. 2.- El "Administrador" pulsa el botón "Eliminar". 3.- En la base de datos se localiza al usuario implicado. 4.- Se almacena en la base de datos un registro de histórico sobre la eliminación

4.7.7 Gestión de mediciones

Módulo mediante el cual un “Administrador” podrá realizar todas las operaciones posibles (crear y añadir nuevos elementos, eliminar elementos, actualizar el peso de cada uno de los elementos dentro de su elemento superior, actualizar el valor de las medidas actualizando así las mediciones, indicar que los indicadores han sido revisados dando por revisadas las mediciones, y monitorizar el estado de los elementos) sobre los elementos Medición.

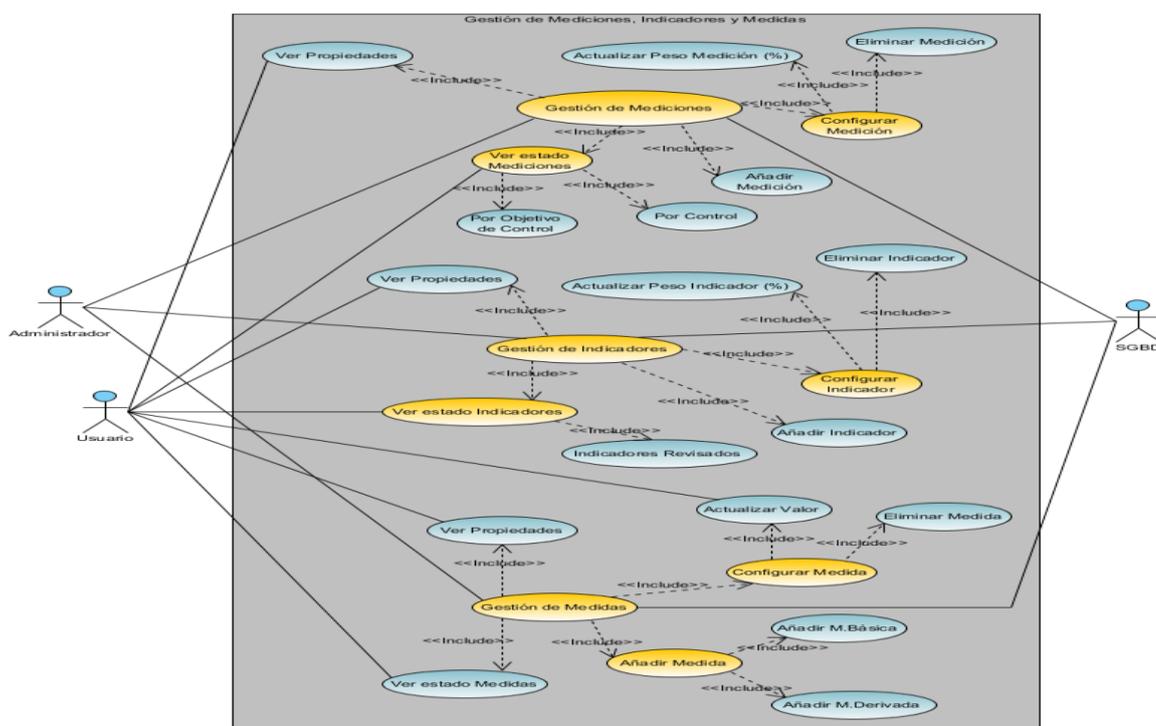


Figura 48 Diagrama de casos de uso de Gestión de Mediciones

4.7.8 Consultar informe

Módulo mediante el cual, tanto un “Administrador” como un “Usuario”, podrá consultar un informe-resumen con los valores de todos los elementos presentes en la Organización.

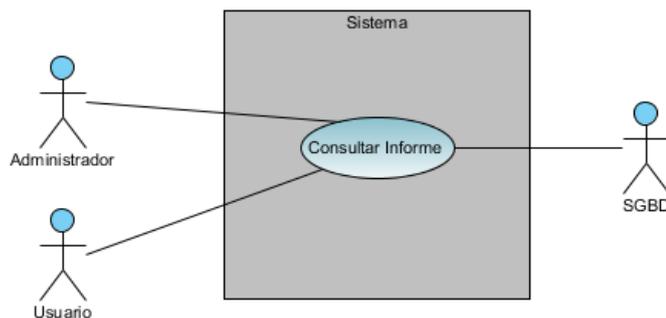


Figura 49 Diagrama de casos de uso de Consultar Informe

Al “Administrador” o “Usuario” que ejecute este caso de uso se le facilitarán los valores referidos a la Organización a nivel general, los de todos los Dominios, Objetivos de Control, Controles, Mediciones, Indicadores y Medidas, todos ellos presentados de una manera ordenada y organizada en tablas gracias a las cuales, el “Administrador” o “Usuario”, podrá ver toda la información referida a la Organización en su conjunto de una vez, y relacionarla el mismo.

4.7.9 Consultar información

Módulo mediante el cual, tanto un “Administrador” como un “Usuario”, podrá consultar un manual de usuario de la herramienta

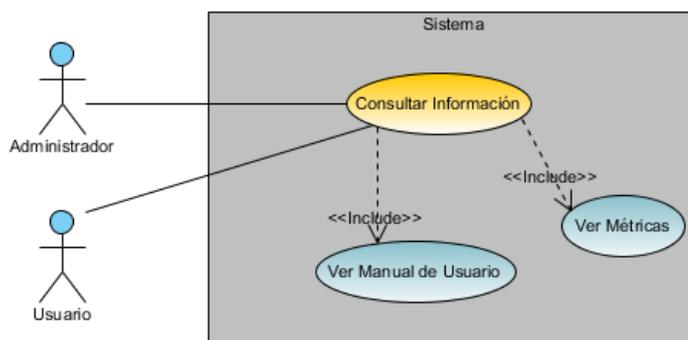


Figura 50 Diagrama de casos de uso de Consultar Información

4.8 Diseño

En esta fase se diseñará la *arquitectura del sistema* y, para ello, se utilizarán *diagramas UML* y de *entidad-relación* para el modelado de la base de datos. Así, la arquitectura del sistema, al no ser necesario realizar ninguna clase de ingeniería inversa ni reutilización de software antiguo, se podrá plantear como una *implementación Top-Down*, desde la especificación de requisitos hasta los productos software finales.

Al seguir este proceso de diseño, se logra obtener un software altamente reutilizable, mantenible, con alto grado de trazabilidad e independencia tanto de la plataforma en la que se ejecute como del escenario final de uso.

4.8.1 Capas de la arquitectura

En este apartado se muestran las diferentes capas diseñadas para la herramienta. En la *Figura 52*, se ofrece el modelo de tres capas de primer nivel utilizado.

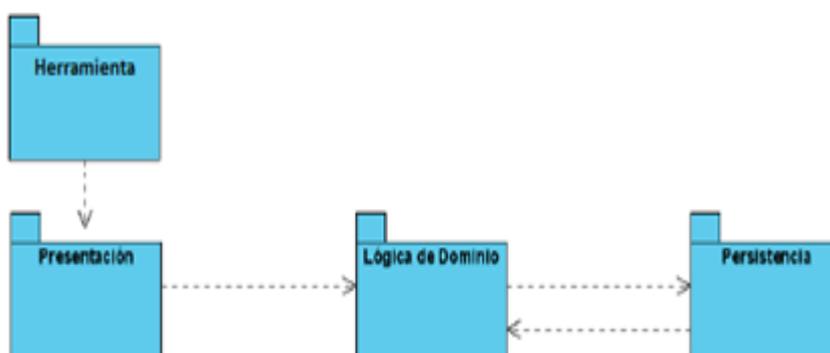


Figura 51 Diagrama general de arquitectura multicapa en la Herramienta

4.8.2 Diagrama de clases

A continuación, se muestra el diagrama de clases general de la lógica de dominio (Fig. 53).

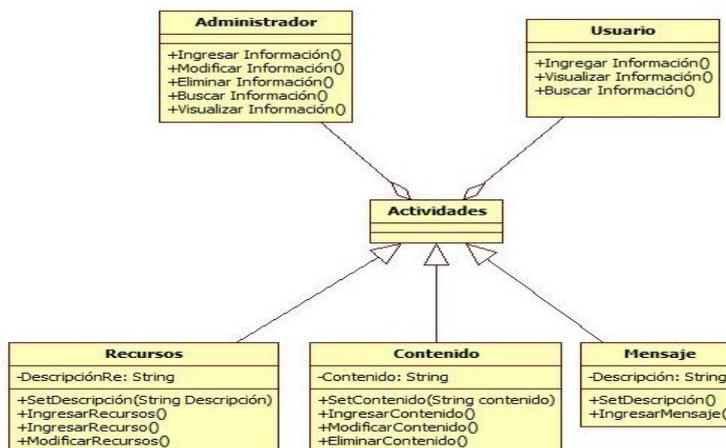


Figura 52 Diagrama de clases general de la lógica de dominio

4.8.3 Diagrama de paquetes

A continuación, se muestra el diagrama de paquetes (Fig. 54).

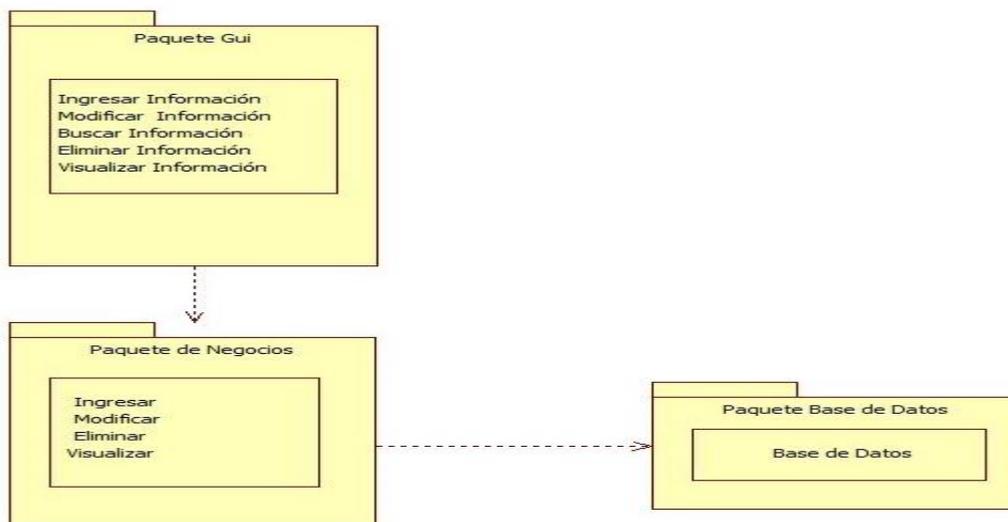


Figura 53 Diagrama de clases general de la lógica de dominio

4.8.4 Diagrama de estado

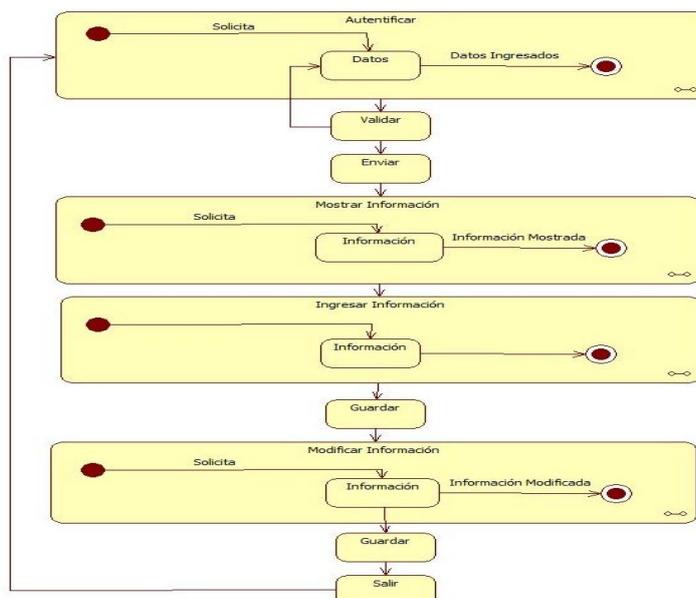


Figura 54 Diagrama de clases general de la lógica de dominio

4.8.5 Base de datos en Visual Studio

Los TableAdapters comunican la aplicación con la base de datos, conectándose a ésta, ejecutando consultas y *devolviendo un DataTable* relleno con los datos requeridos en la consulta. También permiten la inserción y actualización de elementos de la BD. Un TableAdapter puede contener todas las consultas que se desee, siempre y cuando éstas devuelvan datos con el mismo esquema, mostrándose cada consulta agregada como un método público que puede invocarse como cualquier otro. Para ello, cada vez que se quiera invocar alguno de esos métodos.

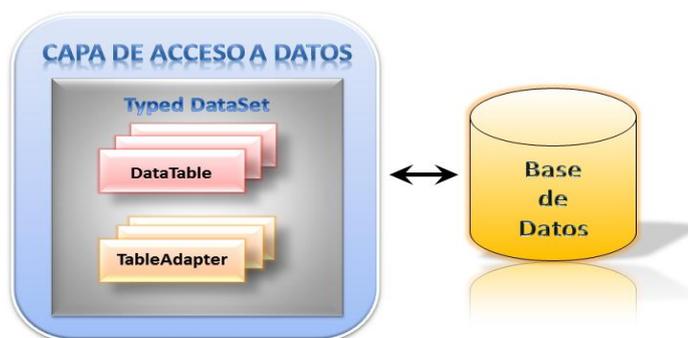


Figura 55 Bases de datos en MS Visual Studio

A continuación, se ofrece el **diagrama de la base de datos (Fig 56)** que da soporte a la herramienta. Para el diseño del esquema de esta base de datos se ha seguido el **patrón *Una clase, una Tabla***, que permite la construcción del esquema de la base de datos partiendo de la lógica de negocio.

4.8.6 Aspectos a señalar

A continuación se mencionan algunos aspectos tenidos en cuenta en el desarrollo de la herramienta.

4.8.7 Interfaz gráfica

Esta interfaz se ha planificado como un *entorno de usabilidad*, en la que los usuarios de manera rápida y cómoda puedan acceder a aquellos datos que necesiten sobre cualquier elemento, y realizar las operaciones deseadas. En su diseño se ha intentado cuidar, tanto su aspecto como la facilidad en la realización de las operaciones interactuando con ella, ofreciendo al usuario en todo momento, únicamente, la información necesaria.

También debe señalarse la adaptabilidad de los formularios de ISM², ocupando siempre la zona centro-superior, al utilizar la herramienta en *resoluciones de pantalla mayores* a la definida para ellos.

infoSOLAR ESPOCH
MEDIDOR DE ÍNDICES UV EN TIEMPO REAL

INICIO QUIENES SOMOS ANTECEDENTES BASE CIENTÍFICA ESTADÍSTICAS GALERÍA CRÉDITOS CONTACTO

MODELO SOFTWARE PARA EL MONITOREO Y DIFUSIÓN DE LOS ÍNDICES DE RADIACIÓN SOLAR EN LA ESPOCH

NÚMERO DE ÍNDICE UV EQUIVALENTE SEGÚN LA NWS - EPA	DESCRIPCIÓN	CATEGORÍA DE EXPOSICIÓN (OMS)	RECOMENDACIONES DE LA OMS PARA PROTEGERSE DE LA RADIACIÓN UV	EXPOSICIÓN GRÁFICA
2	No hay peligro para la persona promedio.	BAJA	No necesita protección. Puede mantenerse en el exterior sin riesgo.	

FECHA/HORA DE MONITOREO: 2015-12-04 01:06:43

ÍNDICE UV	0 - 2	3 - 4	5 - 6	7 - 9	10 +
CATEGORÍA DE EXPOSICIÓN	BAJA	MODERADA	ALTA	MUY ALTA	EXTREMA

LOCALIZACIÓN

Escuela Superior Politécnica de Chimborazo
Antonio de Alcedo

Posición del sensor de radiación UV

CONTADOR DE VISITAS

[contador de visitas](#)

Figura 56 Pantalla Inicial de sesión de la herramienta

4.8.8 Informes y disponibilidad en internet

Otra de las peticiones realizadas fue la de poder obtener una vista general con todos los valores para la Organización en su conjunto. La herramienta ha satisfecho esta petición. Una vez terminada, la herramienta web ha sido subida a un servidor para que pueda ser *accesible y utilizable mediante un navegador Web* instalado en el equipo usado. Así, para acceder a ISM debe hacerse mediante la dirección Web <http://www.infosolarespoch.com>. Todo lo relacionado con el acceso y subida de la herramienta.

El tiempo de desarrollo se ofrece la determinación de los tiempos empleados en el desarrollo del en su conjunto, ayudándose para ello de diagramas de Gantt, y mostrando los tiempos destinados para cada tarea. Estos diagramas de Gantt han sido construidos con la herramienta gratuita *Gantt Designer v2*.

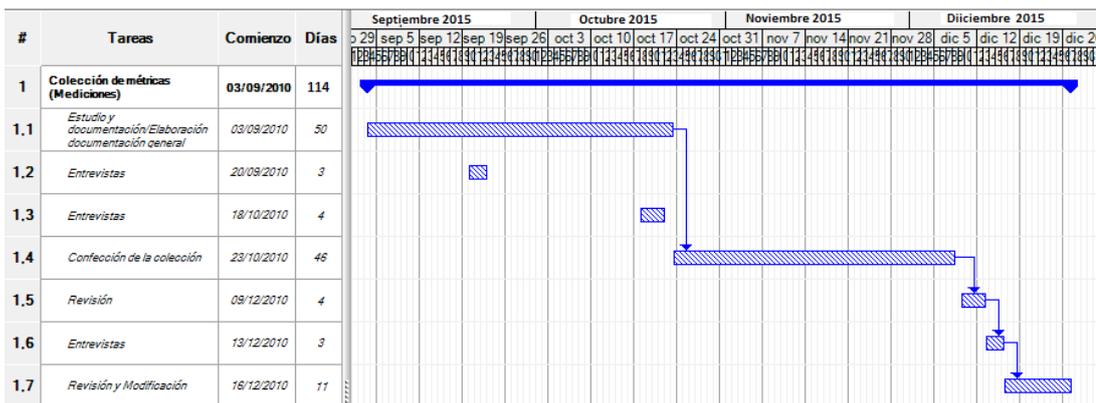


Figura 57 Diagrama de Gantt (Mediciones)

En la *Figura 59* se ofrecen los tiempos correspondientes a la iteración 1 y la iteración 2 del desarrollo de la herramienta ISM².

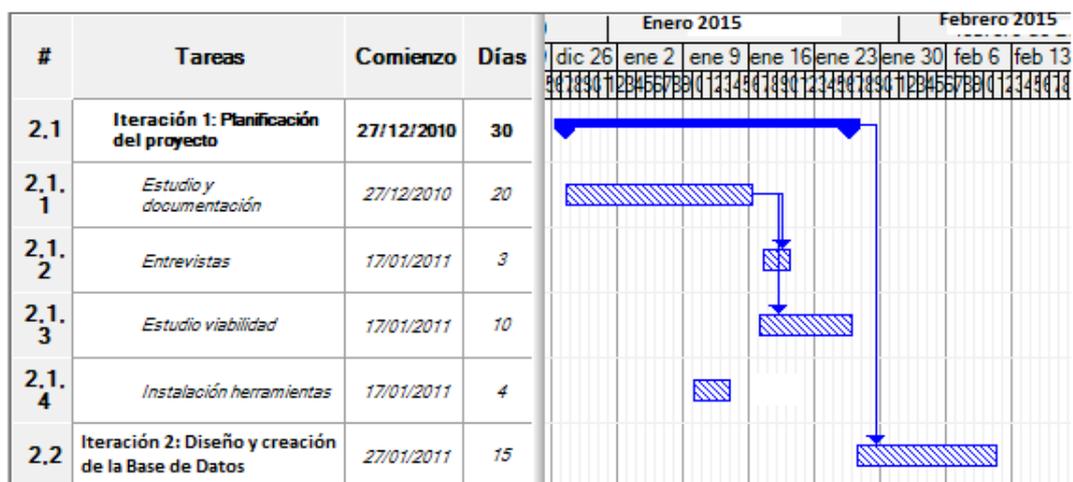


Figura 58 Diagrama de Gantt iteraciones 1 y 2 de la construcción de la herramienta

En la *Figura 60* pueden verse los tiempos dedicados a la iteración 3 del desarrollo de la herramienta ISM².

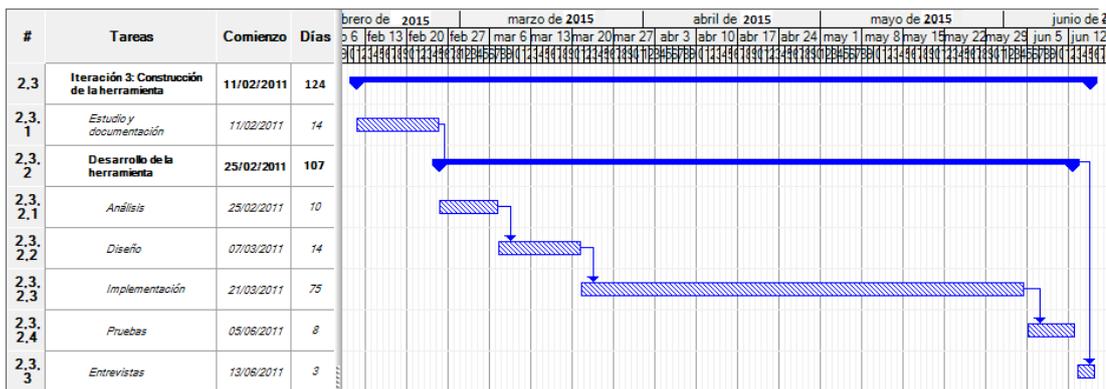


Figura 59 Diagrama de Gantt iteración 3 de la construcción de la herramienta

La *Figura 61* muestra los tiempos para la iteración 4 del desarrollo de la herramienta ISM².

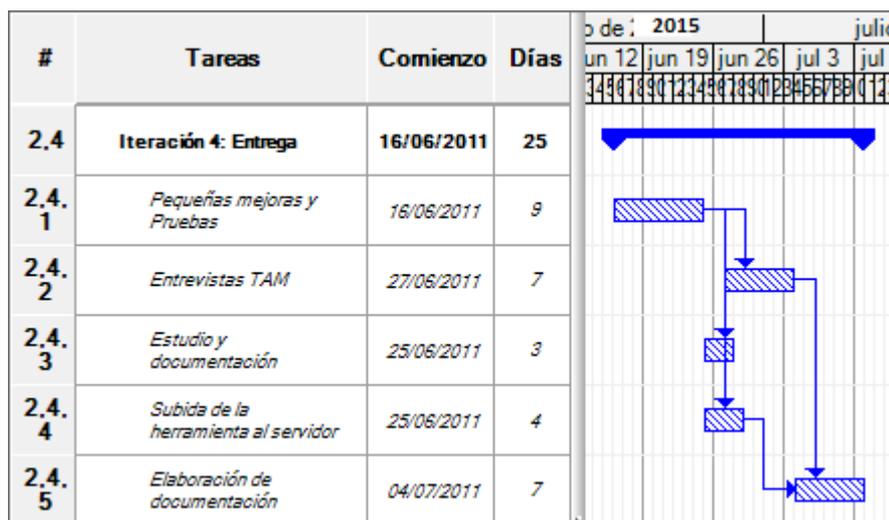


Figura 60 Diagrama de Gantt iteración 4 de la construcción de la herramienta

En la *Figura 62* se ofrecen los tiempos para el caso de estudio y últimas tareas del PFC.

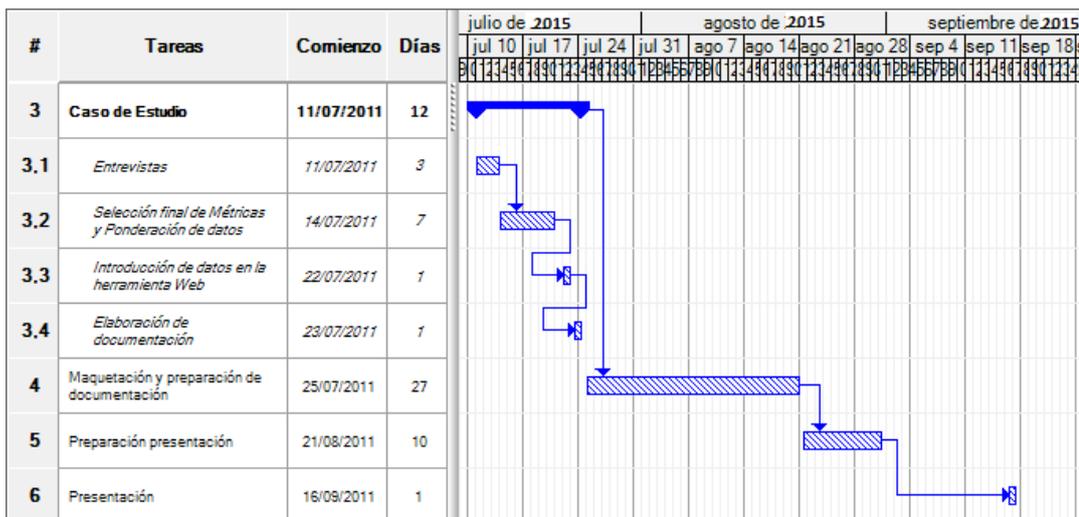


Figura 61 Diagrama de Gantt Caso Estudio – Fin PFC

En total, como ha podido comprobarse, el proyecto inició a desarrollarse el 03/09/2014 y concluyó el 13/12/2015 con su presentación, En total, se han destinado 358 días (**Tabla 4.26**) al desarrollo de todas las partes del PFC.

Tabla 17
Resumen coste temporal del Proyecto

TAREAS	DÍAS
1. Colección de Métricas (Mediciones)	114
2. Construcción de la herramienta web	194
2.1. Iteración 1	30
2.2. Iteración 2	15
2.3. Iteración 3	124
2.4. Iteración 4	25
3. Caso Estudio	12
4. Maquetación y preparación de documentación	27
5. Preparación de presentación	10
6. Presentación	1
TOTAL	358

4.9 Implementación del software y hardware para predicción del índice de radiación solar.

Para el cumplimiento de los objetivos del presente proyecto se parte de la metodología descrita en el Capítulo I, para lo cual fue imprescindible Investigar las técnicas de inteligencia artificial que pueden aplicarse a la predicción de series temporales no lineales, paso seguido se determinó la técnica IA con capacidad de predicción de radiación solar.

La implementación del sistema electrónico de adquisición de datos que permita obtener los datos del potencial de radiación solar se hace referencia en el presente capítulo, con lo cual se logrará construir el modelo software de predicción. Comenzando con la búsqueda de centros meteorológicos en el Ecuador que permitan posteriormente la validación del modelo junto con los Índices establecidos por la OMS según indica la Figura 63, se localiza la página oficial Agencia Espacial Civil Ecuatoriana (EXA) la misma que provee información satelital de la radiación solar en algunas provincias

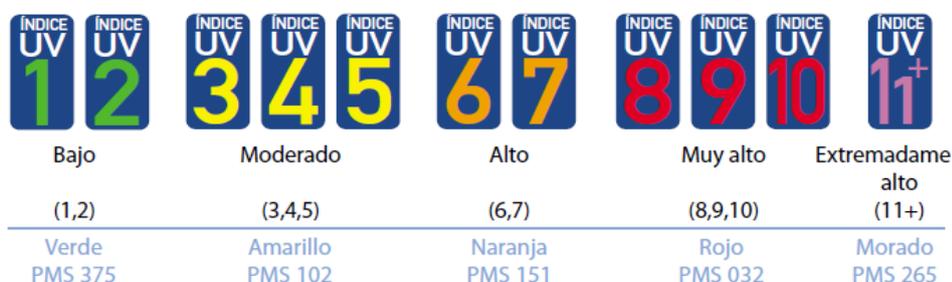


Figura 62 Código internacional de colores para comunicar el IUV.

Fuente: (OMS, 2010)

El índice de radiación UV (IUV) se calcula tomando en consideración. El espectro de acción de referencia de la Comisión Internacional sobre Iluminación (CIE) para el eritema inducido por la radiación UV en la piel humana (ISO 17166:1999/CIE S 007/E-1998) y el IUV se define y calcula para una superficie horizontal, El IUV es adimensional, la fórmula para calcular el IUV es:

$$I_{uv} = K_{er} \int_{250 \text{ nm}}^{400 \text{ nm}} E_{\lambda} \cdot S_{er}(\lambda) d\lambda \quad (4.1)$$

Dónde:

$E\lambda$, es la irradiancia espectral solar expresada en $W/(m^2.nm)$ a la longitud de onda λ .

$d\lambda$, es el diferencial de longitud de onda utilizado en la integración.

$S_{er}(\lambda)$, es el espectro de acción de referencia para el eritema.

K_{er} , es una constante igual a $40m^2/W$.

Hay dos métodos para determinar el IUV, y son mediante mediciones y mediante cálculos basados en modelos matemáticos. En esta investigación se va a predecir el IUV a través de una regresión elaborada con los datos obtenidos mediante el segundo método de determinación del IUV adaptada a SVM, algunos estudios sobre la predicción de series de tiempo lineales se muestran en el capítulo III Tabla 7, sobre comparación de trabajos de predicción de radiación solar, y mediante históricos, para lo cual en primera instancia se construirá un sistema de adquisición de datos de tal forma que provea los datos de entrada para el modelo de predicción.

4.9.1 Sistema de adquisición de datos

Un sistema electrónico de adquisición de datos (SAD), es diseñado para adquirir la información de una variable o variables físicas.

Estas variables físicas que ha sido o han sido convertidas a variables eléctricas generalmente como voltaje o corriente, consta de 5 bloques como se observa en la Figura 64

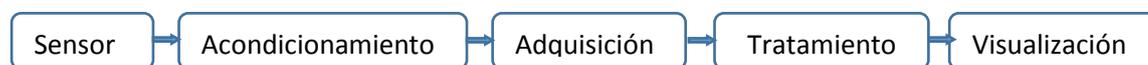


Figura 63 Diagrama de bloques de un sistema SAD.

Fuente: (EXA, 2010)

El sensor, recibe en su entrada una variable física y convierte esta en una variable eléctrica (corriente o voltaje). El acondicionamiento de la señal, acondiciona la señal de salida del sensor a las características del bloque de adquisición, el acondicionamiento puede consistir en: filtrado, amplificación, conversión y, aislamiento eléctrico de la señal. La adquisición, consiste en

convertir la señal analógica en digital, mediante un convertidor analógico a digital (CA/D).

4.9.2 Sensores de la radiación UV

Un sensor es un dispositivo que toma como entrada una variable física y convierte esta en una variable eléctrica, generalmente en un voltaje o una corriente. Los sensores más adecuados son los que tienen una relación lineal entre la variable física y la variable eléctrica, deben ser convertidos al índice IUV respectivo, además se debe diseñar un circuito de acondicionamiento de la señal eléctrica.

Así como también la respectiva amplificación, este tipo de sensores son más complicado utilizar por la circuitería electrónica adicional que requieren, como amplificadores operacionales, resistencias, capacitores, 2 circuitos típicos se ilustra en la Figura 65.

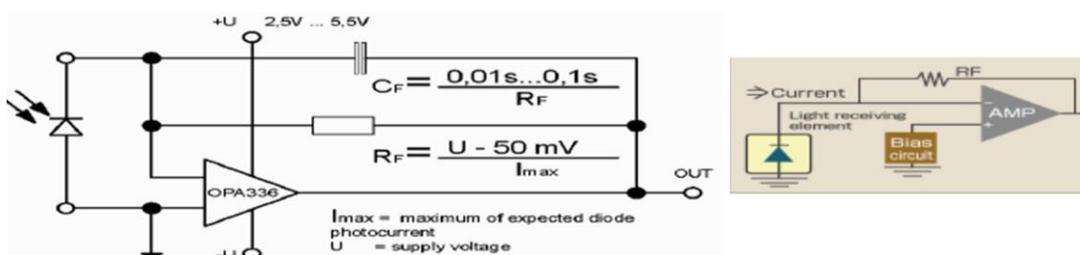


Figura 64 Circuitos de acondicionamiento de la señal del sensor UV

Fuente: (Sic UV, 2010)

Otros sensores más sofisticados incluyen dentro del sensor el acondicionamiento y amplificación de la señal por lo que pueden directamente ser utilizados. En la parte izquierda de la Figura 66 en la columna "Chip Area" se puede ver la Figura de una hormiga, lo cual permite comparar el tamaño del chip (sensor) con el tamaño de una hormiga. y se muestra una captura de la pantalla de la página web llena con las siguientes características del sensor:

- Área de la superficie: 4 mm^2 .

- Banda espectral: banda ancha.
- Empaquetamiento: siso90.
- Característica especial: giga.

SiC UV Photodiode - Configurator

Configure your SiC UV Photodiode here. All SiC Photodiodes can be permanently operated at up to 170°C (338°F). The selection criteria are Chip Area, Spectral Response, Packaging and additional Special features. Moving of the cursor on a picture gives a short information about the particular feature. Please do a selection of each column starting with the left column. The Configurator informs about the product's name and price, the datasheet link and gives the opportunity to create an order message. Please download a comprehensive Tutorial [here](#).

Chip Area	Spectral Response	Packaging	Special
S (0,06 mm ²) € 25,00	- (Broad Band) + € 0,00	18 € 0,00	Standard standard version + € 0,00
M (0,20 mm ²) € 49,00	A (UVA) + € 24,00	18ISO90 + € 6,00	Lens (concentrator) + € 14,00
D (0,50 mm ²) € 49,00	B (UVB) + € 24,00	18S + € 4,00	MEGA (0,5 W/cm ²) + € 30,00
L (1,00 mm ²) € 105,00	C (UVC) + € 24,00	5 + € 4,00	GIGA (7 W/cm ²) + € 30,00
XL (4,00 mm ²) € 404,00	E (UV-Index) + € 34,00	5ISO90 + € 6,00	

Selection Reset

Figura 65 Opciones de selección del fotodetector.

Fuente: (Sic UV, 2010)

En la Figura 67 se ve el resultado del foto detector es el SGO1XL-5ISO90GIGA, el precio del dispositivo es 440 euros.

Chip Area	Spectral Response	Packaging	Special
S (0,06 mm ²) € 25,00	- (Broad Band) + € 0,00	18 € 0,00	Standard standard version + € 0,00
M (0,20 mm ²) € 49,00	A (UVA) + € 24,00	18ISO90 + € 6,00	Lens (concentrator) + € 14,00
D (0,50 mm ²) € 49,00	B (UVB) + € 24,00	18S + € 4,00	MEGA (0,5 W/cm ²) + € 30,00
L (1,00 mm ²) € 105,00	C (UVC) + € 24,00	5 + € 4,00	GIGA (7 W/cm ²) + € 30,00
XL (4,00 mm ²) € 404,00	E (UV-Index) + € 34,00	5ISO90 + € 6,00	

Selection Reset

Product Name	Product Net Price
SG01XL-5ISO90GIGA	€ 440,00
Download Datasheet SG01XL-5ISO90GIGA	Show Contact Details Create Order E-Mail

Figura 66 Características seleccionadas del fotodetector.

Fuente: (Sic UV, 2010)

En la Figura 68 se muestra las diferentes bandas del espectro UV, si el sensor es de banda ancha será sensible a todo el espectro UV, si es solo sensible al UVA solo responderá a la banda de frecuencias UVA, en general un sensor puede responder a toda la banda UV o selectivamente solo a una o a varias de ellas simultáneamente, por Ej. el sensor ML 8511 responde a las bandas UVA: 315-400nm y UVB: 280-315nm

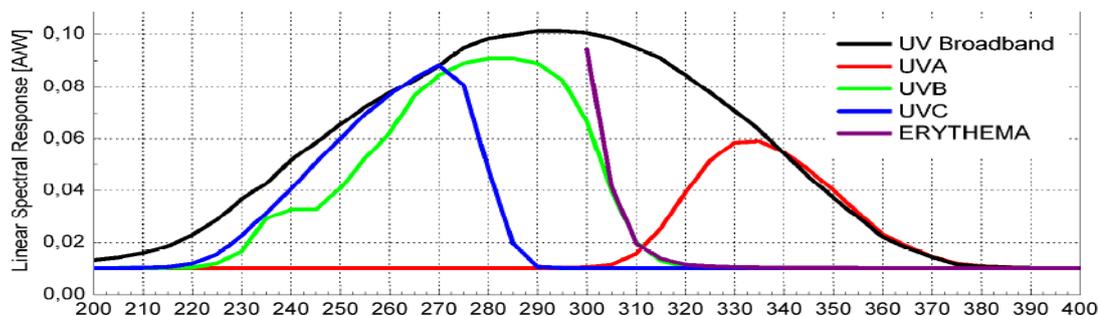


Figura 67 Respuesta espectral de un fotodetector.

Fuente: (Sic UV, 2010)

4.9.3 EL UVI-01 DETECTOR DE RAYOS ULTRAVIOLETA

El UVI-01 fabricado por REYAX TECHNOLOGY CO. LTD., es sensible a la banda UVA y UVB, y genera una salida de voltaje que da directamente el índice ultravioleta (IUV) es apropiado para instrumentos de medición de la radiación ultravioleta, sus características relevantes son:

- Intervalo de temperatura ambiente de operación: mínimo -25°C , máximo $+70^{\circ}\text{C}$.
- Intervalo de temperatura de almacenamiento: mínimo -40°C , máximo $+70^{\circ}\text{C}$.
- Máxima temperatura de la unión: $+260^{\circ}\text{C}$.
- Humedad: mínimo 5 %RH, máximo 85 %RH.
- Humedad: mínimo 5 %RH, máximo 85 %RH.

En la Figura 69 se puede observar el circuito de aplicación recomendado por el fabricante en los dos sensores expuestos anteriormente, como se puede ver el circuito es muy sencillo y fácil de implementar, evitando la utilización de amplificadores operacionales, por lo tanto no requiere de una fuente de poder y el consumo de energía del sensor es cero vatios, esta característica es fundamental sobre todo para mediciones remotas y continuas.

El valor de la resistencia puede ser elegido para modificar la pendiente de la recta que da la relación entre el voltaje y el IUUV, en la Figura 69 se ve la recta para diferentes valores de resistencia.

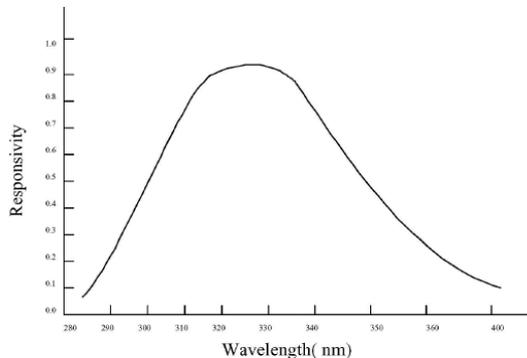
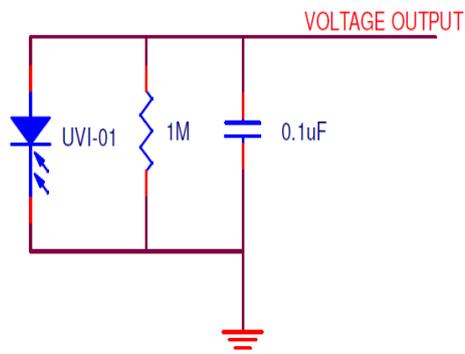


Figura 68 Circuito de aplicación para el SAD/
 Fuente: (Sic UV, 2010)

4.9.4 SIMULACIÓN EN PROTEUS DE LA PROPUESTA DEL CIRCUITO PARA MEDICIÓN DEL INDICE DE RADIACIÓN SOLAR

En la Figura 69 se puede muestra la simulación lógica del circuito mencionado en el apartado de la Figura 70 del Circuito de aplicación para el SAD, en la que toma como referencia el valor medido, el valor esperado y el error en el valor medido, en la simulación física del circuito se puede observar en la Figura 70.

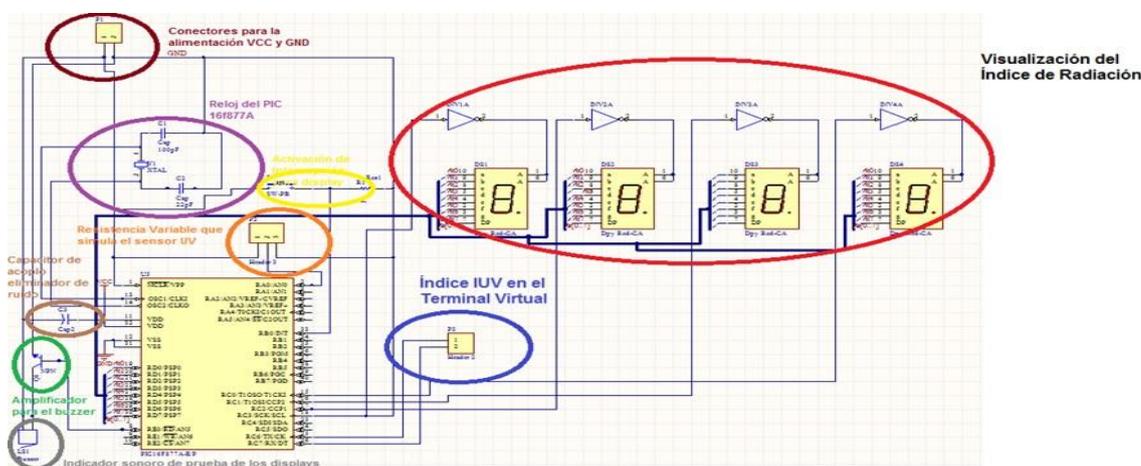


Figura 69 Simulación circuito lógico del SAD

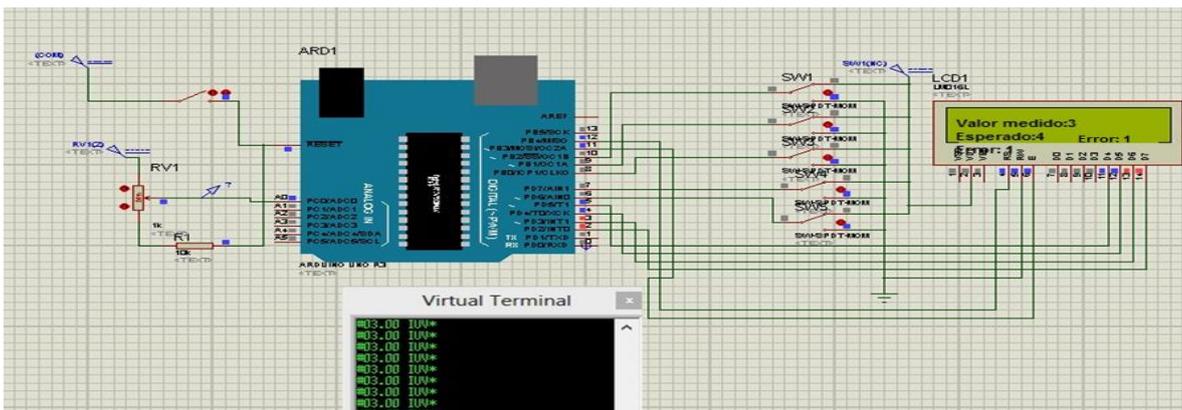


Figura 70 Simulación circuito físico del SAD

En la Figura 71 se puede observar el diagrama de conexiones del SAD, el mismo que contiene un LCD y un display, para mostrar los datos medidos.

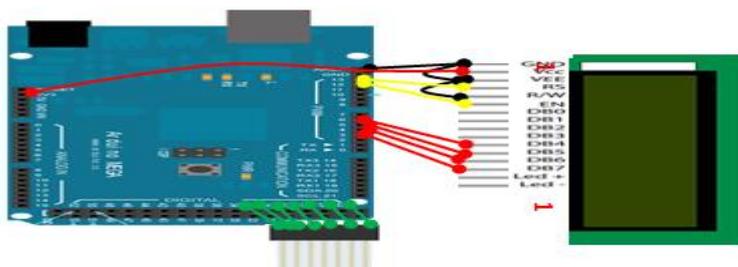


Figura 71 Conexión del LDC a los puertos del Arduino.

4.10 APROXIMACIÓN DEL PCB EN ALTIUM DEL CIRCUITO PARA MEDICIÓN DEL INDICE DE RADIACIÓN SOLAR

En cumplimiento de las normas, criterios, la organización de componentes digitales y analógicos y la etapa de visualización de datos se simula la construcción del diagrama circuital del SAD como se ilustra en la Figura 73

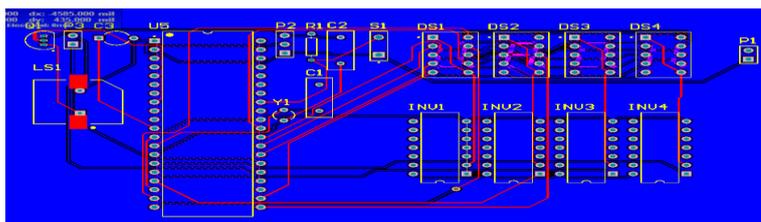


Figura 72 Construcción del diagrama circuital del SAD

4.10.1 SOFTWARE DEL DISPOSITIVO ELECTRÓNICO DE PREDICCIÓN DEL IUV.

La predicción del IUV se realiza mediante la evaluación de la ecuación de predicción del IUV, que fue desarrollada en el capítulo 2, para que el Arduino Mega 2560 realice todas las operaciones así como la lectura y generación de datos se basa en la creación de un programa.

El software para Arduino se desarrolla a través del Sketch, con facilidades para escribir, editar y, compilar un programa que se va a descargar en un Arduino, El diagrama de flujo se puede observar en el anexo 1 y el código del programa para la predicción del IUV en la Epoch está desarrollado en el lenguaje para Arduino mediante sketch y se puede revisar en el Anexo 2.

4.10.2 Selección del punto de referencia en la medición

El punto de referencia escogido en Riobamba es la Epoch. Como se muestra en la Figura 74 Debido a que en la Institución existe un centro meteorológico que provee suficiente información sobre las otras variables climáticas como nubosidad, precipitaciones entre otras, permitiendo utilizarlas para la validación del modelo. Los datos serán enviados inalámbricamente al data center de la Institución de forma que se asegure la continuidad de la serie de tiempo para la predicción.



Figura 73 Ubicación punto de referencia ESPOCH.

Fuente: (Google Earth, 2015)

4.10.3 PREDICCIÓN DE LA RADIACIÓN SOLAR EMPLEANDO MODELO KERNEL GAUSIANO DE FUNCIÓN BASE RADIAL (RBF)

De acuerdo a Baldeón y coautores (2013), los datos climáticos se obtuvieron de la base de datos de la investigación “MODELO DE PREDICCIÓN IUV Y SU APLICACIÓN A LA SALUD”, esta base de datos contiene registros en el periodo comprendido entre enero del 2012 y marzo del 2015, teniendo en cuenta que la medición es del potencial de radiación solar y el tiempo en sus 24 horas del día por cada año, convirtiéndose en variables de entrada al modelo.

Son Kernels de propósito general utilizados de acuerdo a diversas implementaciones que se han realizado de las SVM según se indica en el Capítulo III tabla 7 se ha reportado que, para el caso de clasificación y regresión, el Kernel de función de base radial RBF ha sido el que mejor se ajusta evitando el sobre entrenamiento. (Karatzoglou et al., 2006)

$$K(x, x') = \exp(-\sigma \|x - x'\|^2). \quad (4.2)$$

4.10.4 Clasificación

El detector de radiación ultravioleta tiene 5 rangos numerados del 1 al 5 y son los *valores del número de índice* (VNI), según se indica en la Figura 75, los datos de la radiación UV en forma de VNI se toman dentro de la politécnica y, se tabulan siguiendo el formato de la tabla 8.

Exposure Category	Index Number	Valor del número del Índice
MINIMAL	0-2	1
LOW	3-4	2
MODERATE	5-6	3
HIGH	7-9	4
VERY HIGH	10+	5

Figura 74 Índice UV según la NWS y EPA

Fuente: (OMS, 2004.)

Tabla 18
Formato de recolección de datos

AÑO/MES/DIA			
Hora	VNI	Hora	VNI
6		12,5	
6,5		13	
7		13,5	
7,5		14	
8		14,5	
8,5		15	
9		15,5	
9,5		16	
10		16,5	
10,5		17	
11		17,5	
11,5		18	
12		18,5	

Los resultados diarios de los datos se tabulan en tablas según el formato de la tabla VIII, los datos se toman cada día desde las 6h00 hasta las 18h30 en intervalos de media hora.

La tabla VIII contiene los valores más altos de VNI obtenidos en el mes de enero, aquí se puede observar que los valores más altos de VNI están entre las 10h30 y las 17h00, Correspondiendo en la escala NWS y EPA 2004 al valor ALTO, y su equivalente en la escala modificada corresponde a MUY ALTO, los valores más bajos están entre las 6h00 y las 8h30, así como también entre las 17h30 y las 18h30 y representan a un VNI mínimo que su equivalente en la escala modificada corresponde a BAJO. El mayor peligro de radiación UV está entre las 10h30 y las 16h00 en el mes de enero

No se han registrado valores de VNI iguales a 5, es decir valores codificados como *muy altos* en la escala antigua y como *extremos* en la escala actual. Sin embargo los valores registrados son peligrosos y hay que tomar las medidas de protección recomendadas por la OMS, como se indica en la Figura 76

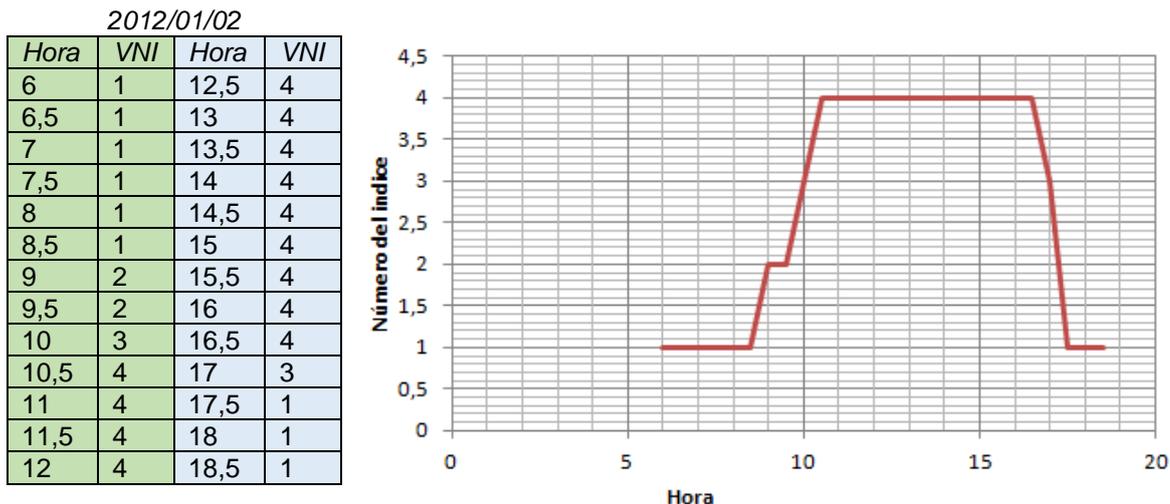


Figura 75 VNI en función de la hora para el valor mayor VNI medido en enero.

Sin embargo el rango puede variar y en su modo general estaría entre las 10h30 y las 15h30 para días con cierto grado de nubosidad. Como no es posible disponer de todas las variables para la predicción, como son nubosidad, lluvia entre otras por practicidad y sin pérdida de precisión significativa, se considera.

$$UVI = UVIO * CMF * (1 + 0.08 * \Delta h) \quad (4.3)$$

En donde:

UVI, es el índice ultravioleta pronosticado.

UVIO, es el índice IUV para cielo claro y a nivel del mar.

CMF, es el factor de corrección según el tipo de nubes.

Δh , es la altura a la cual se va a pronosticar el IUV.

A pesar de que la ecuación (3,51) hace referencia directa y exclusiva al IUV, se va a utilizar esta para corregir el modelo del voltaje, esto es factible debido a que el IUV calculado es directamente proporcional al voltaje generado por el sensor como se demostró en la Figura 77.

Valor Constate

$$IUV = 78.1 v.$$

En donde el voltaje está en milivoltios (mv) y v es la ecuación 3.51 (v(t)).

Se dividen los 2 lados de la ecuación por la constante 78.1, y se tiene la ecuación 3.53

$$UVI / 78.1 = UVI0 / 78.1 * CMF * (1 + 0.08 * \Delta h) \quad (4.4)$$

$UVI / 78.1$, es el voltaje (v) de la ecuación 2.2, y $UVI0 / 78.1$ sería el voltaje medido a nivel del mar. La ecuación 2.4 se transforma en la ecuación 2.54.

$$v = v_0 * CMF * (1 + 0.08 * \Delta h) \quad (4.5)$$

En donde: v es el voltaje pronosticado o calculado y, v_0 el voltaje medido a nivel del mar.

El modelo del voltaje para el sensor UV-01 queda de la siguiente manera.

$$v = (a_1 * e^{-((\frac{t-b_1}{c_1})^2)} + a_2 * e^{-((\frac{t-b_2}{c_2})^2)}) * CFM \quad 6 \leq t \leq 18 \quad (4.6)$$

Puesto que $CFM = 1$, para las octavas 0-2 y nubosidad alta, media y baja, y octavas 3-4 y nubosidad alta y media y, para las octavas 5-6 con nubosidad alta, el voltaje de predicción es:

$$v = v_1 \quad (4.7)$$

Las ecuaciones 2.53 y 2.55 son el aporte original de esta investigación. Estas ecuaciones predicen el voltaje generado por la radiación UV-A y UV-B para cielo claro y nublado y, partir de estas ecuaciones se puede predecir el IUV, estas ecuaciones son locales y pero válidas para cualquier lugar en donde se hagan las mediciones del voltaje.

Sin embargo en el proceso de transformación del voltaje v_1 al IUV todos los sensores deben dar el mismo resultado del IUV.

Por lo tanto no importa el modelo del sensor (hay un modelo para cada tipo de sensor) que se utilice ya que el objetivo final es determinar el IUV y, todos los sensores deben coincidir en este valor. Incluso un mismo sensor

puede tener varios modelos, como es el caso del IV-01, cuyo modelo depende de la resistencia de carga RI como se demostró en el capítulo anterior. Se calculan los IUV con la ecuación 3.51 para las ciudades de Quito y Cuenca con la finalidad de verificar el error que se tendría al utilizar la ecuación y por ende el error que se podría producir al evaluar el IUV, por Ej. :

Quito:

Altura= 2.8 Kms.

CMF= 1 (cielo claro).

UVI₀= 12.

$$\text{IUV (Quito)} = 12 * 1 * (1 + 0.08 * 2.8) = 14.69. \quad (4.8)$$

El IUV se expresa como un entero redondeado, por lo tanto el IUV calculado es 15. El valor medido de acuerdo al historial de la Agencia Espacial Civil Ecuatoriana (EXA) es 16, para los primeros días del enero.

$$\text{IUV (Cuenca)} = 12 * 1 * (1 + 0.08 * 2.560) = 14.46. \quad (4.9)$$

El IUV para Cuenca que se encuentra a 2560 metros sobre el nivel del mar, es 15. El valor medido de acuerdo al historial de la Agencia Espacial Civil Ecuatoriana (EXA) es 16, para los primeros días de enero. Los resultados del IUV para las 2 ciudades, nos da una buena idea de que la ecuación 2.1 da un error de una unidad en el cálculo del IUV, como se indicó antes esto no es raro debido a la incertidumbre de los parámetros atmosféricos.

4.10.5 Pre procesamiento de datos

Para determinar la tendencia de los datos se grafica el NIUV en función del tiempo, en la Figura 77 se pueden ver varios de estos gráficos, la tendencia es claramente hacia una parábola.

Tabla 19

VALORES MÁS ALTOS DE NIUV OBTENIDOS EN ENERO

Hora	6	6,5	7	7,5	8	8,5	9	9,5	10	10,5	11	11,5	12
NIUV	1	1	1	1	1	1	2	2	3	4	4	4	4

Hora	12,5	13	13,5	14	14,5	15	15,5	16	16,5	17	17,5	18	18,5
NIUV	4	4	4	4	4	4	4	3	3	3	1	1	1

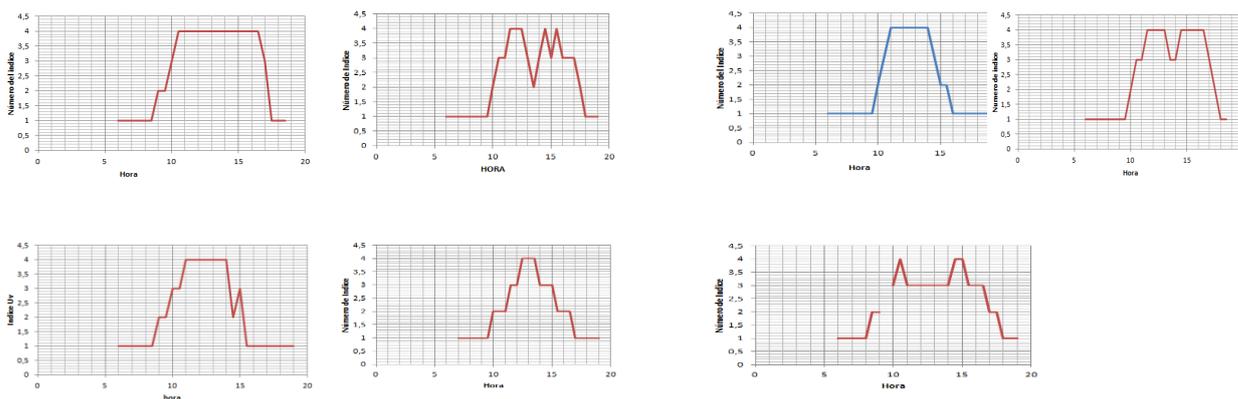


Figura 76 Gráficos del NIUV en función del tiempo.

Fuente: (Baldeón. W, Mora. V , 2012)

Para determinar la relación matemática entre los puntos graficados se utiliza como primera aproximación la *parábola de mínimos cuadrados*, cuya ecuación es.

$$NIUV = Y(t) = a_0 + a_1 \cdot t + a_2 \cdot t^2 \tag{4.10}$$

Las constantes a_0 , a_1 , a_2 , se determinan resolviendo simultáneamente el sistema de ecuaciones 2 en donde $N = 26$, es el número de mediciones realizadas cada día.

$$\begin{aligned} \sum_{n=1}^{26} (Y) &= a_0 \cdot 26 + a_1 \cdot \sum_{n=1}^{26} (t) + a_2 \cdot \sum_{n=1}^{26} (t^2) \\ \sum_{n=1}^{26} (tY) &= a_0 \cdot \sum_{n=1}^{26} (t) + a_1 \cdot \sum_{n=1}^{26} (t^2) + a_2 \cdot \sum_{n=1}^{26} (t^3) \end{aligned} \tag{4.11}$$

$$\sum_{n=1}^{26} (t^2 Y) = a_0 \cdot \sum_{n=1}^{26} (t^2) + a_1 \cdot \sum_{n=1}^{26} (t^3) + a_2 \cdot \sum_{n=1}^{26} (t^4)$$

Como los valores NIUV varían día a día y de hora a hora se determinan los valores representativos de los datos medidos (las tablas son extensas) mediante su promedio y esto se indica en la tabla X, el promedio se redondean puesto que el NIUV debe ser un número entero.

Tabla 20
Promedio De Los Datos Niuiv Medidos

Hora	6	6,5	7	7,5	8	8,5	9	9,5	10	10,5	11	11,5	12
NIUV	1	1	1	1	1	1	1	1	2	3	3	4	4

Hora	12,5	13	13,5	14	14,5	15	15,5	16	16,5	17	17,5	18	18,5
NIUV	4	4	3	3	3	3	3	3	2	2	2	3	1

Fuente: (Baldeón. W, Mora. V , 2012)

El gráfico de los datos de la tabla IV se puede ver en la Figura 78, por lo que el modelo a determinarse debe representar a ese conjunto de puntos dispersos.

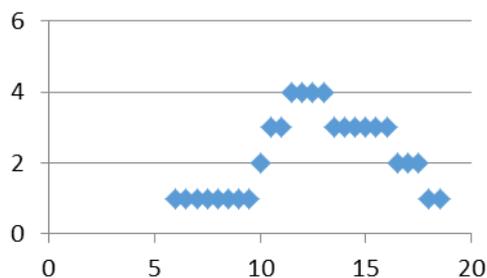


Figura 77 Gráfico del NIUV en función del tiempo para la tabla 4

Para determinar los valores de a_0 , a_1 y a_2 de la ecuación 4.11, se utiliza la caja de herramientas de ajuste de curvas disponible en MATLAB, obteniendo los siguientes valores:

$$a_0 = -8.275$$

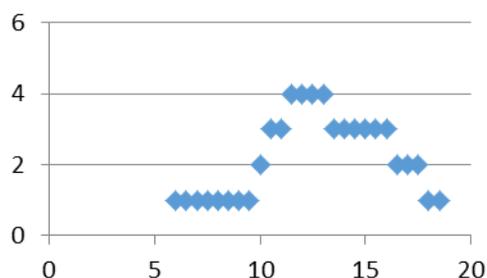
$$a_1 = 1.774$$

$$a_2 = -0.06838$$

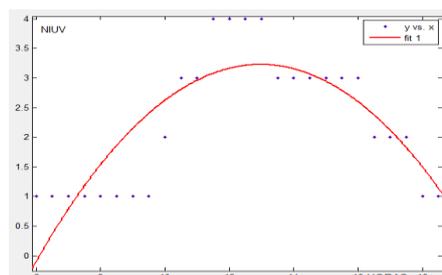
Con estos valores el modelo matemático queda de la siguiente manera, ecuación 3.61

$$Y(t) = - 8.275 + 1.774 * t - 0.06838 * (t)^2 \quad (4.12)$$

embargo al graficar la ecuación 3, el resultado es el grafico b de la Figura 79, como se observa sobre en el punto central no es una muy buena aproximación puesto que el valor debería estar muy cerca de 4, por lo que se decide probar un polinomio de mayor grado, se prueban polinomios de grado 3, 4, 5, 6, 7, 8, 9 y 10 siendo este último una mejor aproximación a la curva deseada por lo que el modelo matemático queda como se indica en la ecuación 4.



a



b

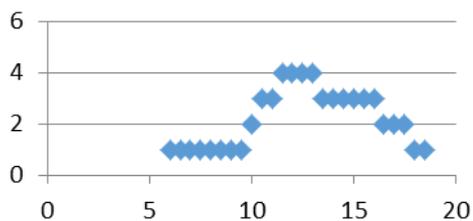
Figura 78 Gráficos del NIUV a) forma deseada, b) forma obtenida

$$Y(t) = a_9t^9 + a_8t^8 + a_7t^7 + a_6t^6 + a_5t^5 + a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0 \quad (4.13)$$

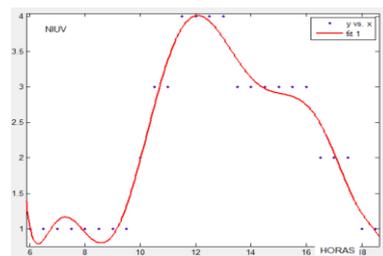
$$Y(t) = -0.507t^9 + 0.506t^8 + 3.44t^7 - 3.157t^6 - 7.96t^5 + 6.789t^4 + 6.596t^3 - 6.364t^2 - 0.617t + 3.992$$

Para $6 \leq t \leq 18.5$, y valores enteros de $Y(t)$, iguales a: 1,2,3,4,5. Los valores fraccionarios resultado de calcular $Y(t)$ deben ser aproximados a valores enteros; puesto el NIUV es entero por definición.

La función t esta normalizado a 12.25 con una desviación estándar de 3.824. Los coeficientes tienen límites de confianza del 95%. Como se puede ver en la Figura 79, el gráfico b es una mejor aproximación al gráfico a, de la misma Figura



a)



b)

Figura 79 Gráficos del NIUV a) forma deseada, b) forma obtenida

La divergencia de las curvas a y b de la Figura 80 sobre todo entre las 6h00 y 9h00 no es de importancia puesto que los valores fraccionarios de $Y(t)$ deben ser ajustados al mejor número entero entre 0 y 5. El número de índice 10+ a pesar de que las nubes actúan como difusoras y amplificadoras de rayos solares no se ha detectado en Riobamba. El NIUV y por ende el IUV es variable y su máximo valor alcanza alrededor del mediodía.

4.10.6 Regresión en función del tiempo

Con mucha frecuencia en trabajos de investigación es necesario para comprender un fenómeno, estudiar la relación que existe entre las variables involucradas en dicho fenómeno.

En el presente estudio las variables involucradas son: la radiación solar, el voltaje, el tiempo (hora del día) y, el IUV. La naturaleza y el grado de relación que hay entre variables, puede ser analizado estadísticamente mediante dos técnicas: regresión y correlación.

El análisis de datos mediante regresión, se utiliza para determinar la relación probable entre variables, relación que es establecida mediante una ecuación matemática, mediante la cual se puede *predecir o estimar* el valor de la variable dependiente para un valor dado de la variable independiente.

La radiación solar que varía en función del tiempo (hora del día) es convertida en voltaje mediante el sensor UV-01, y el voltaje del sensor es

transformado en el IUV. Las variables involucradas directamente en este estudio son: el voltaje, el tiempo y el IUV. Los datos recolectados son migrados a Excel (copiados del monitor serial del PC y pegados en Excel).

Una muestra se indica en la tabla IV, como la tabla es extensa, no se indican los 145 valores sino solo algunos de ellos. En el anexo 1, se puede ver los datos del voltaje en función del tiempo, que se recolectaron para este estudio y se almacenaron en Excel.

Tabla 21
Algunos datos del voltaje migrados a Excel.

Viernes 16-01-2015	
Hora	Voltaje
6,00	0
6,05	0
6,10	0
...	...
7,00	0,001
7,05	0,002
...	...
10,00	0,122
10,05	0,155
10,10	0,149
...	...
11,00	0,163
11,05	0,16
11,10	0,16

Viernes 16-01-2015	
Hora	Voltaje
12,00	0,148
12,05	0,151
12,10	0,135
...	...
13,00	0,133
13,05	0,132
15,10	0,059
16,00	0,035
16,05	0,033
16,10	0,024
...	...
17,00	0,003
17,05	0,004
17,10	0,005
...	...
18,00	0

Antes de proceder a realizar el diagrama de dispersión, se realiza un cambio de escala en el eje del tiempo, de los voltajes recolectados en función de las horas, esto es necesario ya que entre cada hora hay 60 divisiones, es decir 60 minutos, las mediciones se realizan cada 5 minutos o 12 mediciones por hora, mismas que no están distribuidas uniformemente.

El anexo C contiene las mediciones del voltaje tomados del sensor durante los meses indicados.

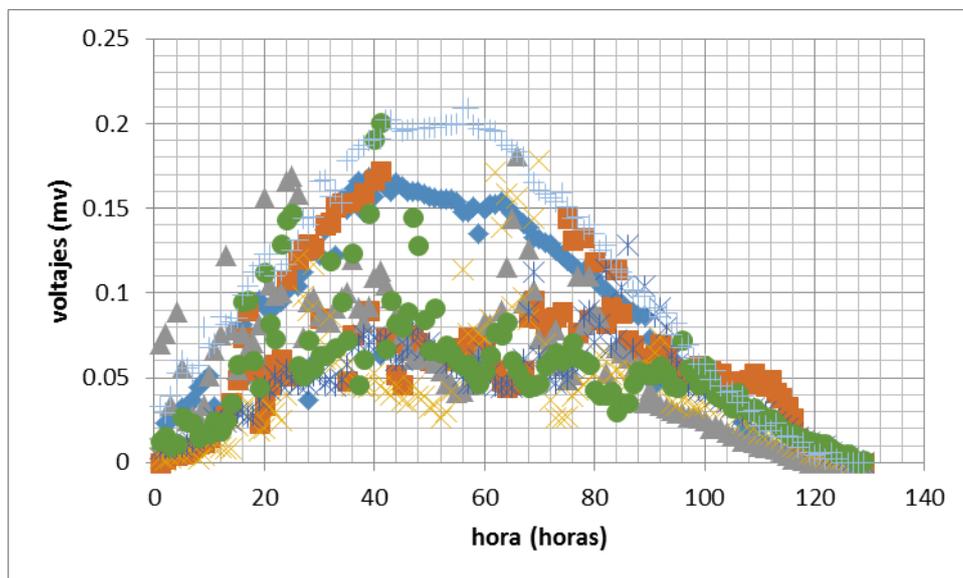


Figura 80 Variación del voltaje en función del tiempo (H.T).

Fuente: (Baldeón. W, Mora. V , 2015)

En la Figura 81 se ve el diagrama de dispersión del voltaje para cielo claro es decir sin la presencia de nubes, la Figura incluye la presencia de nubes.

La única manera de tomar medidas a cielo claro es esperar por un día sin nubes, esto ocurre cuando el cielo es completamente azul durante todo el día.

4.11 Modelos implementados con Matlab

En la Figura 82, se observa el gráfico obtenido mediante la herramienta cftool de Matlab, de los 4 modelos matemáticos posibles para la curva de la Figura 83, el modelo general Fourier 1.

Los dos modelos de Gauss tienden a cero antes de las 6H00 y después de las 18H00, evitando valores negativos, el modelo 1 de Gauss tiene el mismo problema que el modelo 1 de Fourier, no modelan bien entre las 6 y

las 7 de la mañana, el modelo 2 de Gauss representa mejor a la curva en este intervalo, por lo tanto este es el elegido.

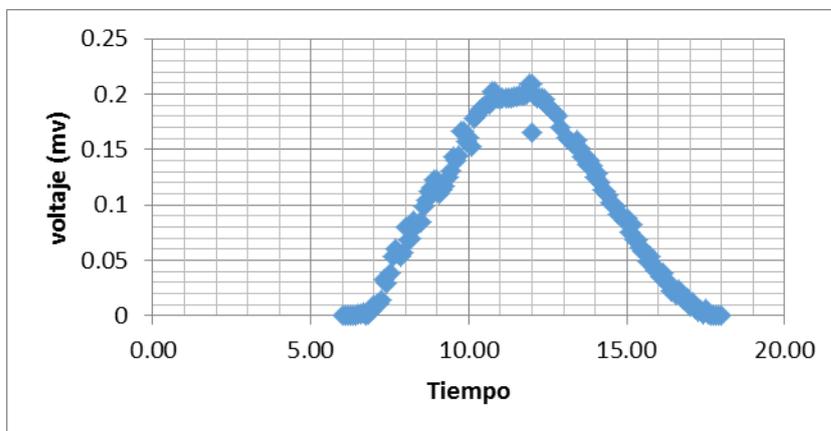


Figura 81 Medidas del voltaje del sensor a cielo claro.

Los modelos de cada curva de la Figura 82 se indican a continuación tal como son generados por Matlab.

4.11.1 Modelo general Model gauss1

$$f(x) = a1 \cdot \exp(-((x-b1)/c1)^2) \quad (4.14)$$

Coefficients (with 95% confidence bounds): a1 = 0.2063 (0.2031, 0.2094)

$$b1 = 11.65 (11.61, 11.69)$$

4.11.2 Modelo general model gauss2:

$$f(x) = a1 \cdot \exp(-((x-b1)/c1)^2) + a2 \cdot \exp(-((x-b2)/c2)^2) \quad (4.15)$$

Coefficients (with 95% confidence bounds):

$$a1 = 0.1686 (0.1264, 0.2108)$$

$$b1 = 12.46 (11.67, 13.26)$$

$$c1 = 2.901 (2.453, 3.349)$$

$$a2 = 0.08764 (0.01377, 0.1615)$$

$$b2 = 9.838 (9.28, 10.4)$$

$$c2 = 2.171 (1.761, 2.581)$$

4.11.3 Modelo general model fourier1:

$$f(x) = a_0 + a_1 \cos(x*w) + b_1 \sin(x*w) \quad (4.16)$$

Coefficients (with 95% confidence bounds):

$$a_0 = 0.1001 \quad (0.0979, 0.1023)$$

$$a_1 = 0.09941 \quad (0.09771, 0.1011)$$

$$b_1 = -0.01041 \quad (-0.02267, 0.001853)$$

$$w = 0.5295 \quad (0.5192, 0.5399)$$

4.11.4 Modelo GENERAL MODEL FOURIER2:

$$f(x) = a_0 + a_1 \cos(x*w) + b_1 \sin(x*w) + a_2 \cos(2*x*w) + b_2 \sin(2*x*w) \quad (4.17)$$

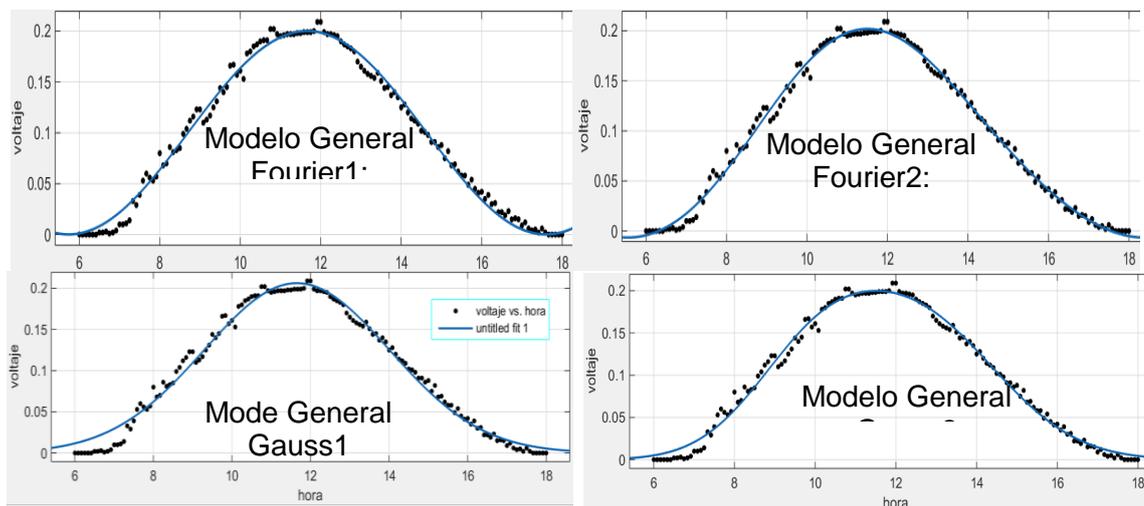


Figura 82 Alternativas de modelos matemáticos

Finalmente el modelo matemático que representa al voltaje del sensor para cielo claro es:

$$v(t) = a_1 * e^{-\left(\frac{t-b_1}{c_1}\right)^2} + a_2 * e^{-\left(\frac{t-b_2}{c_2}\right)^2} \quad 6 \leq t \leq 18 \quad (4.18)$$

En donde:

$$a_1 = 0.1686 \quad (0.1264, 0.2108)$$

$$b_1 = 12.46 \quad (11.67, 13.26)$$

$$c_1 = 2.901 \quad (2.453, 3.349)$$

$$a_2 = 0.08764 \quad (0.01377, 0.1615)$$

$$b2 = 9.838 \quad (9.28, 10.4)$$

$$c2 = 2.171 \quad (1.761, 2.581)$$

El gráfico de este modelo gaussiano se ve en la Figura 84, el voltaje máximo corresponde a 0.2 a las 12am.

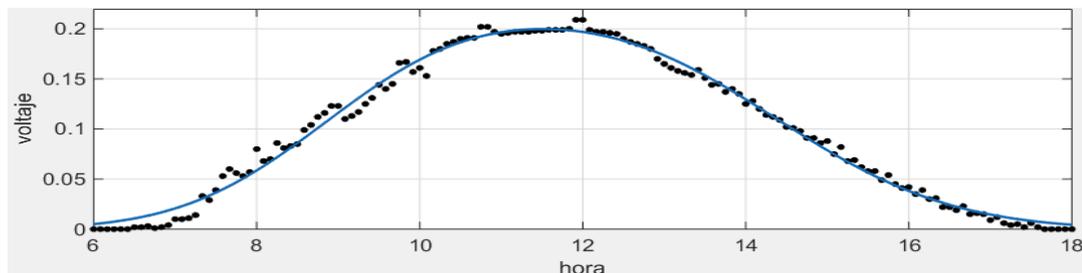


Figura 83 Grafico del modelo general Gauss 2.

La curva continua en color azul es el gráfico dado por Matlab de la ecuación 3.67 y la curva en puntos de color negro, es la curva de la Figura 84 que se está modelando. La ecuación es una contribución original de esta investigación.

4.12 Corrección del modelo del voltaje del sensor mediante la inclusión de la nubosidad.

Para incluir la nubosidad en el modelo matemático del voltaje del sensor, se toma como base la ecuación 2.1 indicada antes y que se vuelve a escribir a continuación.

$$UVI = UVI0 * CMF * (1 + 0.08 * \Delta h) \quad (2.1) \quad (4.19)$$

A pesar de que la ecuación 3.68 hace referencia directa y exclusiva al IUV (UVI en inglés), se va a utilizar esta para corregir el modelo matemático del voltaje, esto es factible debido a que el IUV calculado es directamente proporcional al voltaje generado por el sensor como se demostró en la sección 1.5.2, la relación está dada por la ecuación 3.70.

$$IUV = 78.1 \text{ v.}$$

En donde el voltaje está en mili voltios (mv) y v es la ecuación 2.2 ($v(t)$).
Se dividen los 2 lados de la ecuación 2.1 por la constante 78.1, y se tiene la ecuación 3.69.

$$UVI / 78.1 = UVI0 / 78.1 * CMF * (1 + 0.08 * \Delta h) \quad (4.20)$$

$UVI / 78.1$, es el voltaje (v) de la ecuación 2.2, y $UVI0 / 78.1$ sería el voltaje medido a nivel del mar. La ecuación 2.4 se transforma en la ecuación 2.5.

$$v = v_0 * CMF * (1 + 0.08 * \Delta h) \quad (4.21)$$

En donde: v es el voltaje pronosticado o calculado y, v_0 el voltaje medido a nivel del mar.

La ecuación 3.70 corresponde a la corrección del voltaje v_0 , con la nubosidad y la altura.

CFM= corrección por nubosidad.

$(1 + 0.08 * \Delta h)$ = corrección por altura.

El voltaje v que genera el sensor incluye entonces el efecto de la altura, es decir el término entre paréntesis de la ecuación 2.5.

Por lo tanto la ecuación 2.5, se puede escribir como:

$$v = v_1 * CFM$$

Donde v_1 , es el voltaje medido y que incluye la altura sobre el nivel del mar a la que es medido ese voltaje.

El modelo matemático del voltaje de la ecuación 2.6, para el sensor UV-01 queda de la siguiente manera.

$$v = \left(a_1 * e^{-\left(\frac{t-b_1}{c_1}\right)^2} + a_2 * e^{-\left(\frac{t-b_2}{c_2}\right)^2} \right) * CFM \quad 6 \leq t \leq 18 \quad (4.22)$$

Puesto que $CFM = 1$, para las octavas 0-2 y nubosidad alta, media y baja, y octavas 3-4 y nubosidad alta y media y, para las octavas 5-6 con nubosidad alta, el voltaje de predicción es: $v = v_1$

La ecuación 3.71 es el aporte original de esta investigación. Estas ecuaciones predicen el voltaje generado por la radiación UV-A y UV-B para cielo claro y nublado y, partir de estas ecuaciones se puede predecir el IUV. Las ecuaciones 3.71 es local sin embargo es válida para *cualquier lugar* en donde se hagan las mediciones del voltaje.

Se calculan los IUV para las ciudades de Quito y Cuenca con la finalidad de verificar el error que se tendría al utilizar la ecuación 2.1 y por ende el error que se podría producir al evaluar el IUV.

Para tener una idea de la exactitud de la ecuación en el pronóstico del IUV se calcula el IUV para Quito y Cuenca los resultados son:

Quito:

Altura= 2.8 Kms.

CMF= 1 (cielo claro).

UVI_o= 12.

$$IUV (\text{Quito}) = 12 * 1 * (1 + 0.08 * 2.8) = 14.69. \quad (4.23)$$

El IUV se expresa como un entero redondeado, por lo tanto el IUV calculado es 15. El valor medido de acuerdo al historial de la Agencia Espacial Civil Ecuatoriana (EXA) es 16, para los primeros días del enero.

$$IUV (\text{Cuenca}) = 12 * 1 * (1 + 0.08 * 2.560) = 14.46. \quad (4.24)$$

El IUV para Cuenca que se encuentra a 2560 metros sobre el nivel del mar, es 15. El valor medido de acuerdo al historial de la Agencia Espacial Civil Ecuatoriana (EXA) es 16, para los primeros días de enero. Los resultados del IUV para las 2 ciudades, nos da una buena idea de que la ecuación 2.1 da un error de una unidad en el cálculo del IUV.

Como se indicó antes esto no es raro debido a la incertidumbre de los parámetros atmosféricos. La predicción de series de tiempo es un importante problema de investigación debido a sus implicaciones en ingeniería.

Siendo importante el desarrollo de nuevos modelos y su comparación con aproximaciones previas en términos de la precisión del pronóstico de la radiación solar.

Recientemente, las máquinas de vectores de soporte (SVM) han sido usadas para la predicción de series de tiempo, pero las experiencias han sido alentadoras en términos de precisión de los modelos. El método kernel propuesto es usado para estimar varias SVM con el fin de pronosticar series futuras.

CAPÍTULO V

VALIDACIÓN DEL MODELO SOFTWARE

5. Introducción

La variable observada en el presente trabajo es el índice de radiación Solar, a cielo despejado, los datos fueron obtenidos los 2 primeros años (marzo 2011-marzo 2013) a través de un sistema de adquisición de datos comercial con el sensor UVI-01 Ultraviolet rays detector, y los datos de los dos años subsiguientes (abril 2013-abril2015) fueron obtenidos a través de un sistema de adquisición de datos diseñado y creado por la Autora de la Presente Tesis,

Considerando que la adecuación de los modelos no se evalúa con los datos utilizados en su ajuste o entrenamiento, sino con otro conjunto de valores (test) de la variable 1, toen cuenta que no se ha realizado ningún estudio ni medición en Riobamba-Ecuador antes.

Los datos son almacenados en sistema de Adquisición de datos y enviada en formato texto a la web en tiempo real [www.http://nfosolarespoch.com](http://nfosolarespoch.com) en base a mediciones in-situ, adaptadas al Sistema NWS-EPA, Normas de la OMS y comparadas con EXA (Agencia Espacial Civil Ecuatoriana), que aunque no posee mediciones en Riobamba, se tomó como referencia las mediciones de Guayaquil, con el parámetro importante de que las mediciones de EXA son tomadas desde el Satélite, y que pueden ser observadas desde el monitor Nacional de radiación ultravioleta.

5.1 Criterios de evaluación del modelo

La adecuación de los modelos no se evalúa con los datos utilizados en su ajuste o entrenamiento, sino con el conjunto de variables explicativas voltaje obtenido y minutos Esta prueba permite la comparación de los resultados (valores de predicción) entre diferentes modelos y de esta forma seleccionar el modelo que arroje los menores errores de predicción.

Para la comparación general entre los modelos se ha usado la raíz del error cuadrático medio, RMSE, ya sea en su valor real, ecuación (5.1), o en su valor normalizado a la potencia instalada, ecuación (5.2). En estas ecuaciones N representa el número total de datos y k el horizonte de predicción.

Una vez implementado el Hardware y el Software del sistema electrónico de predicción del IUV en la ESPOCH, se prueba el sistema para validarlo, realizar el procedimiento de corrección y puesta a punto del sistema, para lo cual se hace necesaria:

- La validación del sistema electrónico en cuanto a la exactitud de los resultados generados por Sistema Electrónico de Adquisición de Datos versus los valores versus los valores de predicción.
- La validación del Software del sistema Electrónico, como la interfaz de usuario.
- La validación de la predicción del índice IUV en la página web Institucional.

5.2 Validación de la exactitud de los resultados que entrega el sistema electrónico en base a los cálculos del modelo

La prueba consiste en comparar los resultados que generan Matlab y el arduino al evaluar la ecuación:

$$v = (a1 * e^{-((\frac{t-b1}{c1})^2)} + a2 * e^{-((\frac{t-b2}{c2})^2)}) * CFM \quad (5.1)$$

Se eligen 2 valores del tiempo el uno a las 9 de la mañana por ser el mínimo valor y el otro a la hora del medio día solar que son las 12h00.

Para las 9 de la mañana, $t=9$ y, para el medio día solar, $t=12$, Arduino tienen una exactitud de 2 cifras decimales, los resultados se ven en la tabla 5.1, de donde se puede afirmar que el Arduino genera los mismos resultados que Matlab para dos cifras decimales, por lo tanto el cálculo de la ecuación que realiza el Arduino es exacto.

Los programas tanto en Matlab como para el Arduino que se utilizaron para llenar la tabla 5.1.

Calculo de la ecuación: $v = (a1 * e^{-((\frac{t-b1}{c1})^2)} + a2 * e^{-((\frac{t-b2}{c2})^2)}) * CFM$ (5.2)

Tabla 22
Resultados de las pruebas de simulación.

CFM	Resultado del SAD		Resultado Matlab:		Exactitud para 2 cifras significativas
	t=9	t=12	t=9	t=12	
1	0.01	0.20	0.0050	0.1999	100%
0.9	0.00	0.18	0.0045	0.1712	100%
0.8	0.00	0.19	0.0040	0.1575	100%
0.5	0.00	0.10	0.0025	0.0985	100%
0.4	0.00	0.08	0.0020	0.0788	100%
0.2	0.00	0.04	0.0010	0.0394	100%

5.3 Validación del software del sistema electrónico.

La validación de la interfaz de usuario se realiza mediante el programa que es básicamente el mismo que está indicado en el capítulo IV en el apartado 4.3.2

En la figura 5.1 se observa la ejecución del programa en el computador, y los resultados se muestran en el monitor a través del puerto serial COM 7.

```

// PROGRAMA DESARROLLADO PARA EL ARDUINO MEGA 2560, SE USA UN TECLADO GENERAL
// ESTE PROGRAMA EVALUA LAS ECUACIONES MEDIANTE LAS CUALES SE PREDICE EL UV
// TAMTO MEDIANTE EL MONITOR SERIE DE ESTE SKETCH SOBRE UN PC, ASI COMO YA
// LA SELECCION DEL TIPO DE HUMEDAD SE REALIZA MEDIANTE EL TECLADO.

#include <EEPROM.h> //se añade la libreria EEPROM, para leer y escribir
#include <LiquidCrystal.h> // se añade la libreria LiquidCrystal, que permite
LiquidCrystal lcd(12, 11, 5, 4, 3, 2); // se define e inicializa los pines
// que el LCD utilizara del arduino.

#include <Keypad.h> //añadimos la libreria keypad, para poder utilizar el teclado
const byte FILAS = 4; // se define el número de filas del teclado
const byte COLUMNAS = 3; //se define el número de columnas del teclado
char keys[FILAS][COLUMNAS] = { // se crea una matriz con los caracteres de
  {'1','2','3'},
  {'4','5','6'},
  {'7','8','9'},
  {'*','0','#'}
};

```

COM7

```

La nubosidad elegida es:
Octava=0-2_A, M, B
La predicción del UV para la ESPOCH ES:
Hora: 6.00__suv: 2
Hora: 6.05__suv: 2
Hora: 6.10__suv: 2
Hora: 6.15__suv: 2
Hora: 6.20__suv: 2
Hora: 6.25__suv: 2
Hora: 6.30__suv: 2
Hora: 6.35__suv: 2
Hora: 6.40__suv: 3
Hora: 6.45__suv: 3
Hora: 6.50__suv: 3
Hora: 6.55__suv: 3

```

Figura 84 El programa de simulación corriendo sobre un PC.

En la figura 85 se puede observar las capturas de pantalla con el sistema funcionando bajo las pruebas a las que es sometido el sistema, y del mismo modo en la tabla 5.2 se muestra la respuesta del sistema a los estímulos

programados, como se puede ver la respuesta del sistema es la adecuada, por lo tanto el software responde de manera adecuada.

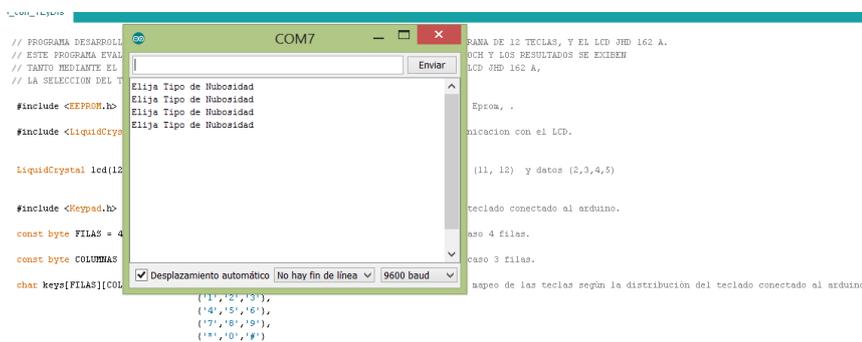


Figura 85 Captura de pantalla con el mensaje del sistema al usuario
 Fuente: (Baldeon W, Mora V, 2014)

Tabla 23
Respuesta y resultado de la validación del software del sistema a pruebas de teclado en el cálculo del IUUV.

Mensaje al usuario	Tecla activada	Respuesta del software al mensaje del usuario	Respuesta del sistema electrónico.
Sin Nubosidad	Ninguna	Espera por una tecla	La esperada
Sin Nubosidad	9	Octavas=7-8__A	Calcula y Muestra IUUV La esperada
Sin Nubosidad	8	Octavas=3-4-_B	Calcula y Muestra IUUV La esperada
Sin Nubosidad	7	Octavas=7-8__B	Calcula y Muestra IUUV La esperada
Sin Nubosidad	9	Octavas=3-4__B	Calcula y Muestra IUUV La esperada
Sin Nubosidad	5	Octavas=5-9__B	Calcula y Muestra IUUV La esperada
Sin Nubosidad	4	Octavas=7-8__B	Calcula y Muestra IUUV La esperada
Sin Nubosidad	3	Octavas=5-9__A	Calcula y Muestra IUUV La esperada
Sin Nubosidad	2	Octavas=3-4__M	Calcula y Muestra IUUV La esperada
Sin Nubosidad	1	Octavas=0-2__A, M, B	Calcula y Muestra IUUV La esperada
Sin Nubosidad	0	Octavas=3-4__A	Calcula y Muestra IUUV La esperada
Sin Nubosidad	*	Octavas=7-8__M	Calcula y Muestra IUUV La esperada
Sin Nubosidad	#	Octavas=7-8__lluvia	Calcula y Muestra IUUV La esperada

5.4 Validación de la predicción del iuv en la epoch

Como se indica en el apartado 5.2 y 5.2.1 La validación se realiza comparando para una misma hora, el valor de IUVM medido con el valor IUVP pronosticado, y este calculado en términos porcentaje nos entrega el error producido por el sistema electrónico.

El cálculo del error se realiza mediante la siguiente ecuación:

$$\%ERROR = (IUVM - IUVP)/100 \quad (5.1)$$

Dónde:

IUVM= IUVM medido.

IUVP= IUVP pronosticado.

Como se indicó en el capítulo II el error que generan los modelos de predicción de IUVM, es en general del orden de una unidad de IUVM, por tal razón según (Lehmann, E. L.; Casella, George, 1998) "El error cuadrático de un estimador mide el promedio de los errores al cuadrado, es decir, la diferencia entre el estimador y lo que se estima. El error es una función de riesgo, correspondiente al valor esperado de la pérdida del error al cuadrado o pérdida cuadrática. La diferencia se produce debido a la aleatoriedad o porque el estimador no tiene en cuenta la información que podría producir una estimación más precisa".

Por lo tanto el error esperado de este sistema electrónico de predicción es de una unidad de IUVM.

5.5 Error de predicción

La ecuación (4.2) detallada en el apartado 4.3 muestra el esquema de medición, cualquier cambio en UVIO es corregido por CMF debido a la nubosidad, y que constituyen valor de entrada en el modelo, al final se traduce en un error de medición.

$$UVI = UVIO * CMF * (1 + 0.08 * \Delta h) \quad (5.2)$$

Es evidente además que si los cambios en UVI son pequeños el Índice acoplado al estándar de la OMS no es relevante, es así que por ejemplo en la figura 86 se puede observar que el valores más altos de VNI ocurre a las 11h30 y cae a las 6y30 pm, los valores expresan el VNI medido y constante, lo que permite definir un error mínimo de predicción

2015/03/06

Hora	VNI	Hora	VNI
6	1	12,5	4
6,5	1	13	4
7	1	13,5	3
7,5	1	14	3
8	1	14,5	4
8,5	1	15	4
9	1	15,5	4
9,5	1	16	4
10	2	16,5	4
10,5	3	17	3
11	3	17,5	2
11,5	4	18	1
12	4	18,5	1

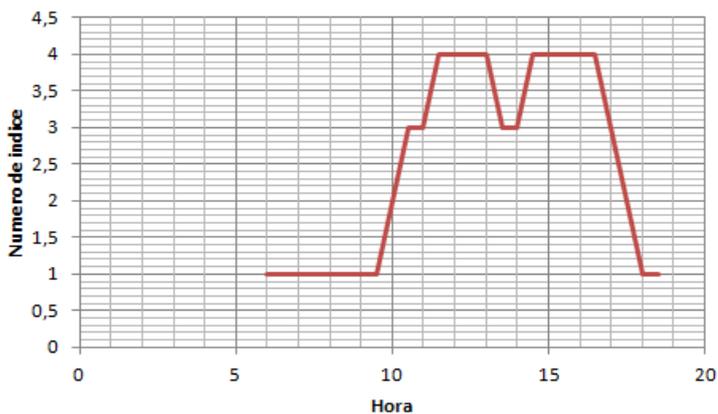


Figura 86 VNI medido en función de la hora.

Fuente: (Baldeón W, Mora V, 2014)

En la figura 87 se puede observar una referencia de los valores de predicción para el día 6 de marzo del 2015 y se puede notar la celda resaltada en rojo equivalente medido en tiempo real a las 10 am de la mañana con un índice 2VNI y el valor pronosticado de 2VNI, asimismo el valor medido a las 17 horas con el índice 3VNI y el valor de predicción equivalente a 3

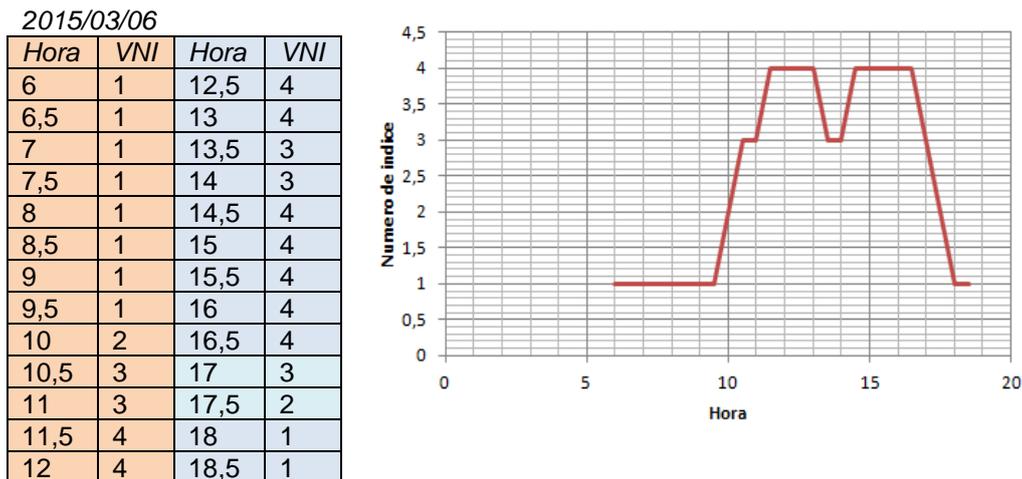


Figura 87 VNI pronosticado en función de la hora.

Fuente: (Baldeón W, Mora V, 2014)



Figura 88 Representación de los valores máximos VNI

Fuente: (Baldeon W, Mora V, 2014)

Del análisis de los datos representados linealmente en la Figura 89 se concluye que durante la primera semana de febrero del año en curso los valores máximos de VNI se producen entre las 11h30 y las 14h00, por lo tanto estas son las horas más críticas para la salud, sin embargo el rango puede variar y en su modo general estaría entre las 10h30 y las 15h30 para días sin nubosidad.

El valor pronosticado permite guardar un archivo de las predicciones al igual que los valores medidos y enviados a la web en base al modelo estimado con el factor de corrección y el ajuste de valores en cada caso sin

nubosidad, en la que El pronóstico promedio proporciona el error típico de las predicciones del valor deseado.

Con las variables Hora y VNI se genera el archivo mediciónTR.txt y que han sido analizadas en el apartado 4 y el archivo prediccionTR se estiman los valores pronosticados, se aplica el modelo de regresión lineal simple que explica el comportamiento del VNI en función de la hora y se realice la estimación en las horas de mediciones definidas (6:00am a 6:30 pm).

Debido a que los días con o sin nubosidad en un año calendario a partir de las 6:00pm el índice IUUV es de 1VNI constante. El modelo permite generar predicciones para el valor esperado a 12 horas sin tomar en cuenta las horas de 6:00 pm a 6:00 am del siguiente día debido a que los valores son constantes y el IUUV es igual a 1 siempre.

Para obtener el intervalo de confianza de los pronósticos y así lograr contrastar los valores se condiciona el valor $X_0=Y_0$ y se calcula el error de la predicción, el cual como ya se explicó en el apartado 5.2 dependerá del valor de la predicción. del valor esperado de Y para $X=X_0$, se tomará en cuenta la siguiente modelo:

$$v = (a_1 * e^{-((\frac{t-b_1}{c_1})^2)} + a_2 * e^{-((\frac{t-b_2}{c_2})^2)}) * CFM \quad 6 \leq t \leq 18$$

La hipótesis de la presente investigación indicaba que “Si se crea un modelo software basado en técnicas de inteligencia artificial capaz de predecir series temporales no lineales. Entonces es posible obtener una predicción de la radiación solar en 3 días posteriores a la medición en la ciudad de Riobamba El resumen del modelo es el siguiente

Tabla 24 Descripción del cálculo del error en la predicción del IUUV

Modelo	R	R ²	R corregido	Error de la estimación
Mediciones volatje/hora				

	De (6:00am a			
	6:00pm) 8 meses			
20160	43680	1907942400	1907898960	0.0012

De las variables predictoras, en el cuadro resumen del modelo se observa que: $r=43680$ $R^2=1907942400$ y el error de estimación es 0.012 del modelo estimado por lo que el modelo proporciona un buen ajuste.

Cabe destacar que las mediciones fueron realizadas a cielo claro tomando en cuenta el tiempo de medición de 6am a 6pm, debido a que el siguiente periodo de tiempo es decir de 6pm a 6am la radiación solar es constante y equivale al nivel 1 según los rangos de la OMS indicados en el apartado 4.

Es importante destacar que los valores medidos según la Escala de la OMS poseen un rango definido de mediciones dentro de cada nivel, es así que las mediciones en el rango 10+ equivalen a mediciones de potencial de voltaje entre 850 y 1024 Watts/s por lo que la variación de los valores corresponde a la escala 1 de la OMS, permitiendo de esta manera tener más exactitud en la predicción del nivel en la escala OMS.

6 Conclusiones

1. El desarrollo de la herramienta web servirá como medio de difusión del índice de radiación solar en Riobamba, y cuya utilización permitirá una correcta valoración en las personas con cáncer de piel.
2. El sistema para adquisición de datos es muy sensible a cielo claro
3. La técnica apropiada para predicción e variables metereológicas son las Redes neuronales Backpropagation y los SVM en series de tiempo no lineales por la excelente precisión en los modelos.
4. El método propuesto fue usado para estimar varias SVM con el fin de pronosticar series futuras, estos resultados son comparados con las estadísticas reportadas de las mediciones de los 3 años anteriores al actual.
5. Los puntos de referencia para la medición, deben estar ubicados al menos en el norte y en el sur en un espacio abierto donde los edificios y la proyección de la sombra no afecte la medición.
6. En la predicción de variables metereológicas las características de las redes neuronales artificiales son muy apropiadas debido a que se dispone de un conjunto básico datos de entrada para encontrar patrones de predicción.
7. La técnica apropiada para predicción e variables metereológicas son las Redes neuronales Backpropagation y los SVM en series de tiempo no lineales y su excelente precisión en los modelos.
8. El método propuesto fue usado para estimar varias SVM con el fin de pronosticar series futuras, estos resultados son comparados con las estadísticas reportadas de las mediciones de los 3 años anteriores al actual.
9. Las técnicas de inteligencia artificial ANN Backpropagation y los SVM, son específicas para predicción de variables metereológicas ya que comparten algunas ventajas como Aprendizaje, Auto organización, Tolerancia a fallos, y Flexibilidad permitiendo mayor exactitud en las predicciones de radiación solar, la exactitud y el menor valor de error son característicos en las aplicaciones de estas técnicas IA.

10. Los valores intermedios entre una medición y otra no posee mayor incidencia en el valor asignado por la escala OMS, más sin embargo los valores superiores e inferiores de la escala definen un cambio de valor de incidencia solar.
11. La incertidumbre de los parámetros atmosféricos en la predicción de series de tiempo es un importante problema de investigación debido a sus implicaciones en ingeniería, Siendo importante el desarrollo de nuevos modelos y su comparación con aproximaciones previas en términos de la precisión del pronóstico de la radiación solar.
12. El pronóstico promedio proporciona el error típico de las predicciones del valor deseado, El modelo permite generar predicciones para el valor esperado a 3 días sin tomar en cuenta las horas de 6:00 pm a 6:00 am del siguiente día debido a que los valores son constantes y el IUV es igual a 1 siempre.
13. Los modelos kernel en las máquinas de vectores de soporte (SVM) son usadas para la predicción de series de tiempo, pero las experiencias han sido alentadoras en términos de precisión de los modelos.
14. Los resultados obtenidos de la presente investigación pretenden servir de aporte local, y puede ser utilizada como línea base por el Ministerio de Salud en la prevención de cáncer de piel en la población de Riobamba, el Ministerio de Electricidad y Energía Renovable, y posteriormente pudriera extrapolarse a otras regiones del País.
15. La modelación del índice de radiación es complicado, sin embargo el error de predicción es 0
16. Se desarrolló un modelo Software local que incluye la altura nivel del mar a cielo claro sin nubes.
17. El modelo puede ser usado en cualquier parte del Ecuador.
18. El modelo es adaptable a cualquier sensor dentro de la banda UVA y UVB
19. El sistema de adquisición de datos simplificó el diseño electrónico gracias a la ayuda de Arduino
20. El modelo Software provee información en tiempo real de la irradiación solar en la ciudad de Riobamba, a través del servidor de la Escuela Superior Politécnica de Chimborazo.

7 Recomendaciones

1. Se puede crear una red de sensores inteligentes basados en el voltaje de un sensor, con el objetivo de enviar en forma inalámbrica la información desde diferentes puntos de la ciudad y del país.
2. El sistema puede conectarse a una red nacional con el fin de proveer información en tiempo real vía móvil.
3. Con la ayuda de Matlab se puede analizar y procesar cada una de las señales para mejorar el factor de corrección.
4. El monitoreo de la radiación solar en Riobamba debe ser constante debido a que el IUV va cada vez en aumento según la escala OMS el valor más alto en Riobamba es de 17. Es decir un riesgo notable en la salud.
5. Se debe investigar los factores como nubosidad, contaminación, que pueden incidir en la medición y en el modelo de la radiación IUV reflejada.
6. Se debe instalar el dispositivo de medición en parques, parques, estadios lugares de gran concurrencia y en donde las personas hagan deporte al aire libre
7. Es necesario crear un prototipo electrónico que simule la piel según el genotipo para determinar bajo experimentación el tiempo de exposición solar promedio.
8. Se debe crear una App del modelo de predicción.

8. REFERENCIAS BIBLIOGRÁFICAS

- [1] G. Zhang, B. Patuwo and M. Hu. "Forecasting with artificial neural networks: the state of the art". *International Journal of Forecasting*. Vol. 14, pp. 35-62. 1998
- [2] González-Rodríguez, A. F. (2014). *Modelo para la predicción de la radiación solar a partir de redes neuronales artificiales*, Escuela de Ingeniería de Antioquia.
- [3] H. Tong. "Non-linear time series a dynamical system approach". Claredon Press Oxford. 1990
- [4] EPA, environmental protection agency. a guide to the uv, index
- [5] Juez García, R., J. H. Hechavarría Miyares, et al. (2001). "Cáncer de piel y ocupación." *Revista Cubana de Medicina*
- [6] D. Renné, "Semi-annual status report: November 2009. task 36: Solar resource knowledge management," tech. rep., IEA. IEA SHC Task 36 Solar Resource Knowledge Management., 2009.
- [7] A. Sfetsos y A. Coonick, "Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques," *Solar Energy*, vol. 68, no. 2, pp. 169–178, 2000.
- [8] Y. Kemmoku, S. Orita, S. Nakagawa, y T. Sakakibara, "Daily insolation Forecasting using a multi-stage neural network," *Solar Energy*, vol. 66, no. 3, pp. 193 – 199, 1999.
- [9] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
- [10] E. Lorenz, J. Hurka, D. Heinemann, y H. G. Beyer, "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, no. 1, pp. 2–10, 2009.
- [11] M. R. Arahál, M. Berenguel Soria, y F. Rodríguez Díaz, *Técnicas de predicción con aplicaciones en Ingeniería*. Sevilla: Universidad de Sevilla. Servicio de publicaciones, 2006
- [12] *Inteligencia Artificial Un Enfoque Moderno*, 2da Edición – Russell y Norvig
- [13] Baldeon Wilson, Mora Verónica "Índice de radiación uv en Riobamba y su peligro en la salud" *Editorial ESPOCH*, ISSN 24532
- [14] Laboratorio Nacional de Investigación de Energía en su artículo publicado en la revista Scielo Laboratorio Nacional de Investigación
- [15] Zhang, Peter: "Neural Networks in Environmental Forecasting". 1ª edición, Idea Group Publishing, E.U.A. 2004
- [16] José R. Hilera y Victor J. Martínez, "Redes Neuronales Artificiales", Alfaomega-Rama, 2000.
- [17] VÁZQUEZ E., CHACÓN O. L. y ALTUVE H. J. (1999). *Aplicación del análisis de componente principal y redes neuronales en la identificación de osciladores de potencia*. *Ingenierías*, 2 (5):3-12.
- [18] DORFFNER, G. (1996). *Neural Networks for Time Series Processing*. *Neural Network World*, 4: 447-468.

- [19] Labajo, A.L.; Martín, Q. y Labajo, J.L. *Departamento de Física General y de la Atmósfera. Universidad de Salamanca.*
- [20] Zhang, Peter: "Neural Networks in Environmental Forecasting". 1ª edición, Idea Group Publishing, E.U.A. 2004
- S. Y. Kung, "Digital Neural Networks", Prentice Hall, 1993.
- [20] J. Cao y X. Lin, "Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 8, pp. 1255–1263, 2008.
- [21] E. M. Crispim, P. M. Ferreira, y A. E. Ruano, "Prediction of the solar radiation evolution using computational intelligence techniques and cloudiness indices," *International Journal of Innovative Computing, Information and Control*, vol. 2, p. 2, 2008.
- [22] C. Paoli, C. Voyant, M. Muselli, y M.-L. Nivet, "Forecasting of preprocessed daily solar radiation time series using neural networks," *Solar Energy*, vol. 84, no. 12, pp. 2146–2160, 2010.
- [23] Baldeon, W. Mora, V "Un sistema Electronico para la predicción de Radiación UV" *Perspectivas Epoch*, 3ra Ed, 2015.
- [24] Z. Ramedani, M. Omid, A. Keyhani, S. Shamshirband, y B.

ANEXOS

ANEXO-A:

(Programa referencial que evalúa las ecuaciones mediante las cuales se predice el IUUV en la Espoch)

```
// PROGRAMA DESARROLLADO PARA EL ARDUINO MEGA 2560, SE USA  
UN TECLADO GENERICO DE MEMBRANA DE 12 TECLAS, Y EL LCD JHD 162  
A.
```

```
// ESTE PROGRAMA EVALUA LAS ECUACIONES MEDIANTE LAS CUALES  
SE PREDICE EL IUUV EN LA ESPOCH Y LOS RESULTADOS SE EXIBEN
```

```
// TANTO MEDIANTE EL MONITOR SERIE DE ESTE SKETCH SOBRE UN  
PC, ASI COMO YAMBIEN EN EL LCD JHD 162 A,
```

```
// LA SELECCION DEL TIPO DE NUBOSIDAD SE REALIZA MEDIANTE EL  
TECLADO.
```

```
#include <EEPROM.h> //se añade la librería EEPROM, para leer y  
escribir en la Eprom, .
```

```
#include <LiquidCrystal.h> // se añade la librería LiquidCrystal, que permite la  
comunicación con el LCD.
```

```
LiquidCrystal lcd(12, 11, 5, 4, 3, 2); // se define e inicializa los pines de control  
(11, 12) y datos (2,3,4,5)
```

```
// que el LCD utilizara del arduino.
```

```
#include <Keypad.h> //añadimos la librería keypad, para poder utilizar el  
teclado conectado al arduino.
```

```
const byte FILAS = 4; // se define el número de filas del teclado: en este  
caso 4 filas.
```

```
const byte COLUMNAS = 3; //se define el número de columnas del teclado:  
en este caso 3 filas.
```

```
char keys[FILAS][COLUMNAS] = { // se crea una matriz con los caracteres del  
teclado, mapeo de las teclas según la distribución del teclado conectado al arduino
```

```
{'1','2','3'},
```

```
{'4','5','6'},
```

```
{'7','8','9'},
```

```
{'*','0','#'}
```

```
};
```

```
byte rowPins[FILAS] = {34,32,30,28}; //se F1 8, F2 7  
indica a que pines digitales del arduino se conectas las filas del teclado.
```

```
byte colPins[COLUMNAS] = {26,24,22}; //se indica a C1 4,  
C2 3que pines digitales del arduino se conectan las columnas del teclado.
```

```
Keypad keypad = Keypad( makeKeymap(keys), rowPins, colPins, FILAS,  
COLUMNAS ); //se crea un objeto de la libreria keypad para usar sus funciones
```

```
float v; // v, es el modelo matematico del voltaje generado por el  
sensor IUUV-1.
```

```
float x=6; // x, es la hora del dia sin escalamineto.  
float expo; // expo, es la ecuacion del exponente del modelo  
matemático del voltaje del sensor.
```

```
x=x+0.0833; //x, es el tiempo y el valor 0.0833  
es el equivalente a 5 minutos del tiempo de un reloj.
```

```
hora=k+in; // genera las horas.  
in=in+0.05; // Genera los minutos de las horas  
(hora), cuenta cada 5 minutos.
```

```
if (in > 0.900) { in = 0; k = k+1;} //  
inicializa la cuenta de los minutos e incrementa la hora en una unidad.
```

```
Serial.print("Hora: "); //imprime el mensale  
Hora.
```

```
Serial.print(hora,2); // imprime la hora  
Serial.print("___iuv; // imprime el mensaje iuv.  
Serial.println(iuv); //imprime el valor de iuv.  
delay(3000);
```

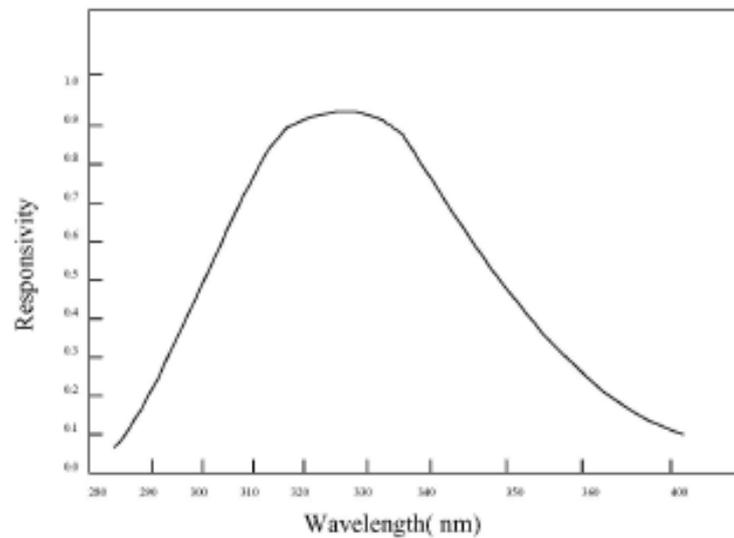
```
if (hora >17.55) { k = 9; in=0; x=9;}//  
inicializa los contadores temporales, horas,x e in.  
} } }
```

ANEXO-B:

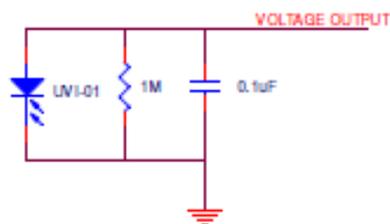
(Datasheet Sensor)

UVI-01 Ultraviolet rays detector

Wave length vs. Responsivity



Application Circuit



REYAX TECHNOLOGY CO.,LTD.

Address : 6F.-3, No.6, Aly. 30, Ln. 210, Wende Rd., Neihu Dist., Taipei City 11476,
Taiwan

Web : <http://www.reyax.com>

Tel : +886-2-2627-2777

Fax : +886-2-2627-2781

E-mail : sales@reyax.com

ANEXO-C:

Pantalla principal del modelo Software para la predicción del IUV.



[INICIO](#) [QUIENES SOMOS](#) [ANTECEDENTES](#) [BASE CIENTÍFICA](#) [ESTADÍSTICAS](#) [GALERÍA](#) [CRÉDITOS](#) [CONTACTO](#)

MODELO SOFTWARE PARA EL MONITOREO Y DIFUSIÓN DE LOS ÍNDICES DE RADIACIÓN SOLAR EN LA ESPOCH

NÚMERO DE ÍNDICE UV EQUIVALENTE SEGÚN LA NWS - EPA	DESCRIPCIÓN	CATEGORÍA DE EXPOSICIÓN (OMS)	RECOMENDACIONES DE LA OMS PARA PROTEGERSE DE LA RADIACIÓN UV	EXPOSICIÓN GRÁFICA
1	No hay peligro para la persona promedio.	BAJA	No necesita protección. Puede mantenerse en el exterior sin riesgo.	

FECHA/HORA DE MONITOREO: 2015-06-17 20:10:27

INDICE UV	0 - 2	3 - 4	5 - 6	7 - 9	10 +
CATEGORÍA DE EXPOSICIÓN	BAJA	MODERADA	ALTA	MUY ALTA	EXTREMA

LOCALIZACIÓN



Posición del sensor de radiación UV

CONTADOR DE VISITAS

0000934

Fuente: (Mora. V, 2015))

ANEXO-D:

Programa de simulación es el que se indica a continuación.

```
// PROGRAMA DESARROLLADO PARA PROBAR EL SOFTWARE DEL ARDUINO
MEGA 2590, SE USA UN TECLADO GENERICO DE MEMBRANA DE 12 TECLAS.
// ESTE PROGRAMA EVALUA LAS ECUACIONES MEDIANTE LAS CUALES SE
PREDICE EL IUV EN LA ESPOCH Y LOS RESULTADOS SE EXIBEN
// MEDIANTE EL MONITOR SERIE DE ESTE SKETCH SOBRE UN PC.
// LA SELECCION DEL TIPO DE NUBOSIDAD SE REALIZA MEDIANTE EL TECLADO.
p

#include <Keypad.h>          //añadimos la librería keypad, para poder utilizar el teclado
conectado al arduino.

const byte FILAS = 4;      // se define el número de filas del teclado: en este caso 4
filas.

const byte COLUMNAS = 3;   //se define el número de columnas del teclado: en este
caso 3 filas.

char keys[FILAS][COLUMNAS] = { // se crea una matriz con los caracteres del teclado,
mapeo de las teclas según la distribución del teclado conectado al arduino
    {'1','2','3'},
    {'4','5','9'},
    {'7','8','9'},
    {'*','0','#'}
};

byte rowPins[FILAS] = {34,32,30,28}; //se F1 8, F2 7 indica a que
pines digitales del arduino se conectas las filas del teclado.
byte colPins[COLUMNAS] = {29,24,22}; //se indica a C1 4, C2 3que
las 9h00 y hasta las 18h00.
float hora=9; //hora, guarda las horas y minutos, cuenta las horas cada 5
minutos.
float in=0; // Genera los minutos de las horas, cuenta los minutos , cada 5
minutos.

x=x+0.0833; //x, es el tiempo y el valor 0.0833 es el
equivalente a 5 minutos del tiempo de un reloj.
hora=k+in; // genera las horas.
in=in+0.05; // Genera los minutos de las horas (hora),
cuenta cada 5 minutos.

if (in > 0.900) { in = 0; k = k+1;} // inicializa la
cuenta de los minutos e incrementa la hora en una unidad.

Serial.print("Hora: "); //imprime el mensale Hora.
Serial.print(hora,2); // imprime la hora
Serial.print("___iuv;// imprime el mensaje iuv.
Serial.println(iuv); //imprime el valor de iuv.
delay(3000);

if (hora >17.55) { k = 9; in=0; x=9;}// inicializa los
contadores temporales, horas,x e in.
```



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y TRANSFERENCIA DE TECNOLOGÍA

MAESTRÍA EN INGENIERÍA DE SOFTWARE CUARTA PROMOCIÓN

Se certifica que el presente trabajo fue realizado por la Sra. Ing. Verónica Elizabeth Mora Chunllo. En la ciudad de Latacunga a los 7 días del mes de marzo de 2016

Aprobado por:

Dr. Luis Enrique Sánchez
DIRECTOR DEL PROYECTO

Ing. Javier Montaluisa MSc.
COORDINADOR DE LA MAESTRÍA
EN INGENIERÍA DE SOFTWARE

Ing. Javier Montaluisa MSc.
OPONENTE

Dr. Rodrigo Vaca
SECRETARIO ACADÉMICO