



DETERMINACIÓN DEL RECURSO SOLAR SOBRE FACHADAS VERTICALES EN CASTILLA Y LEÓN MEDIANTE MÉTODOS DE INTERPOLACIÓN

Memoria de Tesis Doctoral presentada para optar al título de Doctor por la Universidad de Burgos por Carmen Rodríguez Amigo, siendo las directoras de la misma la Dra. D^a Montserrat Díez Mediavilla y la Dra. D^a Cristina Alonso Tristán.

Vº. Bº. de las Directoras

Dra. Montserrat Díez Mediavilla

Dra. Cristina Alonso Tristán

Burgos, 30 de Octubre de 2015



UNIVERSIDAD DE BURGOS

ESCUELA POLITÉCNICA SUPERIOR

PROGRAMA DE DOCTORADO: INGENIERÍA TÉRMICA
(RD 778/1998)

MEMORIA DE TESIS DOCTORAL

TÍTULO:

**Determinación del recurso solar sobre fachadas verticales en
Castilla y León mediante métodos de interpolación**

AUTORA: Carmen Rodríguez Amigo

DIRECTORAS:

Dra. Montserrat Díez Mediavilla

Dra. Cristina Alonso Tristán

*¿Tu verdad? No, la Verdad,
y ven conmigo a buscarla.*

La tuya, guárdatela.

LXXXV

Proverbios y Cantares

Antonio Machado

A mi madre.

30 años después.

Agradecimientos.

A mis Directoras de Tesis por el coraje que tuvieron al ofrecirme trabajar con ellas, a Cristina por su empuje y su ánimo, y sobre todo a Montse que, además de enseñarme radiación solar, me ha mostrado como es capaz de dirigir y acompañar en el trabajo con buen humor y una paciencia casi infinita.

De todos los apoyos recibidos quiero destacar el de mis compañeros del grupo SWIFT. En particular agradecer el trabajo de David, sin su ayuda ArcGis hubiera sido mucho más difícil, y a Ángela y a Fernando por su colaboración con los datos. A Maribel, por su callada amistad. A Amaya por su ayuda.

A Luis Mena por permitirme utilizar una foto suya en la portada.

A toda mi familia por su apoyo y ayuda en todo momento, por su resignación los días en los que estaba “ausente” y su paciencia los días en los que mejor lo hubiera estado.

A todos:

MUCHAS GRACIAS.

ÍNDICE

| | |
|---|-----|
| ÍNDICE | I |
| ÍNDICE FIGURAS | III |
| ÍNDICE TABLAS | VI |
| INTRODUCCIÓN..... | 1 |
| I.1 MOTIVACIÓN Y OBJETIVOS | 1 |
| I.2 SITUACIÓN INICIAL..... | 9 |
| I.3 LÍNEA DE EXPOSICIÓN | 12 |
| CAPÍTULO 1. FUNDAMENTOS | 15 |
| 1.1 INTRODUCCIÓN..... | 15 |
| 1.2 MÉTODOS DE INTERPOLACIÓN..... | 15 |
| 1.2.1 Revisión bibliográfica | 17 |
| 1.2.2 Métodos de interpolación..... | 28 |
| 1.2.3 SPLINES..... | 46 |
| 1.2.4 Métodos de interpolación en ArcGis | 61 |
| 1.3 MODELOS DE RADIACIÓN | 79 |
| 1.3.1 Revisión bibliográfica | 80 |
| 1.3.2. Conceptos generales de radiación solar | 86 |
| 1.3.3 Clasificación de los modelos de radiación | 90 |
| 1.3.4 Modelos de radiación solar empleados en este trabajo..... | 93 |
| CAPÍTULO 2. DATOS | 101 |
| 2.1 INTRODUCCIÓN..... | 101 |
| 2.2 DATOS DE LA AEMET..... | 105 |
| 2.2.1 Fuentes de datos de radiación global. | 105 |
| 2.3 DATOS DE LA RED SIAR..... | 106 |
| 2.3.1 Datos de la región | 106 |
| 2.3.2. Datos de las zonas circundantes | 112 |
| 2.4 DATOS DE LA ESCUELA POLITÉCNICA SUPERIOR DE BURGOS | 114 |
| 2.4.1 Descripción de la estación..... | 115 |
| 2.4.2 Calibración y Mantenimiento..... | 124 |
| 2.4.3. Tratamiento de los datos | 125 |
| CAPÍTULO 3. METODOLOGÍA | 129 |

| | |
|---|-----|
| 3.1 INTRODUCCIÓN | 129 |
| 3.2 TRATAMIENTO DE LOS DATOS DE LA RED SIAR | 130 |
| 3.2.1 Datos de la red SIAR | 130 |
| 3.2.2 Datos de la AEMET | 149 |
| 3.3 MÉTODOS DE INTERPOLACIÓN CON ARCGIS | 149 |
| 3.3.1 Análisis previos | 150 |
| 3.4 MODELOS DE RADIACIÓN SOLAR | 162 |
| 3.4.1. Proceso de cálculo | 163 |
| CAPÍTULO 4. ANÁLISIS ESTADÍSTICO DE LOS DATOS | 167 |
| 4.1 INTRODUCCIÓN | 167 |
| 4.2 ESTUDIO DE LOS DATOS DE IRRADIACIÓN | 167 |
| 4.3 IRRADIACIÓN Y COORDENADAS GEOGRÁFICAS DE LAS ESTACIONES | 177 |
| 4.4 ESTUDIO DE LA NORMALIDAD DE LOS DATOS | 188 |
| 4.5 ESTUDIO DE LA VALIDACIÓN DE RESULTADOS | 192 |
| CAPÍTULO 5. RESULTADOS | 167 |
| 5.1 INTRODUCCIÓN | 195 |
| 5.2 ANÁLISIS DE LOS MÉTODOS DE INTERPOLACIÓN | 196 |
| 5.2.1 Control 1 (C1) | 197 |
| 5.2.2 Control 2 (C2) | 199 |
| 5.2.4 Control 4 (C4) | 203 |
| 5.2.5 Control 5 (C5) | 206 |
| 5.2.6 Control AEMET (C6) | 209 |
| 5.2.7 Control AEMET y SIAR (C7) | 212 |
| 5.2.8 Discusión de los resultados | 215 |
| 5.3 ANALISIS DE LOS MODELOS DE RADIACIÓN | 218 |
| 5.3.1 Resultados en la dirección Norte | 219 |
| 5.3.2 Resultados en la dirección Sur | 222 |
| 5.3.3 Resultados en la dirección Este | 225 |
| 5.3.4 Resultados en la dirección Oeste | 227 |
| 5.4 RESUMEN DE LOS RESULTADOS | 229 |
| CONCLUSIONES | 231 |
| BIBLIOGRAFÍA | 237 |

ÍNDICE FIGURAS

| | |
|--|-----|
| Figura I.1. Consumo de energía eléctrica. Elaborado con datos de http://data.worldbank.org/indicador | 2 |
| Figura I.2. Consumo de energía per cápita | 3 |
| Figura I.3. Comparativa de las emisiones de CO ₂ entre España y U.E. según datos de http://datos.bancomundial.org/indicador | 4 |
| Figura I.4. Comparativa de las emisiones de CO ₂ entre U.E. total mundial según datos de http://datos.bancomundial.org/indicador | 5 |
| Figura I.5. Evolución del porcentaje de población urbana elaborado con datos de http://data.worldbank.org/indicador/SP.URB.TOTL.IN.ZS | 7 |
| Figura 1.1. Polinomio de Hermite en el intervalo [0, 1]..... | 59 |
| Figura 2.1. Península Ibérica http://www.ign.es/ign/img/actividades/cartografia/mapa_1M.jpg | 101 |
| Figura 2.2. Potencial solar fotovoltaico de Europa http://re.jrc.ec.europa.eu/pvgis/ | 102 |
| Figura 2.3. Mapa de Castilla y León dentro de la península Ibérica. | 103 |
| Figura 2.4. Distribución de las estaciones de la AEMET en la Península Ibérica..... | 105 |
| Figura 2.5. Distribución de estaciones de la red SIAR en España. http://eportal.magrama.gob.es/websiar | 108 |
| Figura 2.6. Emplazamiento de la instalación de medida. Escuela Politécnica Superior de Burgos. (Google Maps)..... | 116 |
| Figura 2.7. Vistas de la estación radiométrica del grupo SWIFT ubicada en la EPS (Burgos) | 117 |
| Figura 2.8: Veleta, anemómetro (Izquierda) y Pluviómetro (derecha)..... | 119 |
| Figura 2.9: Izquierda: Sensor de temperatura y humedad relativa. Derecha: Barómetro. . | 120 |
| Figura 2.10: Subestación radiométrica. Izquierda: seguidor solar. Centro: difusa vertical MK6. Derecha: Global vertical. | 121 |
| Figura 2.11: Datalogger Campbell Scientific CR-3000..... | 122 |
| Figura 2.12. Irradiancia global vertical de las cuatro orientaciones: norte, sur, este, y oeste, para el día 15 de agosto de 2014 | 123 |
| Figura 2.13. Irradiancia horizontal global, directa y difusa para el día 15 de agosto de 2014 | 124 |
| Figura. 3.1. Distribución de las 53 estaciones de la red SIAR en Castilla y León. (Fuente: http://www.inforiego.org/opencms/opencms) | 131 |
| Figura 3.2. Valores medios anuales diarios de irradiación global horizontal de las nueve provincias de Castilla y León | 134 |
| Figura 3.3. Valores medios de irradiación media anual diaria de las 46 estaciones de la red SIAR empleadas para el estudio..... | 136 |
| Figura 3.4. Valores medios de irradiación de cada año frente a los valores medios de irradiación de los 7 años | 137 |
| Figura 3.5. Comparación entre valores medios de irradiación anual media (MJ·m ⁻²) de los 7 años con: a) media de 6 años (eliminando el año 2013) y b) media de 6 años (eliminando el año 2009). | 138 |
| Figura 3.6. Distribución del número de valores de irradiación anual media diaria en intervalos..... | 139 |

| | |
|--|-----|
| Figura 3.7. Irradiación media anual diaria (GHI / MJ·m ⁻²) para las estaciones con latitud extrema en la zona de estudio (Burgos, Valle de Losa 42.967 y Salamanca, Ciudad Rodrigo 40.590) | 140 |
| Figura 3.8. Valores medios del índice de claridad, K_p , por provincias | 142 |
| Figura 3.9. Distribución del número de días en (%) en cada intervalo del índice de claridad por provincias..... | 143 |
| Figura 3.10. Distribución del número de días en (%) en cada intervalo del índice de claridad por estaciones para Burgos..... | 144 |
| Figura 3.11. Distribución del número de días en (%) en cada intervalo del índice de claridad por estaciones para la provincia de Valladolid | 147 |
| Figura 3.12. Estaciones empleadas para hacer el estudio. Interiores (SIAR), exteriores (varias fuentes) círculos las estaciones de la AEMET..... | 154 |
| Figura 3.13. Estaciones base para interpolar y las cinco del Control1: León, León, Palencia, Valladolid y Vizcaya | 156 |
| Figura 3.14. Mapas de interpolación para SPLIN REGULARIZADO con peso 0.01: a) 12 puntos y b) 35 puntos..... | 157 |
| Figura 3.15. Mapas de interpolación Control2, con menor RMSE: a) Kriging Ordinario Circular 35 puntos y menor MAPE, b) Kriging Ordinario Lineal 35 puntos | 157 |
| Figura 3.16. Mapas de interpolación con IDW: a) 12 puntos, b) 35 puntos | 158 |
| Figura 3.17. Mapas de interpolación con Kriging: a) Universal Lineal de 12 puntos menor RMSE, b) Ordinario Gaussiano de 35 puntos mayor RMSE..... | 158 |
| Figura 3.18. Mapa de interpolación con Vecino Natural | 159 |
| Figura 3.19. Representación de las 71 estaciones base de interpolación (SIAR + diferentes medios) y cuatro de control (AEMET) | 160 |
| Figura 3.20. Mapa de interpolación Kriging Universal Cuadrático 35 puntos menor RMSE usando 71 estaciones como base y cuatro estaciones de la AEMET como control..... | 160 |
| Figura 3.21. Mapa de irradiación de Castilla y León | 162 |
| Figura 4.1. Gráfico de valores medios de Irradiación Global Anual (MJ·m ⁻²) con intervalo de confianza del 95 %..... | 169 |
| Figura 4.2. Valores entre el primer y tercer cuartil y puntos atípicos (MJ·m ⁻²) | 172 |
| Figura 4.3. Gráfica de aberrantes..... | 174 |
| Figura 4.4. Gráfica de Caja y Bigotes para Irradiación Global Media de 7 años (MJ·m ⁻²) | 177 |
| Figura 4.5. Regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)..... | 181 |
| Figura 4.6. Regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud) | 184 |
| Figura 4.7. Regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud longitud y altitud) | 186 |
| Figura 4.8. Gráfica de Caja y Bigotes para 42 estaciones (MJ·m ⁻²)..... | 189 |
| Figura 4.9. Gráfico de Probabilidad Normal..... | 190 |
| Figura 4.10. Traza de Densidad de los valores de Irradiación Global Media de 7 años con las 42 estaciones que garantizan una distribución normal | 191 |
| Figura 4.11. Histograma de los valores de Irradiación Global Media de 7 años con las 42 estaciones que garantizan una distribución normal | 191 |

| | |
|---|-----|
| Figura 5.1. Mapas de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenidos para el control C1 con los metodos de interpolación: a) KOL35 y b) KOEf35..... | 199 |
| Figura 5.2. Mapas de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenidos para el control C2 con los métodos de interpolación: a) KOL35 y b) KUC35..... | 201 |
| Figura 5.3. Mapa de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenido para el control C3 con los métodos de interpolación KOC35..... | 203 |
| Figura 5.4. Mapa de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenido para el control C4 con los métodos de interpolación KUL12..... | 206 |
| Figura 5.5. Mapas de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenidos para el control C5 con los métodos de interpolación: a) KUC35 y b) SR(0.001)12..... | 209 |
| Figura 5.6. Mapa de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenido para el control C6 con los métodos de interpolación KUC35..... | 212 |
| Figura 5.7. Valores de RMSE de media mensual diaria en superficies orientadas hacia el Norte..... | 220 |
| Figura 5.8. Valores de RMSE de irradiación media diaria ($\text{GVI}/\text{MJ}\cdot\text{m}^{-2}$) sobre superficies vertical orientado hacia el Norte..... | 221 |
| Figura 5.9. Valores de MAE de media mensual diaria en superficies orientadas hacia el Sur..... | 223 |
| Figura 5.10. Valores de MAE de irradiación media diaria ($\text{GVI}/\text{MJ}\cdot\text{m}^{-2}$) sobre superficie vertical orientado hacia el Sur..... | 225 |
| Figura 5.11. Valores de MAE de media mensual diaria en superficies orientadas hacia el Este..... | 226 |
| Figura 5.12. Valores de RMSE de irradiación media diaria ($\text{GVI}/\text{MJ}\cdot\text{m}^{-2}$) sobre superficie vertical orientado hacia el Este..... | 227 |
| Figura 5.13. Valores de MAE de media mensual diaria en superficies orientadas hacia el Oeste..... | 228 |
| Figura 5.14. Valores de MAE de irradiación media diaria ($\text{GVI}/\text{MJ}\cdot\text{m}^{-2}$) sobre Superficie vertical orientado hacia el Oeste..... | 229 |

ÍNDICE TABLAS

| | |
|--|-----|
| Tabla I.1. Consumo de energía procedente de combustibles fósiles en %. Elaborado con datos de http://data.worldbank.org/indicador | 2 |
| Tabla I.2. Energía importada (en % de la energía usada) Elaborado con datos de http://data.worldbank.org/indicador/EG.IMP.CON.S.ZS | 7 |
| Tabla 2.1. Relación de sensores climáticos. | 118 |
| Tabla 2.2. Fichero de datos con los que se trabaja | 127 |
| Tabla 3.1. Valores máximos (RGM) y mínimos (RGm) de irradiación global horizontal diaria ($\text{MJ}\cdot\text{m}^{-2}$). Se indica la fecha y la estación en la que se ha recogido el valor..... | 133 |
| Tabla 3.2. Valores medios del índice de claridad, K_t , diario para cada provincia..... | 141 |
| Tabla 3.3. Distribución de los porcentajes en los tres intervalos del índice de claridad para cada una de las nueve estaciones de León. | 145 |
| Tabla 3.4. Distribución de los porcentajes en los tres intervalos del índice de claridad para cada estación de Zamora. | 148 |
| Tabla 3.5. Tabla con datos de las estaciones | 152 |
| Tabla 3.6. Errores de Splin con 12 puntos, regularizado y de tensión con estaciones de AEMET, con 44 y 40 estaciones del SIAR..... | 161 |
| Tabla 3.7. Valores de radiación global vertical orientación sur para los cuatro modelos (circunsolar, isotrópico, Klucher y Hay) y experimentales ($\text{MJ}\cdot\text{día}^{-1}$) | 165 |
| Tabla 4.1. ANOVA para Irradiación Media Anual ($\text{MJ}\cdot\text{m}^{-2}$) por año | 168 |
| Tabla 4.2. Medias para Irradiación Media Anual ($\text{MJ}\cdot\text{m}^{-2}$) por año con intervalos de confianza del 95 %..... | 169 |
| Tabla 4.3. Pruebas de Rangos Múltiples para Irradiación por año | 170 |
| Tabla 4.4. Diferencia de valores medios de Irradiación Global sobre plano horizontal | 171 |
| Tabla 4.5. Verificación de Varianza | 172 |
| Tabla 4.6. Estimado de Localización..... | 173 |
| Tabla 4.7. Estimados de Escala..... | 173 |
| Tabla 4.8. Intervalos de confianza del 95.0 % para la media | 174 |
| Tabla 4.9. Valores Ordenados | 175 |
| Tabla 4.10. Coeficientes de la regresión simple de la Irradiación Global Media 7 años con la latitud | 178 |
| Tabla 4.11. Análisis de Varianza en el ajuste lineal de la Irradiación Global Media (7 años) con la latitud..... | 178 |
| Tabla 4.12. Estadísticos del ajuste lineal de la Irradiación Global Media 7 años con la latitud | 179 |
| Tabla 4.13. Regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)..... | 180 |
| Tabla 4.14. Análisis de Varianza de la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud) | 180 |
| Tabla 4.15. Estadísticos de la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)..... | 181 |
| Tabla 4.16. Residuos atípicos de la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)..... | 182 |

| | |
|--|-----|
| Tabla 4.17. Regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud) | 182 |
| Tabla 4.18. Análisis de Varianza en la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud)..... | 183 |
| Tabla 4.19. Estadísticos en la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud) | 183 |
| Tabla 4.20. Residuos atípicos en la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud) | 184 |
| Tabla 4.21. Coeficientes de la regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud, longitud y altitud) | 185 |
| Tabla 4.22. Análisis de Varianza de la regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud, longitud y altitud)..... | 185 |
| Tabla 4.23. Estadísticos de la regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud, longitud y altitud) | 185 |
| Tabla 4.24. Residuos atípicos de la regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud, longitud y altitud)..... | 186 |
| Tabla 4.25. Coeficientes de la regresión múltiple de la Irradiación Global Media 1 año con tres variables independientes (latitud, longitud y altitud) para los años 2007 a 2013 | 187 |
| Tabla 4.26. Tabla resumen de los valores representativos de la relación entre la irradiación y las tres variables independientes, latitud, longitud y altitud | 188 |
| Tabla 4.27. Resumen Estadístico para Irradiación Media 7 años | 188 |
| Tabla 4.28. Resultados de la Prueba de Kolmogorov-Smirnov | 190 |
| Tabla 5.1. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C1..... | 197 |
| Tabla 5.2. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C1(BI030, LE020, LE060, PA080 y VA070), con todos los métodos de interpolación | 198 |
| Tabla 5.3. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C2..... | 199 |
| Tabla 5.4. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C2 (LE020, PA080, SO020, VA060 y ZA060), con todos los métodos de interpolación | 200 |
| Tabla 5.5. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C3..... | 201 |
| Tabla 5.6. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C3 (LE020, BU050, SA030, VA080 Yza080), con todos los métodos de interpolación..... | 202 |
| Tabla 5.7. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C4..... | 204 |
| Tabla 5.8. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C4 (LE020, PA030, SG010, VA020 Y ZA040), con todos los métodos de interpolación | 205 |
| Tabla 5.9. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C5..... | 207 |
| Tabla 5.10. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C5 (LE020, SA030, VA020, ZA060 y VA080), con todos los métodos de interpolación | 208 |
| Tabla 5.11. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C6..... | 210 |

| | |
|---|-----|
| Tabla 5.12. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C6 (León, Salamanca, Soria y Valladolid), con todos los métodos de interpolación | 211 |
| Tabla 5.13. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C7 | 213 |
| Tabla 5.14. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C7 (León, Salamanca, Soria y Valladolid, LE080, PA020, VA030 y ZA010), con todos los métodos de interpolación | 214 |
| Tabla 5.15. Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) experimental e interpolada en los controles C6 y C7 | 215 |
| Tabla 5.16. Resumen de los métodos de interpolación con menores RMSE y MAE calculados en los siete controles. | 216 |
| Tabla 5.17. Resumen de los métodos de interpolación con mayores RMSE y MAE calculados en los siete controles | 217 |
| Tabla 5.18. Errores observados (diferencia entre valor experimental e interpolado) para las estaciones de la red SIAR empleadas en más de un control..... | 217 |
| Tabla 5.19. Errores anuales para los cuatro modelos de radiación vertical dirección Norte | 219 |
| Tabla 5.20. Valores de MAE mensuales para los cuatro modelos de irradiación para la orientación Norte. | 221 |
| Tabla 5.21. Errores anuales para los cuatro modelos de radiación vertical dirección Sur .. | 222 |
| Tabla 5.22. Errores anuales para los cuatro modelos de radiación vertical dirección Este. | 226 |
| Tabla 5.23. Errores anuales para los cuatro modelos de radiación vertical dirección Oeste | 227 |
| Tabla 5.24. Tabla resumen del comportamiento de los modelos considerados más adecuados para el cálculo de IGV en las direcciones cardinales..... | 229 |

INTRODUCCIÓN

I.1 MOTIVACIÓN Y OBJETIVOS

La primera fuente de energía utilizada por el hombre fue la luz, pero la primera que el hombre fue capaz de usar de forma controlada fue el fuego. Una vez que aprendió a encenderlo y apagarlo voluntariamente lo dominó y lo utilizó para cocinar, calentarse, fabricar utensilios etc. A partir de ese momento la búsqueda de distintas fuentes de energías y el aumento de su consumo ha sido incesante, siendo mayoritario el consumo de energía procedente de combustibles fósiles. Esto tiene repercusión en el medio ambiente y aunque no se puede menospreciar la utilización de energías renovables desde tiempos remotos, como el uso del viento para la navegación a vela ya desde el antiguo Egipto o la utilización del agua por los griegos para las ruedas hidráulicas, la influencia que la combustión de fósiles en dicho medio ambiente es muy importante, sobre todo desde la revolución industrial, que aumentó grandemente su consumo y en los últimos 50 años en los que se ha acelerado de forma exponencial.

La población mundial ha pasado de 6000 millones en el año 2000 a los 7300 millones en 2015, y según la previsión de la O.N.U. alcanzará los 9600 millones en 2050. El crecimiento en Europa ha sido proporcionalmente menor, pasando de 728 millones en el año 2000 a 789 millones en el 2015.

Este aumento en el número de habitantes, junto al mayor desarrollo, lleva aparejado un gran aumento del consumo energético, ver Figura I.1 (según el Banco Mundial la media del consumo por habitante en el año 2011 fue de 3045 kWh, y en la zona euro de 6599 kWh.).

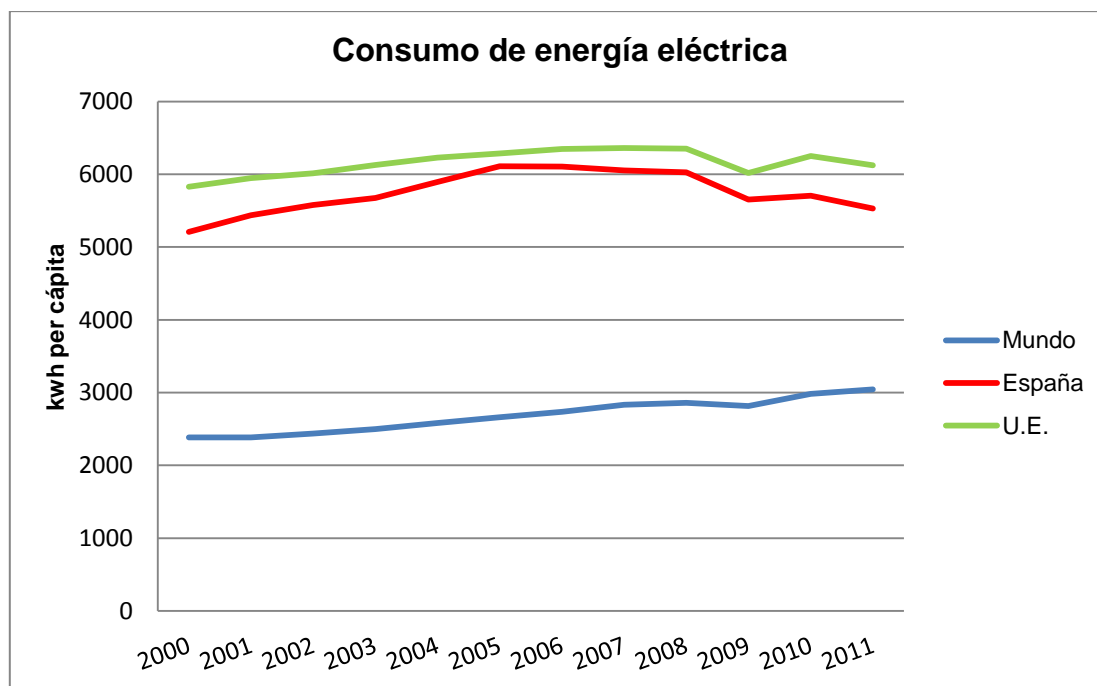


Figura I.1. Consumo de energía eléctrica. Elaborado con datos de <http://data.worldbank.org/indicator>

Teniendo en cuenta el elevado porcentaje de combustibles fósiles utilizados para la generación de energía (Tabla I.1), este aumento en el consumo supone un grave peligro para el medio ambiente.

Tabla I.1. Consumo de energía procedente de combustibles fósiles en %. Elaborado con datos de <http://data.worldbank.org/indicator>

| | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| Mundo | 80.87 | 80.96 | 81.20 | 80.92 | 80.66 | 80.80 | 81.23 | 81.29 |
| España | 83.46 | 82.50 | 83.19 | 81.80 | 79.79 | 76.10 | 76.64 | 75.89 |
| U. E. | 78.22 | 77.92 | 77.71 | 77.02 | 75.72 | 75.15 | 74.46 | 73.70 |

Según International Energy Outlook 2013, el consumo mundial de energía se incrementará en un 56 % entre 2010 y 2040. Los países de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) incrementarán su consumo en el 17 % y los no OCDE en un 90 % (los Estados asiáticos no pertenecientes a la OCDE representarán el 60 % del incremento). Es de señalar que los combustibles fósiles (incluida una cuota considerable correspondiente al carbón) seguirán representando casi el 80 % del consumo energético del mundo hasta 2040.

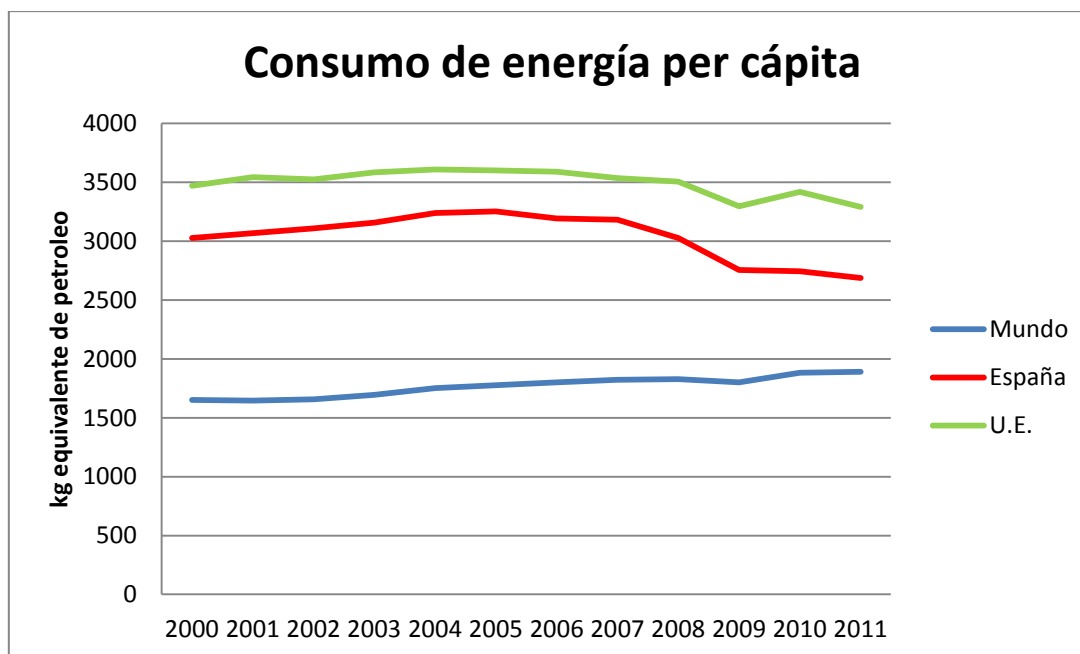


Figura I.2. Consumo de energía per cápita

El Parlamento Europeo en sus resoluciones de 5 de febrero de 2014, sobre un marco para las políticas de clima y energía en 2030, y de 26 de noviembre de 2014, sobre la Conferencia 2014 de las Naciones Unidas sobre el Cambio Climático – COP 20 de Lima, Perú (1-12 de diciembre de 2014), establece una serie de consideraciones a tener en cuenta en cualquier proyecto que involucre la generación o el transporte de energía y las emisiones de CO₂.

El Tratado de Funcionamiento de la Unión Europea (TFUE) establece dentro de sus objetivos su política medioambiental, la seguridad en el suministro energético y las fuentes de energía nuevas y renovables, y que esa política medioambiental debe contribuir a luchar contra el cambio climático.

En 2011 el Consejo Europeo volvió a confirmar el objetivo de la Unión Europea de reducir los gases de efecto invernadero entre un 80 y un 95 % antes de 2050, en comparación con los niveles de 1990, para mantener el cambio climático por debajo de los 2 °C. Teniendo en cuenta que las emisiones de CO₂ en 1990 en la U.E. eran de 8.5 Toneladas Métricas per cápita, el objetivo para 2050 es que esas emisiones se mantengan entre 0.8 y 1.7 t per cápita.

Aunque las cifras de Eurostat indican que la Unión Europea redujo sus emisiones de CO₂ en un 16.97 % entre 1990 y 2011, descendiendo paulatinamente las mismas, por lo que está en vías de alcanzar su objetivo de 2020, la situación en España difiere mucho de seguir el mismo camino, pasando de emitir 5.5 t de CO₂ per cápita en 1990 a emitir 8.1 t en 2004 y aunque posteriormente se produjo una reducción importante de las mismas (coincidiendo con la crisis económica que estamos sufriendo desde ese año), como queda de manifiesto en la figura I.3 las emisiones aún ahora se mantienen por encima de las de 1990. Para alcanzar los objetivos climáticos de 2050 la Unión Europea en conjunto, y España en particular, debe realizar unas reducciones más ambiciosas de las emisiones de CO₂ que las conseguidas hasta el momento.

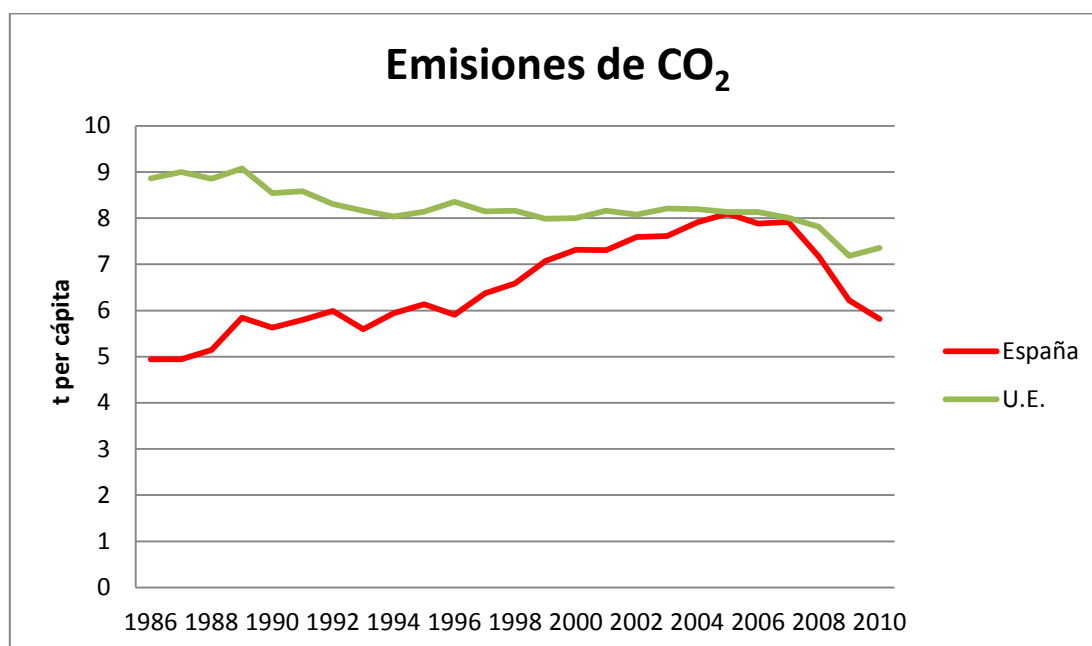


Figura I.3. Comparativa de las emisiones de CO₂ entre España y U.E. según datos de <http://datos.bancomundial.org/indicador>

Las cifras de la División de Estadísticas de las Naciones Unidas demuestran que las emisiones mundiales de CO₂ aumentaron en más del 50 % entre 1990 y 2010, y según el informe del Banco Mundial «Turn Down the Heat»; las actuales trayectorias de emisión están en vías de provocar un calentamiento de 2 °C en un plazo de 20 a 30 años y un calentamiento de 4 °C en 2100.

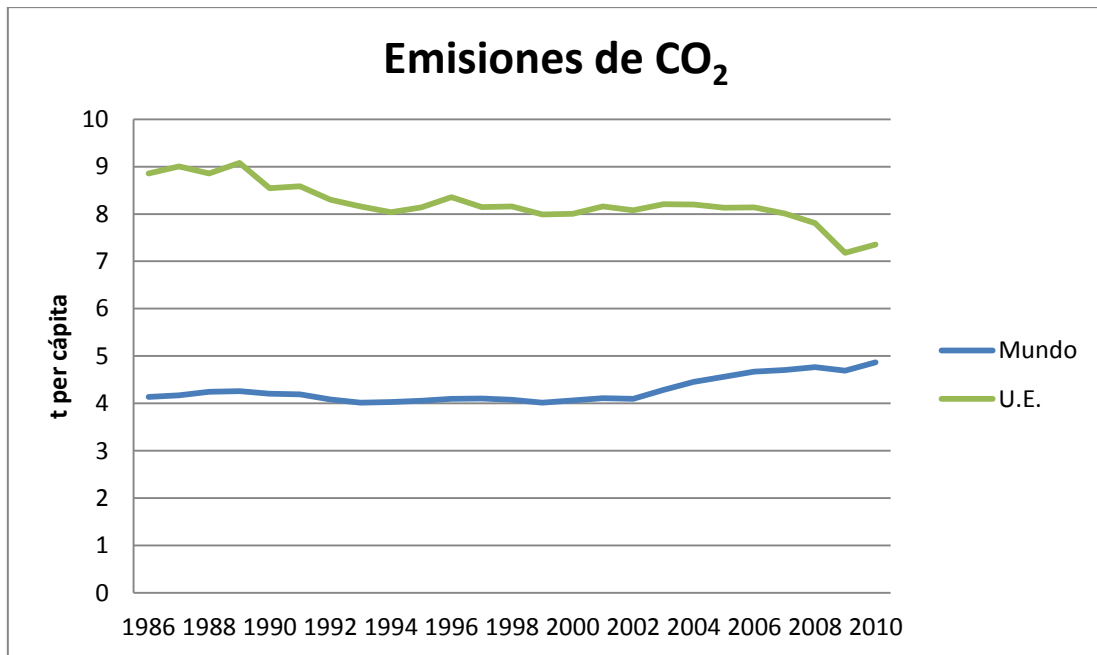


Figura I.4. Comparativa de las emisiones de CO₂ entre U.E. total mundial según datos de <http://datos.bancomundial.org/indicador>

Este cambio ha propiciado repercusiones importantes en todos los continentes y los océanos. Ejemplos de estas repercusiones las podemos encontrar en la acidificación de los océanos que pone de manifiesto *Lauvset et al. (2015)*, con las consecuencias que este hecho tiene, como el efecto sobre el blanqueo de los arrecifes de coral *Van Hooidonk et al. (2014)*.

El calentamiento producido tendrá como consecuencias alteraciones en la hidrología como lo ponen de manifiesto *Taye et al. (2015)* en su estudio sobre la afectación de la cuenca del Nilo Azul por el cambio climático, o establece *Yan et al. (2015)* en el estudio de la respuesta del Río de las Perlas al cambio climático, que prevé el aumento del estrés hídrico en la zona bajo su influencia.

El aumento del nivel del mar tendrá efectos de distinta índole en las zonas costeras, en la que los manglares de la zona intermareal de China pueden verse afectados con pérdidas de sus hábitats desde un 18 % hasta un 25 % según los distintos escenarios de cambio climático: *Li et al. (2015)*, o las playas del Mar Negro que pueden perder el 50 % de su anchura o

desaparecer en su totalidad según sea la evolución real *Allenbach et al. (2015)*.

Los efectos establecidos del cambio climático podrán variar de unos escenarios climáticos a otros como indican *Wilcox y Makowski (2014)* en un artículo sobre los efectos del cambio climático en el rendimiento de las cosechas de trigo, o *Williams et al. (2012)* en el régimen de lluvias en el Cuerno de África.

La radiación solar que llega a la superficie terrestre va variando a lo largo del tiempo y ese cambio no es uniforme. Mientras en parte del planeta como Rusia, Estonia y Alemania va disminuyendo, *Chen et al. (2006)*, posiblemente por el aumento de los aerosoles y la contaminación que provocan el efecto denominado “global dimming”, la radiación global media anual en España durante el período 1985-2010 muestra una tendencia a aumentar, *Sanchez Lorenzo et al. (2013)*, con un aumento medio de $3.9 \text{ W}\cdot\text{m}^{-2}$ por década.

El calentamiento global está relacionado con fenómenos extremos recientes, como olas de calor, sequías, inundaciones, ciclones e incendios forestales. Los riesgos asociados a un aumento de temperatura global media mayor a $2 \text{ }^\circ\text{C}$ aumentan desproporcionadamente, y si no se reduce la emisión de estos gases el aumento de temperatura media puede llegar a $5 \text{ }^\circ\text{C}$ en 2100.

Aunque no es el objeto principal de este trabajo, el uso de energías contaminantes tiene también otros efectos perniciosos, como los económicos.

Según el Parlamento Europeo, la U.E. gastó 406000 millones en la importación de combustibles fósiles en 2011 (el equivalente a más de 1000 euros por habitante) y se espera que aumente su dependencia de las importaciones de energía.

Esa dependencia supone que la Unión sea vulnerable a los precios de la energía y a las convulsiones políticas mundiales, comprometiendo la autonomía de la Unión y de los Estados miembros en materia de política

exterior; por lo que es fundamental la concentración de la U.E. en las opciones de resultados garantizados relativas a la eficiencia energética, las energías renovables y las infraestructuras energéticas

Tabla I.2. Energía importada (en % de la energía usada) Elaborado con datos de <http://data.worldbank.org/indicador/EG.IMP.CON.S.ZS>

| | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| España | 78.75 | 77.89 | 78.94 | 78.17 | 76.32 | 73.15 | 74.69 | 74.19 |
| U.E. | 49.33 | 50.39 | 50.99 | 51.13 | 50.38 | 51.25 | 51.35 | 51.97 |

Además el dinero gastado en importar combustibles fósiles apenas contribuye a la inversión, al empleo o al crecimiento en la Unión, por lo que dedicar ese dinero hacia la inversión nacional en eficiencia energética, energías renovables e infraestructuras inteligentes y permitiría crear empleos de calidad y altamente cualificados, que no pueden exportarse ni deslocalizarse, en industrias de construcción, automoción e industrias de tecnología avanzada.

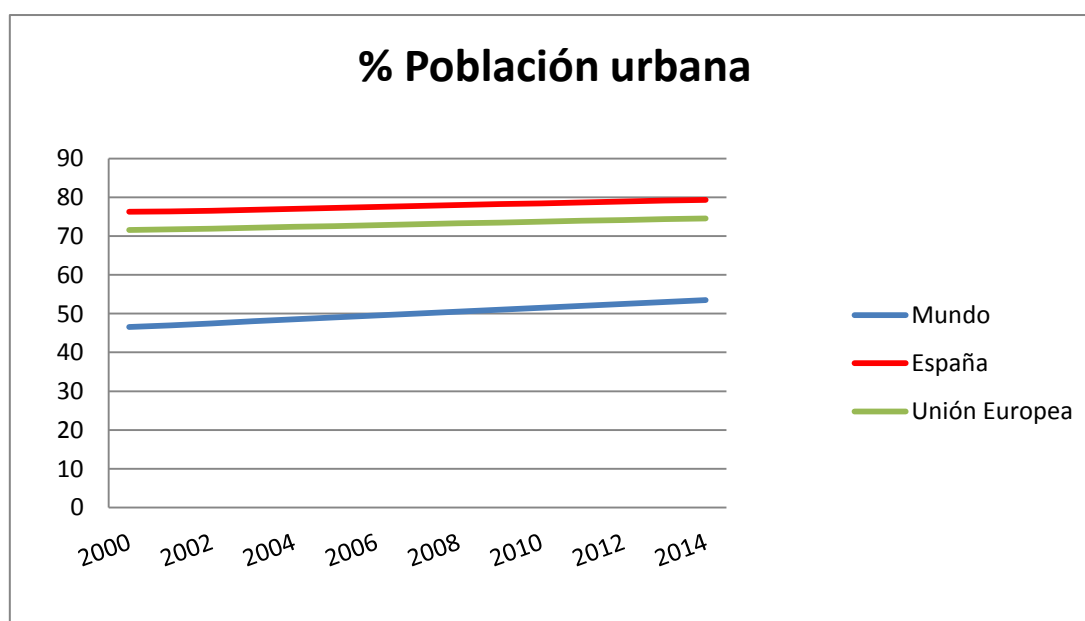


Figura I.5. Evolución del porcentaje de población urbana elaborado con datos de <http://data.worldbank.org/indicador/SP.URB.TOTL.IN.ZS>

Según los indicadores de Eurostat, el 40 % de los habitantes de la U.E. viven en zonas urbanas, pero según el Banco Mundial esa cantidad ascendía al 74 % en 2014 (<http://datos.bancomundial.org/indicador/SP.URB.TOTL.IN.ZS>) y

al 79 % en el caso de España. Estos cambios sociológicos también inciden en un mayor consumo energético y por tanto, según los parámetros actuales, en el incremento de las emisiones contaminantes.

El Quinto Informe de IPCC (*Intergovernmental Panel on Climate Change*) también proporciona posibilidades de mitigación de esos efectos, como es la utilización de energías renovables. El desarrollo de una energía renovable y limpia, incluyendo los recursos de energía solar, se ha convertido en una importante medida para combatir el cambio climático y lograr un desarrollo sostenible.

En muchas zonas del planeta hay numerosas estaciones meteorológicas que proporcionan datos de la irradiación global y difusa recibida en superficies horizontales, pero en general no hay datos disponibles sobre superficies inclinadas y estos son imprescindibles para conocer la capacidad de generación de energía en edificios en construcción y en los ya construidos, así como en la decisión de conjugar la ordenación urbanística de una zona y la posibilidad de generación de energía distribuida, logrando una mayor eficiencia energética.

Las estaciones meteorológicas no están, en general, homogéneamente distribuidas en un territorio, por lo que el primer objetivo de este trabajo ha sido resolver este problema en Castilla y León, proporcionando un mapa de radiación solar global horizontal, obtenido a partir de los datos de estaciones terrestres, que permitiera vaticinar la radiación global media esperable en una localización de la que no se dispusiera de datos.

Para utilizar eficazmente el potencial solar en zonas urbanas, generando energía directamente donde se consume, es imprescindible conocer la relación entre la radiación global incidente en un plano horizontal orientado al sur, por ser de la que habitualmente se disponen datos empíricos, y la radiación global en plano vertical, con cualquier orientación. Obtener la adecuación de los modelos existentes a dicha relación ha sido el segundo objetivo de este trabajo, para lo cual se ha instalado una estación meteorológica en la azotea de la Escuela Politécnica Superior de la

Universidad de Burgos (EPS) y se ha realizado el tratamiento de los datos y la comparación con los modelos escogidos.

I.2 SITUACIÓN INICIAL

El cambio climático está subyacente en mucha de la bibliografía consultada, manifestándose tanto en el interés por conocer el potencial solar disponible en distintos lugares del planeta, con mapas de irradiación obtenidos a partir de bases de datos satelitales o de estaciones en tierra, como en la búsqueda de distintas soluciones para mejorar la eficiencia energética de edificios o el mejor aprovechamiento de la energía solar.

Como ejemplo del primer caso, *Chelbi et al. (2015)* presenta la situación de Túnez, un país con reservas de petróleo, que ha identificado la energía solar como una prioridad en su planificación energética nacional. Realiza un estudio con datos de un número reducido de estaciones (18 entre estaciones de Túnez y Argelia para una superficie de 163610 km²) para estimar, a través de mapas, la radiación global solar media diaria en el país calculada a partir del número de horas de sol. Los resultados obtenidos le permiten afirmar que Túnez tiene un gran potencial solar, convirtiendo este tipo de energía en una de las claves para el futuro energético del país, información de gran importancia para el sector de la energía.

Das et al. (2015) estima la radiación solar en superficie horizontal en Corea del Sur utilizando 79 estaciones de medida para la obtención de distintos parámetros meteorológicos, y 37 con datos de radiación solar. Emplea 22 de ellas para obtener el número de horas de sol y utilizarlas para construir el mapa de radiación solar, empleando datos desde 2001 hasta 2012 para realizar el modelo y 15 estaciones para validar el estudio con los datos de 2013, constatando el fuerte gradiente latitudinal observado.

Ruiz-Arias et al. (2015) corrige los datos obtenidos por bases satelitales con datos de la red SIAR y la AEMET, para presentar mapas mensuales de radiación global y directa de España.

Polo (2015) utiliza la metodología desarrollada por el CIEMAT para presentar los mapas mensuales de irradiación global y directa de España obtenidos a partir de datos de satélites, y utiliza 27 estaciones en tierra de AEMET para la evaluación del error cometido. Los mayores errores en la irradiación global se obtienen en el norte del país habiendo sobreestimado la irradiación. Para la irradiación directa es la estación de La Coruña la que presenta uno de los mayores errores.

Pandey y Katiyar (2015) utilizan los modelos circunsolar, Liu y Jordan, Klucher y Hay para estimar la radiación en una superficie inclinada en la India. Utiliza superficies inclinadas 15° , 30° , 45° y 60° y los datos recogidos el año 2007 para estimar la media mensual horaria de radiación total, como suma de las radiaciones directa difusa y reflejada.

Buscando soluciones para mitigar los efectos del cambio climático aparecen numerosos artículos, entre ellos *Appelbaum (2016)* estudia el rendimiento de células solares bifaciales comparándolas con los paneles solares tradicionales constatando que, en orientación sur y ángulo de inclinación óptimo, el rendimiento es mayor que en posición vertical y orientación este-oeste, y que en este momento las ventajas económicas pueden no ser suficientes para la producción de electricidad a gran escala.

Bahr (2014) propone estrategias que reduzcan la demanda energética de los edificios con persianas provistas de módulos fotovoltaicos. La inclinación óptima para la producción de energía es de 75° en todas las direcciones, pero hay que tener en cuenta en el diseño las sombras sobre los módulos, la reducción de la luz diurna, la relación entre la distancia entre módulos y su longitud etc.

Monger et al. (2016) estudia mediante la interpolación con el método Kriging la variabilidad de la radiación solar en sitios cercanos a un punto, buscando cuantificar la variabilidad de este recurso en el norte de Arizona en plantas de energía fotovoltaica de 10 y 30 MW, que podrían interconectarse a una línea de transmisión o de distribución. Además busca determinar el número

de sensores de irradiancia necesarios y la exactitud de su capacidad para captar la variabilidad de estas plantas de energía.

Li et al. (2014) calcula la radiación media mensual en superficie vertical tomando datos de un año en Beijing en fachadas orientadas al sur, al este y al oeste. Al ser una zona representativa de clima monzónico (cálido, semi-húmedo y semiárido) elige el modelo isotrópico de Liu y Jordan y los anisotrópicos de Hay y de Reindl para comprobar si alguno de ellos es adecuado para la estimación de la radiación. Concluye que el mejor modelo es el de Liu y Jordan y que si se considera la radiación anual es preferible la orientación sur para compensar la mayor demanda energética de los edificios.

Wang et al. (2014) analiza el coste por kWh la sustitución de energía fósil y el nivel de mitigación de CO₂, así como el coste por unidad de reducción de CO₂ de la generación de energía fotovoltaica en 2020 a nivel provincial en China examinando tres tipos de instalaciones fotovoltaicas: Instalaciones a gran escala, integradas en edificios y sistemas fotovoltaicos distribuidos, de capacidad de generación muy baja, utilizados en zonas rurales.

Por último, indicar que la Universidad no es ajena a esta problemática, como lo demuestra la presentación de un Trabajo Fin de Grado para la instalación en la fachada principal de la Facultad de Ciencias de la UBU, con orientación sur, de módulos fotovoltaicos detrás de una superficie de 96.95 m² cubierta por vidrio semitransparente (vidrio oscuro con un 20 % de semitransparencia), que generaría 4215 kWh al año y una potencia pico de 3.6 kW. Soluciones semejantes son factibles en muchos de los edificios de la Universidad y supondrían un importante ahorro en emisiones de CO₂, así como enviar el mensaje de que la Universidad es sensible a uno de los acuciantes problemas que tiene planteada la sociedad actual: la necesidad de orquestar medidas para la viabilidad de la pervivencia humana en el planeta.

I.3 LÍNEA DE EXPOSICIÓN

El trabajo está dividido en tres partes:

1. Descripción de los métodos de interpolación y modelos de radiación sobre superficie inclinada (Capítulo 1) y zona de estudio y bases de datos (Capítulo 2).
2. Trabajo desarrollado con los datos, métodos y modelos (Capítulo 3); análisis estadísticos de los datos (Capítulo 4) y resultados (Capítulo 5).
3. Conclusiones y futuras líneas de investigación abiertas.

En el Capítulo 1, después de un exhaustivo estudio de la bibliografía en este campo, se describen los métodos matemáticos de interpolación más utilizados por los diferentes autores. Se analiza la forma de trabajar con los modelos en el software ArcGis definiendo los conceptos fundamentales para su aplicación. A continuación se hace un recorrido por la bibliografía de los modelos de radiación sobre superficie inclinada, describiendo los diferentes modelos, su aplicación y los resultados de la aplicación de los mismos en diferentes lugares.

Para poder aplicar tanto los métodos matemáticos como los modelos de radiación es necesario disponer de una base de datos de una determinada zona. En el Capítulo 2 se detalla primero la zona en la que se va a realizar el estudio con sus características geográficas y climáticas. La segunda parte del capítulo se dedica a la descripción de las diferentes bases con las que se va a trabajar, datos de la red SIAR, de la AEMET y propios del grupo de investigación SWIFT.

Conocidas las características de los datos y el lugar de aplicación de los mismos, en el Capítulo 3 se detalla la forma de trabajar. Se empieza con un estudio de los datos de la red SIAR, se describen los pasos que se van a seguir para sacar la mayor información acerca del valor de la irradiación media anual diaria y del índice de claridad. La segunda parte se centra en la aplicación de los datos de la red SIAR y AEMET a los métodos de

interpolación seleccionados en el software ArcGis. La última parte del capítulo describe la forma de trabajar con los datos experimentales para aplicar los modelos de radiación.

En el análisis de los datos de la red SIAR se observa un comportamiento peculiar de alguna estación de medida y unas características comunes a todas ellas. Para comprobar esas observaciones, en el Capítulo 4 se realiza un completo análisis estadístico de los datos.

En el Capítulo 5 se exponen los resultados obtenidos en este trabajo. La primera parte se dedica a analizar exhaustivamente los métodos matemáticos de interpolación, lo que permite determinar cuál es el que mejor se adapta a la zona en estudio. En la segunda parte se describen los resultados obtenidos al aplicar los cuatro modelos de radiación con los que se trabaja, a los datos experimentales sobre superficie vertical en las cuatro orientaciones, determinando el que mejor se ajusta a la zona de medida.

La parte final de este trabajo se dedica a exponer las conclusiones del mismo junto con las futuras líneas de investigación abiertas.

CAPÍTULO 1.

FUNDAMENTOS

1.1 INTRODUCCIÓN

En este capítulo se presentan los métodos de interpolación que se van a utilizar para obtener el mapa de irradiación solar de Castilla y León y los modelos matemáticos que permiten pasar de radiación global sobre superficie horizontal a radiación global sobre superficie vertical.

1.2 MÉTODOS DE INTERPOLACIÓN

¿Qué es Interpolación?, muchas son las definiciones que se pueden dar para este concepto, pero cada una de ellas es válida para un determinado campo de la ciencia. Desde un punto de vista matemático, se puede decir que interpolación es obtener nuevos valores, partiendo del conocimiento de un conjunto discreto de puntos en un determinado dominio. En el campo de otras ciencias, incluida la ingeniería, la interpolación consiste en obtener una función que se ajuste al comportamiento de un conjunto de valores obtenido por muestreo.

El problema que se quiere resolver en este trabajo radica en calcular el valor de la radiación solar en un lugar en el que no existe medida real. Las matemáticas pueden resolver este problema clásico, de una forma aproximada, buscando una función fácil de construir y de evaluar, que tome valores próximos a los dados en los puntos donde hay datos y permita asignar valores admisibles en los puntos donde no hay datos. Por lo tanto, la resolución aproximada del problema consiste en encontrar, con los datos de que se dispone, una función que coincida o sea próxima a la función objeto

del problema. Se dice que la función así construida **interpola** a la función dada con respecto a los datos.

El trabajo se desarrolla con datos geoespaciales a partir de los cuales se definirán determinadas funciones que tomarán valores en un dominio espacial. Por esto es recomendable considerar la primera ley geográfica de Tobler (1979): “*Todo está relacionado con todo, pero las cosas cercanas están más relacionadas entre sí que las más lejanas*”.

Existen muchos algoritmos matemáticos para interpolar, según se ha comprobado con la revisión bibliográfica realizada. *Franke, (1982)* relaciona distintas técnicas de interpolación. *Lam, (1983)* presenta una clasificación de los métodos de interpolación en función del grado de exactitud o de aproximación. *Nalder y Wein, (1998)*, *Apaydin et al., (2004)*, *Wu et al., (2013)*, entre otros, utilizan y relacionan distintos métodos de interpolación.

En este capítulo se van a referenciar los métodos de interpolación empleados en el campo de la geoestadística, es decir aquellos que permiten obtener un resultado fácilmente representable por medio de mapas o superficies en los que se puede predecir el valor de la variable en cualquier punto de la región representada.

Primeramente se hace un balance de los métodos encontrados en la revisión bibliografía para describir después, matemáticamente, los métodos de interpolación más empleados por los autores. A continuación se hace una descripción de los métodos empleados en el software de trabajo, ArcGis 10.3, definiéndolos, estableciendo el procedimiento de uso y comentando su funcionamiento. De todos ellos se elegirá el más adecuado para el problema planteado: Calcular los valores de radiación solar global sobre superficie horizontal en la región de Castilla y León a partir de un conjunto de estaciones de medida repartidas no uniformemente en la región.

1.2.1 REVISIÓN BIBLIOGRÁFICA

Una de las formas más sencillas de interpolar es ponderar el inverso de la distancia entre el punto a estimar y los puntos de la muestra. Los pesos asignados a los datos observados son inversamente proporcionales a su distancia al punto en el que se realiza la estimación (Método IDW, *Inverse Distance Weighting*). En este caso, existe una influencia de cercanía entre los valores, que se puede definir de forma determinista y analítica como *Gutierrez-Corea et al., (2014)*. *Güler, (2014)* utiliza este método para obtener mapas de la evapotranspiración en Turquía, utilizando sistemas de información cartográfica. *Apaydin et al., (2004)* usa la interpolación IDW para determinar la mejor distribución de seis parámetros climáticos en Turquía, la radiación solar entre ellos. El valor de la variable se estima tomando el promedio ponderado de los valores vecinos medidos. *Wu et al., (2013)* asigna un peso por medio de una potencia negativa, en este caso de grado dos, de la distancia euclídea entre el valor medido y el estimado. En algunos casos las condiciones de la zona de estudio hacen que el método IDW arrastre errores considerables para la predicción. *Chung y Yun, (2004)*, considerando la influencia que la altitud y la diversidad geográfica en una interpolación de temperatura, calcula los errores a partir de valores de radiación solar en todas sus componentes, basándose en la geometría del lugar.

Una variante del modelo descrito es la que trabaja con el inverso de la distancia al cuadrado (Método IDS, *Inverse Distance Squared*). *Nalder y Wein, (1998)* lo utilizan para estimar la temperatura y pluviosidad en una zona del oeste de Canadá, encontrando que el modelo explica gran parte de las variaciones de un conjunto de valores. Su facilidad de implementación y sencillez de cálculo hace de este método un procedimiento muy atractivo. Estos autores introdujeron una variante denominada GIDS (*Gradient Plus-Inverse Distance Squared*) en el que las regresiones múltiples permiten definir gradientes climáticos y promediar los valores obtenidos mediante el inverso del cuadrado de la distancia. *Price et al. (2000)* utiliza este mismo método para interpolar espacialmente datos climáticos promedio mensuales

en Canadá. En este caso, las múltiples regresiones lineales para estimar los gradientes locales de cada variable se realizan con las estaciones más cercanas y se atribuye una autocorrelación de los datos intrínseca al método. Sin embargo, se advierte que en las regiones montañosas, donde la topografía tiende a reducir las correlaciones espaciales, la previsibilidad de las variables queda afectada. En *Stahl et al., (2006)* se combina la regresión múltiple con el inverso de la distancia al cuadrado. Primero se ajusta el modelo de regresión múltiple, utilizando los coeficientes calculados para ajustar la temperatura, la variable con la que trabaja en este caso, para el punto de predicción en cada estación de validación; posteriormente las temperaturas se promedian utilizando los pesos del inverso de la distancia al cuadrado para cada temperatura que se pretende predecir.

Muchos autores tratan de interpolar valores por medio de modelos de regresión, combinaciones lineales o ajustes polinómicos. En *Karim y Singh, (2014)* se emplea un modelo con una función de regresión y errores independientes de la media nula con varianza común, estudiando el comportamiento con polinomios de diferentes grados desde uno hasta seis. Las variables pueden ser diferentes según el tipo de interpolación empleada: *Wu et al., (2013)*, interpola valores de radiación solar con un modelo lineal múltiple basado en las coordenadas UTM (*Universal Transverse Mercator*), la altitud y las horas de insolación. *Apaydin et al., (2004)* sin embargo, emplea un método de interpolación polinómica, ajustando una superficie definida por un polinomio (Método GPI, *Global polynomial interpolation*) a los datos. De este modo se crea una superficie que varía lentamente, usando polinomios de orden inferior. Un polinomio global de primer orden representa un plano, uno de segundo orden se ajusta a una superficie con una curva, permitiendo el cálculo de superficies que representan valles, y así sucesivamente. Las superficies calculadas de esta manera son muy sensibles a los valores extremos o atípicos, extremadamente altos o bajos y el método no es válido cuando la superficie presenta pendientes y picos. En el mismo trabajo, se emplea un método de interpolación polinómica local (Método LPI, *Local Polynomic Interpolation*), que a diferencia del anterior

(GPI), ajusta con muchos polinomios, cada uno superpuesto a los vecinos próximos. Este método genera superficies que varían de forma más local.

Otra posibilidad consiste en utilizar un método de interpolación para validar una variable que luego se usa en un semivariograma. *Şen y Şahin, (2001)* estiman el valor de irradiación solar en cualquier punto deseado a través de un procedimiento promedio ponderado mediante la técnica de mínimos cuadrados.

Uno de los sistemas más empleados es el Método Kriging, (*Krige, (1951)*) Consiste en un proceso geoestadístico avanzado, basado en métodos estadísticos que utiliza la autocorrelación, generando una superficie de predicción a partir de un conjunto disperso de valores medidos. En algunos casos, se usa un enfoque bayesiano como en *Wilson y Silander, (2014)*, que usan una aplicación de este tipo para interpolar 20 años de datos meteorológicos.

Desde el punto de vista matemático existen diferentes tipos de interpolación Kriging: ordinario, simple y universal. Hay autores que aplican todos ellos con el objetivo de encontrar la mejor resolución e incluso generan nuevos tipos a través de variantes. *Gutierrez-Corea et al., (2014)* realiza una estimación espacial de Irradiancia Global Horizontal (GHI, *Global Horizontal Irradiance*) basada en observaciones oficiales y sensores remotos en el noreste de España a través de varios métodos Kriging, entre ellos el Kriging ordinario. También utiliza dos variaciones de Kriging de regresión con variables auxiliares, que incluyen nueva información para mejorar el proceso apoyando a la variable objeto. *Wu et al., (2013)* emplea siete métodos de interpolación para la estimación espacial de la media mensual de horas de insolación en China. Se emplea el método Kriging ordinario y una versión de regresión con las coordenadas UTM y la altitud en un Kriging ordinario de la regresión residual.

Si se trabaja con valores de radiación solar, el método Kriging puede proporcionar resultados muy adecuados cuando el terreno es homogéneo y las características climáticas similares, pero cuando se trabaja con áreas de

topografía compleja en las que la altitud y la orientación puedan proyectar sombras por la propia variación topográfica, los resultados no son tan buenos. *Alsamamra et al., (2009)* emplea el Kriging ordinario y un método Kriging residual con variables externas, que aportan información complementaria en la interpolación, mejorando los resultados. En este caso, las variables empleadas son la altitud y las sombras producidas por la propia topografía del terreno. *Şen y Şahin, (2001)* trabajan con 29 estaciones meteorológicas terrestres en Turquía, tratando de determinar valores mensuales de radiación solar en puntos en los que no hay estación. Emplea tres procedimientos de interpolación: interpolación lineal, funciones de ponderación clásica y semivariograma acumulativo. En el primero se transfiere la información a una expresión matemática que representa las ponderaciones espaciales en función de las distancias. En el segundo procedimiento, funciones de ponderación clásica, el valor de radiación estimado en un punto es el cociente entre una función de ponderación (que depende del lugar *i*-ésimo y el estimado), y la distancia entre el punto *i*-ésimo en el que se conoce la radiación y el estimado. El tercer método que emplea es el semivariograma acumulativo, que es una alternativa al semivariograma clásico con alguna ventaja en el campo de la dependencia espacial. Para cada distancia se calcula la diferencia media al cuadrado de los valores de radiación medidos. A continuación se hace la suma de todos ellos, obteniendo una función no decreciente. La función no es subjetiva puesto que los promedios se toman dentro de un intervalo concreto que se considera como el valor representativo para ese intervalo.

Para el caso de variables espacio-temporales como la pluviometría, *Kebaili Bargaoui y Chebbi, (2009)* utilizan, junto con el método Kriging ordinario, una variación de Kriging denominado KED (*Kriging External Deviation*) para calcular los mapas de varianza Kriging. El KED es una estimación en 3-D del variograma frente al enfoque 2-D del Kriging ordinario, que presenta según los autores un resultado significativamente mejor. En el caso de disponer de bases de datos continuos pero con alguna laguna, *Jeffrey et al., (2001)* utiliza el método Kriging ordinario para calcular los valores no disponibles en una base de datos climáticos para Australia. *Goovaerts, (2000)* emplea

distintos modelos para el cálculo de la pluviometría al sur de Portugal con diferentes métodos de interpolación, entre ellos: Kriging ordinario, Kriging simple con variación de la media local, Kriging con desviación externa y co-Kriging. Concluye que el Kriging con variación de la media local ofrece mejores resultados para la mayoría de los meses.

Apaydin et al., (2004) emplea el método Kriging ordinario, el más general y ampliamente utilizado, en su estudio sobre el clima en Turquía. En el mismo estudio se aplica Kriging universal y Kriging simple. El primero (universal) supone que puede haber una tendencia predominante de los datos. Esto permite modelarlo por medio de una función determinista, es decir un polinomio. El método Kriging simple utiliza el promedio de todo el conjunto de datos, lo que le puede hacer que sea el menos preciso. Además añade una modificación denominada Kriging disyuntivo, que requiere un normalidad bivalente. Esta suposición es difícil de verificar, y las soluciones son más complejas, tanto matemática como computacionalmente. Para completar el estudio se trabaja con el método CoKriging con sus cuatro variables, utilizando co-variables adicionales. El CoKriging aúna las ventajas de la correlación entre variables, lo que le hace más efectivo cuando las variables están muy correlacionadas.

Nalder y Wein, (1998) utilizan para cada variable climática un modelo de interpolación distinto. Para la aplicación del método Kriging ordinario se supone una estacionalidad de segundo orden, es decir, la función del variograma depende de la separación y no de la localización. En el modelo CoKriging aprovecha la correlación entre dos variables, la variable climática y la altitud, utilizando tres modelos de variograma en cada operación, siendo la altitud una covariable. Define un Kriging sin tendencia, en el que realiza una regresión lineal múltiple, con las variables temperatura y presión, para tener una mejor respuesta a la estacionalidad (que no deja de ser un Kriging residual). El cuarto modelo de Kriging utilizado es el universal, mediante un modelo de variograma generado previamente por la temperatura y la presión, para especificar a continuación un modelo de tendencia de primer orden con la posición y la altitud como variables independientes. Los

resultados obtenidos son muy sensibles a los vecinos aunque la dependencia es menor que en el método Kriging ordinario.

Un método muy frecuente en la mayoría de los trabajos revisados es la interpolación por Splines. Un Spline es una función definida a trozos, por medio de polinomios, generalmente de grado bajo, para evitar las oscilaciones que se pueden producir con polinomios de grado superior. La condición que se impone es que el polinomio verifique todos los valores de entrada. Se buscan funciones suficientemente suaves para que no se produzcan cambios bruscos. Si la función es de una sola variable su gráfica es una curva y si es de dos variables la gráfica será una superficie en el espacio.

Wahba y Wendelberger, (1980), desarrollan un procedimiento matemático usando una función con cuatro variables. Los autores modifican el método Sasaki (*Sasaki et al., (1960)*) en cinco aspectos: a) usan directamente los datos, sin interpolación; b) no aplican valores frontera; c) en lugar de usar datos históricos o suposiciones como parámetros de suavizado, emplean los propios datos fuente; d) establecen un procedimiento que permite mezclar datos medidos experimentalmente con datos de satélite y e) por último emplea la discretización para que no se propaguen los errores.

Jeffrey et al., (2001) utiliza este método para completar su base de datos climática en Australia para 4600 localidades, en un periodo de tiempo que va de 1890 al 1957.

Apaydin et al., (2004) utiliza Splines regularizados completamente (CRS, *Completely Regularized Spline*), definiendo una función de base radial para realizar una interpolación exacta. La superficie se va forzando a través de cada valor medido. Las funciones de base radial incluyen, además de Splines, otras funciones, como la multicuadrática y la multicuadrática inversa. Cada función tiene una forma distinta dando lugar a una superficie de interpolación ligeramente diferente. Estos métodos son conceptualmente similares al ajuste de una membrana flexible que pasa por los puntos correspondientes a los valores medidos, minimizando al mismo tiempo la

curvatura total de la superficie. La función básica seleccionada determina cómo se ajusta la membrana flexible a los valores. A diferencia de otros métodos de interpolación, como el IDW, la interpolación con Splines puede predecir valores por encima del máximo y por debajo del mínimo de todos los valores medidos. Este método se usa generalmente cuando se dispone de abundantes datos y, además, la superficie en la que se va a llevar a cabo la interpolación es más o menos lisa. Los resultados son buenos cuando la superficie no tiene cambios bruscos y no es adecuado cuando hay muchas variaciones o cambios en distancias pequeñas.

El método de Splines suavizados de placa fina se emplea para interpolación de variables climáticas, pero su aplicación implica gran número de operaciones, por lo que las rutinas de cómputo se vuelven inviables para conjuntos de datos con más de 2.000 puntos. Para solventar ese problema *Hancock y Hutchinson, (2006)* generan un algoritmo simple y rápido en el que, haciendo rejillas en la superficie a interpolar, se calcula por aproximaciones numéricas con Splines suavizados de placa fina bivariada. El algoritmo se diseñó precisamente para poder ampliar la aplicabilidad del método a los datos climáticos. Los autores usan varios métodos para realizar la interpolación espacial de un gran conjunto de datos y lo aplican en Australia y África. El primer modelo utilizado es el de Splines de placa fina (TPS, *Thin Plated Spline*), en el que se incorpora una dependencia lineal sobre un submodelo paramétrico dentro del modelo Splin.

La validación cruzada generalizada (GCV, *Generalized Cross Validation*) con Splines de capa fina, usando la metodología de Hutchinson (2000) también se emplea en el trabajo referenciado. Se usan productos de B-splines cuadráticos jerárquicos para discretizar las ecuaciones. El sistema discretizado se resuelve mediante una multirejilla anidada. El procedimiento consiste en una doble iteración: la estimación de la placa fina y la del parámetro de suavización de forma simultánea. Con esta doble iteración se consigue optimizar la suavidad a la vez que se resuelve el sistema discretizado en lugar de resolver el sistema completo para un número de parámetros de suavizado fijos. El sistema discretizado se emplea para

obtener la bivalente discretizada del sistema de Splin de capa fina. Lo importante del método reside en que la rugosidad de la aproximación discretizada se minimiza solamente en la cuadrícula definida y no hay condiciones de suavidad fuera de la misma.

Otro método utilizado por *Hancock y Hutchinson, (2006)* es el MINGCV. Se trata de un algoritmo para aproximar mediante elementos finitos al mínimo de la validación cruzada generalizada del Splin suavizado de placa fina. Este modelo incorpora un proceso de iteración para optimizar el suavizado del parámetro λ , cuyo valor correspondiente al mínimo se consigue mediante una aproximación de series de Taylor de segundo orden. Una variante del método de Splines suavizados de placa fina aparece en *Wu et al., (2013)* para una estimación espacial de la media mensual de las horas de insolación en China.

Price et al., (2000), utiliza el software ANUSPLIN (*Australian National University Splin*) desarrollado por la Universidad Nacional de Australia en código Fortran, programa de uso generalizado. El programa resuelve una interpolación multivariable de datos con ruido, utilizando Splines suavizados de capa fina, en el que se usan habitualmente latitud, longitud y altitud como variables independientes. Los resultados en Canadá superan los logrados mediante la interpolación con GIDS ya que en este caso, el método ANUSPLIN, es capaz de generar gradientes más suaves y por lo tanto con mejor resultados, sobre todo en situaciones en las el número de datos disponible es escaso y con altitud elevada, porque calibra mejor en áreas de pendientes no uniformes.

Un ejemplo adicional se encuentra en *Newlands et al., (2011)*, utilizando el software disponible en (<http://cres.anu.edu.au/outputs/anusplin.php>) de la Universidad Nacional Australiana en Canberra. Este software contiene rutinas numéricas científicas para generar estimaciones en una cuadrícula regular de diversas variables climáticas y meteorológicas y ha sido también aplicado por diferentes grupos de investigación.

La versión ANUSPLIN 4.3 (Hutchinson 2004) ha sido utilizada por otros autores, como *Hijmans et al., (2005)*. Este método es eficiente y fácil de ejecutar y ha sido contrastado en estudios globales con buenos resultados. Adicionalmente, este mismo trabajo utiliza una variante denominada ANUSPLIN-SPLINA, donde cada estación (dato) se emplea como un punto (no como el Splinb que usa cada estación como un subconjunto), generando un Splin de segundo grado que tiene la latitud, longitud y altitud como variables independientes ya que producen errores de validación cruzada globales más pequeños.

En el caso de trabajar con variables agroalimentarias en zonas de topografía compleja (con variación significativa de los parámetros es distancias cortas) y que, además, estén influenciados por factores externos, como puede ser la vegetación, se recurre a una combinación de diferentes métodos para obtener resultados mejores que los valores promedio. Debido a la variabilidad de los valores de trabajo, la interpolación puede ser un método aconsejable para obtener valores adecuados. El problema reside en seleccionar el método óptimo. *Güler, (2014)* emplea, entre otros, el método de interpolación de Splines, el de Splines de capa fina y Splines regularizados completamente para obtener valores de evapotranspiración en Turquía trabajando con 72 estaciones en una región de 865 km².

Newlands et al., (2011) utiliza el método DAYMET, programado para usarse on-line en 2008 por el grupo de Simulation Terradynamic Numerical U.S. de la Universidad de Montana (www.daymet.org). En este método, la variable se determina por una función de interpolación binomial ponderada por la distancia del punto a las estaciones de medida del entorno y por la variación de dicha variable con la diferencia de altitud.

Existen otros métodos de interpolación utilizados en la bibliografía en trabajos de este campo de conocimiento y distintos de los comentados hasta el momento y que consideramos pueden ser de interés para este trabajo. *Zhang et al., (2014)* desarrolla un modelo de cálculo de valores de irradiancia directa a partir de datos de irradiancia global utilizando una base de datos y extendiendo el estudio a la base de datos NASA. El estudio se

realiza con tres métodos diferentes, DIRINDEX, DISC y DIRINT. El primero de ellos es una combinación de dos métodos anteriores, una versión simplificada del modelo SOLIS (*Ineichen, (2008)*), para irradiancia global en cielo claro. El segundo método, DISC, es un modelo cuasi físico en el que las entradas son DNI (*Direct Normal Irradiance*) y se obtienen multiplicando la irradiancia normal en la capa atmosférica, I_0 , por la transmitancia atmosférica directa normal de radiación, que está empíricamente relacionada con la transmitancia atmosférica global horizontal o el índice de claridad. El tercer método, DIRINT, es una modificación del SOLIS realizada por Perez (*Perez et al., (2002)*) que introdujo un factor de las condiciones de insolación, depende del cenit solar, del ángulo cenital de la columna empírica de agua y de un índice de estabilidad. Realmente no son métodos de interpolación, ya que no buscan encontrar valores de variables en un lugar en el que no hay estación de medida, sino más bien modelos de predicción que estiman valores de esa variable en un tiempo futuro a partir del conocimiento de la evolución de la variable en ese punto.

Stahl et al., (2006) realiza una comparación de diferentes métodos de interpolación para la temperatura diaria de una región extensa con una topografía variable y con una densidad de estaciones de medida alta, utilizando 11 métodos diferentes para Canadá. El denominado NN (*Nearest Neighbor*), vecino más próximo con ajuste de la altitud, utiliza la estación más cercana al punto de predicción para especificar la temperatura más cercana conocida. El siguiente método lo denominan LWR-G (*Lapse rate by Weighted Regression with Gaussian filter*), está basado en el DAYMET definido en anteriormente. Este método emplea un algoritmo de densidad de forma iterativa para determinar un conjunto local de estaciones a predecir y cada punto tiene asignado un peso que disminuye con la distancia al punto de predicción.

El tercer método, CLWR-G (*Constrained Lapse rate by Weighted Regression with Gaussian filter*) es simplemente una modificación del anterior contemplando la posibilidad de tener un punto de predicción por encima o por debajo de la malla establecida con las estaciones de medida, porque el

terreno en estudio sea montañoso y la variación de la altitud pudiera ser considerable en algunos casos y o estar fuera del rango establecido. Otro método, el SLT-G emplea el mismo sistema para seleccionar los puntos de predicción que los dos anteriores, pero en este caso la temperatura medida se ajusta primero (utilizando un gradiente de temperaturas mensuales especificadas previamente), para, posteriormente, ajustar la temperatura prevista a la altitud del punto de predicción utilizando la tasa de error especificada.

A continuación emplea dos métodos SLR-K y SLR-KQ, en los que la temperatura se ajusta primero a nivel del mar usando los gradientes térmicos medios mensuales; los pesos los calculan con el Kriging ordinario y se emplea un variograma lineal para evitar problemas en la determinación del modelo de semivariograma óptimo para cada punto de predicción diario. Solo se usa seis de los puntos más cercanos dentro de cada una de las rejillas que rodean al punto de predicción, porque determinan una distancia específica de truncamiento. Los SLR_ID y SLR_IDQ son muy similares a los anteriores excepto en la asignación de pesos, que en este caso se calculan mediante el inverso de la distancia. Se definen los SLR-IDS y SLR-IDSQ semejantes a los dos anteriores pero la expresión para calcular los pesos utiliza el inverso de la distancia al cuadrado. Por último, el MLR es un ajuste de una regresión múltiple usando la longitud, la latitud y la altitud como variables predictores.

Nalder y Wein, (1998) aplican un método más sencillo que el IDW, el NN, utilizando los datos de la estación más próxima para estimar los del punto de prueba, aceptando que la estación de medida es representativa de un área. No es exactamente un método de interpolación, pero se puede emplear como base de pruebas para otros métodos.

Yang et al., (2014) utiliza el diagrama de Hovmöller, una combinación de histogramas, en los que la longitud (o latitud) y el tiempo son los ejes del histograma. En este caso, cada columna del bloque en el histograma está representado por un diagrama a una distancia concreta. Es una forma estadística para representar eficazmente la magnitud de retraso de tiempo

entre dos estaciones. *Riccio, (2005)* usa un procedimiento de inferencia bayesiana que denomina MM5. Es un método de interpolación espacio-temporal de un conjunto de datos medidos formando una distribución irregular. La interpolación se basa en el uso de la distribución predictiva posterior. Se utiliza la aproximación Bayesiana para estimar los parámetros del proceso ARMA estacional. *Antoni et al., (2001)* relaciona valores de cada variable climática observada para la mayoría de las estaciones como una función empírica de la latitud, la longitud y la altitud y los respectivos valores observados para un número limitado de estaciones seleccionadas con una base de datos grande.

1.2.2 MÉTODOS DE INTERPOLACIÓN

1.2.2.1 IDW (Inverse Distance Weighting)

La interpolación del punto problema se realiza asignando pesos a los datos de entorno en función inversa de la distancia que los separa (Inverse Distance Weighting, IDW), la expresión general es:

$$\hat{z}_j = \sum_{i=1}^n k_{ij} z_i \quad (1)$$

donde: \hat{z}_j : es el valor estimado en el nodo j ;

n : es el número de puntos usados en la interpolación;

k_{ij} es el peso asociado al dato i en el nodo j ;

z_i es el valor del punto i -esimo

Los valores de los pesos k_{ij} están comprendidos entre 0 y 1 y la suma de todos ellos tiene que ser igual a 1.

Para establecer una función de proporcionalidad entre el peso y la distancia, la función se puede expresar como:

$$\hat{z}_j = \frac{\sum_i z_i / d_{ij}^\beta}{\sum_i 1 / d_{ij}^\beta} \quad (2)$$

donde: $k_{ij} = \frac{1}{d_{ij}^\beta}$;

β es un exponente de ponderación que controla la forma en la que el peso disminuye con la distancia.

Este método se apoya en el concepto de continuidad espacial, con valores más parecidos para posiciones próximas que se van diferenciando conforme se incrementa la distancia. Desde el punto de vista metodológico, cada valor que tiene una correspondencia con un punto determinado influye sobre los demás de forma local, y esta influencia disminuye proporcionalmente con la distancia. Se puede decir que la influencia que tiene un punto (del que se conoce el valor de la variable), respecto al interpolado (en el que se quiere conocer el valor), es isotrópica.

Este tipo de interpolación se suele emplear cuando el número de valores que se tienen como muestra, es grande. Si la muestra no es densa o son valores muy dispersos los resultados pueden no ser del todo correctos o por lo menos no tienen por qué satisfacer las perspectivas puestas en el método.

1.2.2.2 KRIGING

Es un proceso geoestadístico avanzado que requiere en primer lugar la construcción de una función (variograma) a partir de un conjunto de puntos, en general dispersos, con valores determinados (medidos) y que permite analizar el comportamiento espacial de una propiedad o variable sobre una zona dada. Posteriormente, Kriging utiliza el variograma para ponderar los pesos asignados a los valores de la función a interpolar en los puntos de medición. Inicialmente fue estudiado por Daniel Krige (*Sudáfrica, 1919-2013*), aunque fue realmente Georges Matheron (*Francia, 1930-2000*) quien, basándose en la labor investigadora de Krige aplicada al campo de la geoestadística, formalizó la técnica conocida como Kriging. Fue Gandin

(*Gandin y Hardin, (1965)*) quien lo definió como “interpolación óptima” al aplicarlo a datos meteorológicos. Realmente está basado en modelos estadísticos que incluyen la autocorrelación, es decir que el valor z (del fenómeno estudiado) en una determinada zona influye en la probabilidad de que el fenómeno se presente, o no, en las zonas vecinas. La autocorrelación será positiva si existe mayor probabilidad de presencia en términos similares en las regiones adyacentes, y negativa en caso contrario. Los valores de z varían homogéneamente en la superficie.

Kriging no sólo genera la superficie de predicción, sino que además proporciona una medida de certeza o precisión de la respuesta. Este método presupone que entre los valores medidos hay una correlación espacial que puede ser utilizada para explicar la variación de la superficie a determinar.

Para poder hacer la predicción con el método Kriging es necesario primero descubrir las reglas de la dependencia para poder después realizar las predicciones. Para ello hay que crear un variograma (llamado también en ocasiones semivariograma) y las funciones de covarianza para calcular los valores de dependencia estadística (autocorrelación espacial), que dependen del modelo de autocorrelación (ajustar un modelo). Se dice que Kriging usa dos veces los datos, una para encontrar la correlación y otra para hacer las predicciones.

Un variograma o semivariograma es una herramienta geoestadística básica, que sirve para cuantificar los parámetros y representa la correlación espacial entre los valores. La forma más sencilla de definirlo es como la medida de los cuadrados de las diferencias entre pares de valores de la función en puntos separados una distancia. Las formulas y cálculos correspondientes a variogramas se desarrollan posteriormente.

El ajuste del modelo se hace a través de un semivariograma empírico que proporciona la información sobre la correlación espacial del conjunto de datos. El problema es que dicha información no la da para todas las direcciones y distancias posibles. Para asegurar que las predicciones tengan varianza mínima hay que ajustar un modelo, que no es más que una función

o curva continua, al semivariograma. En definitiva se realiza un análisis de regresión en el que se ajusta una línea o una curva continua a los puntos medidos. El modelo seleccionado influye en los resultados, de hecho la superficie de salida será más o menos suave dependiendo de la forma de la curva en el origen. Como cada modelo está diseñado para un tipo de distribución de datos, se emplean los descriptores **rango**, **meseta** y **nugget**, definidos más adelante, para la selección.

Una vez vista la dependencia o autocorrelación y se haya finalizado el primer uso de los datos (como se ha comentado en párrafos anteriores), se puede hacer la predicción (solución). El proceso es similar al realizado con IDW: se forman las ponderaciones, que en este caso son un poco más sofisticadas, ya que provienen de un semivariograma que se ha desarrollado en función de la naturaleza de los datos. Se realizan predicciones para cada ubicación o celda en las que se divide el área en estudio, basadas en el semivariograma y la disposición espacial de los valores medidos que sean más cercanos. Estas predicciones sirven para crear la superficie continua del fenómeno que se desea estudiar, en este caso, radiación global horizontal.

Dado un punto, u , en el que se quiere conocer el valor de una variable, el planteamiento es considerar $z(u)$, como una combinación lineal de los puntos en los que se conoce el valor de la variable, $z(u_\alpha)$. La expresión general se puede escribir:

$$z^*(u) = \sum_{\alpha=1}^N \lambda_\alpha(u) z(u_\alpha) \quad (3)$$

donde: $z^*(u)$: estimación que se busca para el punto $z(u)$;

$z(u_\alpha)$: puntos conocidos (medidas).

El criterio para establecer los valores de los pesos, λ_α , es que la estimación sea óptima, lo que implica que el estimador sea insesgado y eso se cumple con la siguiente condición:

$$\text{var}[z(u) - z^*(u)] \text{ sea mínima} \quad (4)$$

Existen diferentes versiones del método Kriging. Las más frecuentes son Krigin simple (SK), Krigin ordinario (OK) y Krigin universal (UK). Las diferencias entre ellos se exponen a continuación de forma muy breve:

- a) SK: las medias locales son relativamente constantes y de valor semejante a la media de la población conocida. La media de la población se usa para cada estimación local, en conjunto con los puntos más próximos que se han establecido como necesarios para hacer la suposición de partida.
- b) OK: las medias locales no tienen por qué ser próximas a la media de la población. Sólo se emplean los vecinos para hacer la estimación. En general es la más utilizada en problemas ambientales.
- c) UK: la media local no es conocida y varia suavemente en cada vecindad local. La componente de tendencia se modela como una combinación lineal de funciones específicas de las coordenadas de los puntos. No es más que una extensión de los dos modelos anteriores, pero en este caso dos o más variables se presentan con dependencia espacial

Kriging simple (SK)

En este caso, como hace *Cassiraga (1999)*, la expresión matemática que permite obtener el valor en el punto estimado es:

$$z_{KS}^*(u_0) - m = \sum_{\alpha=1}^n \lambda_{\alpha} [z(u_{\alpha}) - m] \quad (5)$$

donde: m se considera conocido y constante en todo el dominio de estimación, y es igual al valor esperado de la variable aleatoria z ; λ_{α} , *coeficientes de ponderación (pesos)*, vienen determinados al exigir que el estimador sea insesgado y la varianza de los errores sea mínima.

Todo eso conduce al siguiente sistema lineal de ecuaciones:

$$\sum_{\beta=1}^n \lambda_{\beta} C(u_{\alpha} - u_{\beta}) = C(u_{\alpha} - u_0), \quad \alpha = 1, \dots, n \quad (6)$$

donde: $C(h)$ es la covarianza para el vector de separación “ h ”.

En el sistema anterior se tiene a la izquierda de la igualdad las covarianzas entre los datos y a la derecha la covarianza entre los datos y el punto a estimar.

La varianza para este tipo de interpolación, Kriging simple, se puede expresar:

$$\sigma_{KS}^2(u_0) = \sigma^2 - \sum_{\alpha=1}^n \lambda_{\alpha} C(u_{\alpha} - u_0) \quad (7)$$

donde: σ^2 es la varianza correspondiente a los datos, es decir $C(0)$.

El mapa de valores σ_{KS}^2 permite cuantificar la incertidumbre en el campo estimado.

Se puede interpolar por Kriging con media variable. En este caso, el estimador requiere el conocimiento previo de la media, m , y además se supone que este valor es constante sobre todo el campo de estimación. Se puede sustituir la media, m , por alguna variable secundaria, relacionada de alguna forma con la variable principal, teniendo así una media variable localmente, $m(u)$, obtenida a partir de la variable secundaria. Una forma de calcular esta información es discretizar el rango de variación del atributo en K clases e identificar la media local, $m(u)$, para cada localización en la que se quiera estimar, con la media de valores de la variable principal localizados con los valores de la variable secundaria para cada una de sus K clases.

Kriging ordinario (OK):

En este caso la media es desconocida. Se realiza una reestimación local a partir de los datos que intervienen en cada estimación. La expresión de este tipo de interpolación es la siguiente:

$$z_{KO}^*(u_0) = \sum_{\alpha=1}^n \lambda_{\alpha} z(u_{\alpha}) \quad (8)$$

Como la suma de todos los pesos tiene que ser igual a 1 para que el estimador sea insesgado, y además debe cumplir la condición de que el error sea mínimo, el sistema de Kriging ordinario lo podemos expresar como:

$$\left\{ \begin{array}{l} \sum_{\beta=1}^n \lambda_{\beta} C(u_{\alpha} - u_{\beta}) + \mu = C(u_{\alpha} - u_0) \quad \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{\beta} = 1 \end{array} \right. \quad (9)$$

donde: μ es el parámetro de Lagrange introducido en la minimización del error, sujeto a la condición de que la suma de todos los pesos sea 1.

La expresión de la varianza de los errores se puede escribir

$$\sigma_{KO}^2(u_0) = \sigma^2 - \sum_{\alpha=1}^n \lambda_{\alpha} C(u_{\alpha} - u_0) - \mu \quad (10)$$

Kriging universal (UK)

Se le conoce también como Kriging con modelo de tendencia. Considera que la variación espacial de la media local viene dada por un polinomio de grado bajo en función de las coordenadas, pero los parámetros del mismo son desconocidos. La expresión general es la siguiente:

$$m(u) = \sum_{l=0}^L a_l f_l(u) \quad (11)$$

donde: $m(u)$ es la media en el punto de coordenadas u ;

a_l con $l = 0, \dots, L$ son coeficientes desconocidos ;

$f_l(u)$ son monomios de orden bajo, función de las coordenadas.

Un ejemplo de modelo lineal en dos dimensiones tendría la siguiente expresión:

$$m(u) = a_0 + a_1x + a_2y \quad (12)$$

y uno cuadrático sería:

$$m(u) = a_0 + a_1x + a_2y + a_3x^2 + a_4y^2 + a_5xy \quad (13)$$

El estimador Kriging universal tiene la expresión siguiente:

$$z_{KU}^*(u_0) = \sum_{\alpha=1}^n \lambda_{\alpha} z(u_{\alpha}) \quad (14)$$

La condición de “no sesgo” se obtiene en las L restricciones que deben satisfacer los pesos λ_{α} que unidas a la minimización del error da el sistema siguiente:

$$\left\{ \begin{array}{l} \sum_{\beta=1}^n \lambda_{\beta} C(u_{\alpha} - u_{\beta}) + \sum_{l=0}^L \mu_l f_l(u_{\alpha}) = c(u_{\alpha} - u_0), \quad \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{\beta} f_l(u_{\beta}) = f_l(u_0), \quad l = 0, \dots, L \end{array} \right. \quad (15)$$

La restricción para $l = 0$ es la restricción del Kriging ordinario, que corresponde a un modelo $m(x, y) = a_0$.

En este caso, la varianza de los errores por Kriging universal se puede expresar de la siguiente forma:

$$\sigma_{KU}^2(u_0) = \sigma^2 - \sum_{\alpha=1}^n \lambda_{\alpha} C(u_{\alpha} - u_0) - \sum_{l=0}^L \mu_l f_l(u_{\alpha}), \quad \alpha = 1, \dots, n \quad (16)$$

En algunos estudios aparece el CoKriging. Es un método que se emplea para los casos de regresión en los que intervienen varios atributos. Requiere un modelado para la matriz de funciones de covarianza, incluida la covarianza de cada uno de los atributos y las covarianzas cruzadas entre ellos. En este caso también se habla de CoKriging simple, ordinario y universal.

CoKriging

Esta técnica permite incorporar una o varias variables secundarias para realizar la estimación cuando no se conocen las variables primarias para todo el dominio de estimación. Es algo parecido a Kriging con deriva externa o a Kriging con media variable, en los que se emplean variables secundarias disponibles en la localización en la que se está haciendo la estimación. Por eso, la diferencia entre estos métodos y CoKriging no está en el algoritmo, sino en la posibilidad de emplear la información de las variables secundarias para que participen de forma directa en la estimación de la variable principal. Dicho de otro modo: en el caso de CoKriging la información secundaria es una información covariada. De esta forma, CoKriging hace mejor uso de la información disponible, ya que, tanto la magnitud de la variable secundaria, como sus patrones de variación espacial, se consideran en la estimación de la variable principal. Las covarianzas cruzadas permiten recoger la codependencia lineal y por lo tanto no es necesario que estén dadas en todos los puntos en los que se va a trabajar.

La expresión general del método de interpolación CoKriging se recoge en la siguiente ecuación:

$$\begin{aligned}
 z_1^* - m_1(u_0) &= \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} [z_1(u_{\alpha_1}) - m_1(u_0)] \\
 &+ \sum_{i=2}^{N_v} \sum_{\alpha_i=1}^{n_i} \lambda_{\alpha_i} [z_i(u_{\alpha_i}) - m_i(u_{\alpha_i})]
 \end{aligned}
 \tag{17}$$

donde: $z_1(u_{\alpha_1})$ $\alpha_1 = 1, \dots, n_1$ son los n_1 datos principales;

$z_i(u_{\alpha_i})$ $\alpha_i = 1, \dots, n_i, i = 2, \dots, N_v$: son los datos secundarios correspondientes a $(N_v - 1)$ variables secundarias;
 m_1 : es la media de la variable principal;
 m_i : las medias de las variables secundarias;
 λ_α : los coeficientes de ponderación de los datos que intervienen en la estimación.

CoKriging simple

Si se conoce los valores de una variable principal z_1 y de una secundaria z_2 , el estimador por CoKriging simple se puede expresar de la siguiente forma:

$$z_{CKS}^*(u_0) - m_1 = \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} [z_1(u_{\alpha_1}) - m_1] + \sum_{\alpha_2=1}^{n_2} \lambda_{\alpha_2} [z_2(u_{\alpha_2}) - m_2] \quad (18)$$

donde: z_1 es la variable principal;

z_2 es la variable secundaria;

m_1 es la media de la variable principal;

m_2 es la media de la variable secundaria;

n_1 es el número de datos primarios;

n_2 es el número de datos secundarios;

λ_α son los pesos asignados a cada variable que se obtienen resolviendo el sistema de ecuaciones lineales o también llamado sistema CoKriging simple:

$$\left\{ \begin{array}{l} \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} C_1(u_{\alpha_1} - u_{\beta_1}) + \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2} C_{12}(u_{\alpha_1} - u_{\beta_2}) = C_1(u_{\alpha_1} - u_0), \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} C_{12}(u_{\alpha_2} - u_{\beta_1}) + \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2} C_2(u_{\alpha_2} - u_{\beta_2}) = C_{12}(u_{\alpha_2} - u_0), \end{array} \right. \quad (19)$$

$\alpha_1 = 1, \dots, n_1$

$\alpha_2 = 1, \dots, n_2$

Como se puede ver en la expresión (19) están contempladas la covarianza de la variable principal, $C_1(u_{\alpha_1} - u_{\beta_1})$, la de la secundaria, $C_2(u_{\alpha_2} - u_{\beta_2})$ y la covarianza cruzada, $C_{12}(u_{\alpha_1} - u_{\beta_2})$. Aquí también es necesario conocer las

medias de las variables consideradas, igual que en el método Kriging simple. Por lo tanto la varianza de los errores se puede expresar como:

$$\sigma_{CKS}^2(u_0) = \sigma_1^2 - \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} C_1(u_{\alpha_1} - u_0) - \sum_{\alpha_2=1}^{n_2} \lambda_{\alpha_2} C_{12}(u_{\alpha_2} - u_0) \quad (20)$$

CoKriging ordinario

La expresión general de este método de interpolación es:

$$z_{CKO}^*(u_0) = \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} z_1(u_{\alpha_1}) + \sum_{\alpha_2=1}^{n_2} \lambda_{\alpha_2} z_2(u_{\alpha_2}) \quad (21)$$

Los coeficientes de ponderación (pesos) se obtienen de la resolución del siguiente sistema de ecuaciones, también llamado CoKriging ordinario:

$$\left\{ \begin{array}{l} \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} C_1(u_{\alpha_1} - u_{\beta_1}) + \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2} C_{12}(u_{\alpha_1} - u_{\beta_2}) + \mu_1 = C_1(u_{\alpha_1} - u_0), \quad \alpha_1 = 1, \dots, n_1 \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} C_{12}(u_{\alpha_1} - u_{\beta_2}) + \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2} C_2(u_{\alpha_2} - u_{\beta_2}) + \mu_2 = C_{12}(u_{\alpha_2} - u_0), \quad \alpha_2 = 1, \dots, n_2 \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} = 1 \\ \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2} = 0 \end{array} \right. \quad (22)$$

En este caso las dos últimas restricciones son suficientes pero no necesarias, para garantizar una estimación insesgada. Los coeficientes μ_1 y μ_2 son los parámetros de Lagrange introducidos en la minimización.

La expresión para la varianza de los errores, como en el caso anterior, se puede expresar de la siguiente forma:

$$\begin{aligned} \sigma_{CKO}^2(u_0) = \sigma_1^2 - \mu_1 - \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} C_1(u_{\alpha_1} - u_0) \\ - \sum_{\alpha_2=1}^{n_2} \lambda_{\alpha_2} C_{12}(u_{\alpha_2} - u_0) \end{aligned} \quad (23)$$

CoKriging colocalizado

Para la resolución de un CoKriging como el ordinario y el simple es necesario conocer las funciones de covarianza de las variables principal y secundaria y la de covarianza cruzada cada vez que se incorpora una variable secundaria. Esto supone más esfuerzo en el sentido de la modelación y del tiempo de ordenador comparándolas con Kriging, que únicamente necesita la covarianza de la variable principal. Esto hace que en muchos casos no se trabaje con CoKriging. Por otra parte, hay veces que la existencia de un dato secundario localizado en el punto en el que se va a efectuar la estimación, atenúa la influencia de los datos vecinos (*efecto pantalla*)

La alternativa a este problema es el CoKriging colocalizado. Es un coKriging en él que la única información externa que se retiene en cada punto donde se va a realizar la estimación es el dato secundario (dato colocalizado). En ese caso la expresión del estimador se puede poner:

$$z_{CKCS}^*(u_0) - m_1 = \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} [z_1(u_{\alpha_1}) - m_1] + \lambda_2 [z_2(u_0) - m_2] \quad (24)$$

y el sistema para los coeficientes de ponderación es:

$$\left\{ \begin{aligned} \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} C_1(u_{\alpha_1} - u_{\beta_1}) + \lambda_2 C_{12}(u_{\alpha_1} - u_0) &= C_1(u_{\alpha_1} - u_0), \alpha_1 = 1, \dots, n_1 \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} C_{12}(u_0 - u_{\beta_1}) + \lambda_2 C_2(0) &= C_{12}(0) \end{aligned} \right. \quad (25)$$

El sistema (25) es una simplificación del Coos simple donde el número de datos secundarios, n_2 , es igual a 1, pero a pesar de ello hay que añadirle el cálculo de las covarianzas cruzadas $C_{12}(h)$ entre variables, con h igual a la separación entre los puntos considerados. Si se emplea un modelo de Markov se facilita la tarea y por eso existe una versión de este tipo de interpolación denominado, CoKriging colocalizado bajo modelo de Markov.

En este caso el algoritmo que expresa el dato principal, $z_1(u)$, oculta la influencia de cualquier otro dato principal, $z_1(u+h)$, sobre el dato secundario colocalizado, $z_2(u)$. La expresión de este algoritmo es la siguiente:

$$\begin{aligned} E\{Z_2(u)|Z_1(u) = z, Z_1(u+h) = z'\} \\ = E\{Z_2(u)|Z_1(u) = z\} \quad \forall h, z \end{aligned} \quad (26)$$

Esta ecuación expresa que la dependencia de la variable secundaria sobre la primaria está limitada al dato primario colocalizado. Suponiendo esto, se puede demostrar que las covarianzas cruzadas, $C_{12}(h) = C_{21}(h)$, se pueden escribir en función de la covarianza de estos datos principales de la siguiente forma:

$$C_{12}(h) = \frac{C_{12}(0)}{C_1(0)} C_1(h) \quad \forall h \quad (27)$$

Con el modelo de Markov, la covarianza cruzada, $C_{12}(h)$, se puede calcular reescalando la covarianza de la variable principal, siendo el factor de escala el cociente entre $C_{12}(0)$ y $C_1(0)$.

El estimador de este modelo de interpolación, CoKriging colocalizado bajo un modelo de Markov, se puede estandarizar bajo la expresión:

$$\frac{z_{CKCS}^*(u_0) - m_1}{\sigma_1} = \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} \frac{[z_1(u_{\alpha_1}) - m_1]}{\sigma_1} + \lambda_2 \frac{[z_2(u_0) - m_2]}{\sigma_2} \quad (28)$$

Los pesos se obtienen resolviendo el sistema siguiente:

$$\begin{cases} \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} \rho_1(u_{\alpha_1} - u_{\beta_1}) + \lambda_2 \rho_{12}(u_{\alpha_1} - u_0) = \rho_1(u_{\alpha_1} - u_0), \alpha_1 = 1, \dots, n_1 \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} \rho_{12}(0) \rho_1(u_0 - u_{\beta_1}) + \lambda_2 \rho_2(0) = \rho_{12}(0) \end{cases} \quad (29)$$

Para este caso, se cumple que $\rho_{12}(0) = 0$, por lo que $\lambda_2 = 0$ y en esta circunstancia el dato secundario se ignora. Si por el contrario $\rho_{12}(0) = 1$, el sistema se transforma en un Kriging simple en el que hay $(n + 1)$ datos y cuya solución es $\lambda_2 = 1$ y $\lambda_{\beta_1} = 0$, para todo β , es decir, el estimador es igual al valor de la variable secundaria.

En resumen: el CoKriging colocalizado necesita el mismo esfuerzo de implementación que el Kriging con deriva externa, pero presenta como ventaja que permite que el dato secundario participe directamente en la estimación de la variable principal. Otra diferencia con el Kriging con deriva externa es que se construye en base a un parámetro de calibración, el coeficiente de correlación. Esto hace que cuando se elija este modelo pueda verificarse la solución a partir de los datos. Como inconveniente, este modelo requiere que la variable secundaria sea conocida en todos los puntos en los que se quiere estimar la variable principal y eso no es frecuente.

Variogramas

Todos los métodos de interpolación descritos en este apartado, que son los más frecuentes en el campo de variables geofísicas e hidrológicas, necesitan definir una función de correlación espacial denominada **variograma**.

Como se ha dicho en la introducción de Kriging, un variograma no es más que una herramienta geoestadística básica que sirve para cuantificar los parámetros y representa la correlación espacial entre los valores. La forma más sencilla de definirlo es como la media de los cuadrados de las

diferencias de los valores de la función $Z(x)$, entre pares de puntos separados una distancia. Su expresión matemática viene dada por la fórmula:

$$\gamma = \frac{1}{2N} \sum [Z(x) - Z(x + h)]^2 \quad (30)$$

donde: N es el número de parejas de puntos;

h es la distancia entre los dos puntos.

En esencia, aplicar cualquier variedad de Kriging o CoKriging implica que la media y la varianza de la función a estimar sean estacionarias. Eso se traduce en que ambos parámetros estadísticos no dependan de la posición de los puntos en análisis, es decir, solo dependan de la distancia entre ellos. Se puede decir entonces que el variograma es la varianza menos la covarianza en función de la distancia que separa los puntos. Hay autores que al utilizar el método Kriging se refieren a estas funciones denominándolas variograma y otros la denominan semivariograma, incluso en algunos casos usando indistintamente ambas denominaciones y en este trabajo también se hace así. Estas funciones se utilizan para determinar los pesos de la función de estimación.

La expresión que define un variograma o semivariograma para dos dimensiones es la siguiente:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (x_i - y_i)^2 \quad (31)$$

donde: $\gamma(h)$ es el variograma teórico;

$n(h)$ es el número de puntos x_i, y_i separados una distancia

$$h = |x_i - y_i|.$$

En este caso se emplea x_i, y_i por ser el plano euclídeo bidimensional, en general se emplea una variable aleatoria geoestadística como $Z(x)$.

En general la parte fundamental y más delicada del método Kriging es la selección y ajuste de la función objeto al semivariograma. Como ya se ha indicado la interpolación es realizar una estimación de la función $F(x)$ en un

punto X_p a partir de valores conocidos de F en un conjunto de puntos X_i dados, es decir:

$$F(x_p) = \sum_{i=1}^m W_i F(x_i) \quad (32)$$

donde: W_i es el peso o ponderación que se asigna a cada punto x_i .

Este peso se calcula teniendo en cuenta el semivariograma correspondiente, como se indica en el siguiente ejemplo en el que usamos el Kriging ordinario para describir la forma de trabajar con el semivariograma.

El objetivo es calcular los pesos W_i de la ecuación (32). Para ello resolvemos la ecuación matricial $\mathbf{AW} = \mathbf{B}$, que representa un sistema de $m + 1$ ecuaciones lineales con $m + 1$ incógnitas: W_1, W_2, \dots, W_m y λ , siendo A , W y B las matrices:

$$A = \begin{pmatrix} \gamma(h_{11}) & \gamma(h_{12}) & \dots & \gamma(h_{1m}) & 1 \\ \gamma(h_{21}) & \gamma(h_{22}) & \dots & \gamma(h_{2m}) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \gamma(h_{m1}) & \gamma(h_{m2}) & \dots & \gamma(h_{mm}) & 1 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} \quad (33)$$

$$W = \begin{pmatrix} W_1 \\ W_2 \\ \dots \\ W_m \\ \lambda \end{pmatrix} ; B = \begin{pmatrix} \gamma(h_{1p}) \\ \gamma(h_{2p}) \\ \dots \\ \gamma(h_{mp}) \\ 1 \end{pmatrix}$$

donde: $\gamma(h_{ij})$ con $i, j=1, 2, \dots, m$ son los valores del semivariograma que corresponden a la distancia $h_{i,j}$ entre los puntos x_i y x_j .

Los $\gamma(h_{i,j})$ se calculan con los datos experimentales a partir de la ecuación (33) y los $\gamma(h_{i,p})$ con la ayuda de la función analítica ajustada a los puntos del semivariograma experimental para el punto p . Hay que tener en cuenta que la solución del sistema, como se ha comentado varias veces en párrafos anteriores, debe ser insesgada, es decir la suma de todos los pesos debe ser igual a 1. Esta condición permite introducir un grado de libertad más, la

variable λ , que no es más que el multiplicador de Lagrange, que se introduce para minimizar el error de la estimación, cuestión fundamental en este procedimiento de interpolación.

El valor buscado en el punto x_p se calcula con los valores conocidos de F y la ecuación (30). La varianza de la estimación se puede poner como:

$$S_p^2 = W^T B \quad (34)$$

donde: W^T es la traspuesta de W .

Si los errores de estimación siguen una distribución normal, y de acuerdo a las características inherentes a este tipo de distribución, la probabilidad de que el valor verdadero esté en el intervalo $F(x_p) \pm S_p$ es del 68 %, mientras que la probabilidad de que el valor verdadero sea $F(x_p) \pm 2S_p$ es del 95 % siendo S_p la desviación estándar.

El variograma que no es sino una gráfica que describe la variación, en promedio, de la varianza entre dos muestras a una distancia determinada, identifica y modela la correlación espacial y caracteriza las tendencias, la variabilidad y las anisotropías que puedan tener los datos a cualquier distancia, corta o larga. Los elementos que caracterizan cualquier variograma son, la varianza, el alcance, la meseta y el efecto pepita. A continuación se da una breve definición de cada una de ellas.

Varianza: se calcula como la diferencia entre dos puntos para diferentes distancias, también llamadas “Lag”. Se puede decir que es el promedio del grado de diferencia entre pares de puntos.

Alcance (también se llama “Range”): es el punto sobre el eje “X” en el que la varianza se hace plana. Es la distancia a la cual el variograma se estabiliza o distancias a las cuales los valores dejan de estar correlacionados.

Meseta (en alguna bibliografía denominado “Sill”): es la máxima diferencia, es decir mide la varianza total del sistema. Se mide en el eje “Y” y es el punto en el que la covarianza se vuelve plana horizontalmente. Es el valor constante que toma el variograma en distancias mayores al rango

Efecto pepita (también llamado “nugget effect”): es la discontinuidad en el origen. Indica el grado de disimilaridad en la distancia cero. Generalmente aparece debido a errores en las medidas y también puede representar heterogeneidad a pequeña escala.

En la práctica existen diferentes modelos de semivariogramas y la elección para cada trabajo depende del tipo de datos y de la información disponible de la variable a interpolar. A continuación se recogen los modelos más frecuentes que aparecen en la bibliografía.¹

* Efecto pepita:

$$\gamma(h) = \begin{cases} C_0 & \text{si } h > 0 \\ 0 & \text{si } h < 0 \end{cases} \quad (35)$$

* Lineal con meseta:

$$\gamma(h) = \begin{cases} C_0 + C_s \left(\frac{h}{a_l} \right) & \text{si } 0 \leq h \leq a_l \\ C_0 + C_s & \text{si } h > a_l \end{cases} \quad (36)$$

* Esférico:

$$\gamma(h) = \begin{cases} 0 & \text{si } h = 0 \\ C_0 + C_s \left[\frac{3}{2} \frac{h}{a_s} - \frac{1}{2} \left(\frac{h}{a_s} \right)^3 \right] & \text{si } 0 < h \leq a_s \\ C_0 + C_s & \text{si } h > a_s \end{cases} \quad (37)$$

* Exponencial:

$$\gamma(h) = \begin{cases} C_0 + C_s \left(1 - \exp(-3h/a_l) \right) & \text{si } h > 0 \\ 0 & \text{si } h = 0 \end{cases} \quad (38)$$

¹ No se ha hecho más que una simple descripción de las expresiones que caracterizan a cada uno de ellos, no se cree necesario hacer más comentarios ya que la utilización en este trabajo es a través de un sistema en el que ya está definida la herramienta con sus correspondientes modelos, y selecciona el más apropiado en función del tipo de datos y su dispersión.

* Gaussiano:

$$\gamma(h) = \begin{cases} C_0 + C_s \left(1 - \exp(-3h^2/a_l^2) \right) & \text{si } h > 0 \\ 0 & \text{si } h = 0 \end{cases} \quad (39)$$

* Circular:

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 + C_s \left(1 - \frac{2}{\pi} \cos^{-1} \left(\frac{h}{a_c} \right) \right) - \frac{2h}{\pi a_c} \sqrt{\left[1 - \left(\frac{h}{a_c} \right)^2 \right]} & 0 < h \leq a_c \\ C_0 + C_s & h > a_c \end{cases} \quad (40)$$

* Efecto agujero:

$$\gamma(h) = \begin{cases} C_0 + C_s \left(1 - \frac{\text{sen}(h)}{h} \right) & h > 0 \\ 0 & h = 0 \end{cases} \quad (41)$$

1.2.3 SPLINES

El origen de Splin se remonta a la época en que los delineantes usaban una lámina flexible de madera o plástico (“*curvígrafo*”) para hacer pasar una curva por un conjunto de puntos. Este ingenio permite trazar curvas suaves pasando por un conjunto de puntos. La curva se puede expresar por medio de una función matemática, generalmente un polinomio. El problema se presenta en establecer el grado del polinomio que representa la variación de los puntos de forma suave y estable. Si se utilizan polinomios de grado elevado estas funciones pueden no responder al fenómeno con suficiente aproximación, ya que presentan gran número de extremos y/o inflexiones y su cálculo es muy complicado.

Si la función es de una sola variable su gráfica es una curva y si es de dos variables la gráfica será una superficie en el espacio.

La utilización de la interpolación por polinomios, ha demostrado que, si se divide el espacio de trabajo en intervalos pequeños, se pueden utilizar polinomios de menor grado para resolver los problemas de aproximación. De esta forma una función definida a trozos se adapta mejor al fenómeno estudiado. En resumen lo que se hace es utilizar varios polinomios, cada uno en un intervalo, y unirlos bajo condiciones de continuidad para definir la función de interpolación, en lugar de considerar un único polinomio en todo el intervalo. Desde el punto de vista matemático se puede definir como una función continua, diferenciable a trozos, definida en cada subintervalo por medio de polinomios, generalmente de pequeño grado, porque de esta forma se evitan las oscilaciones que se pueden generar en el tratamiento de los datos con polinomios de grado elevado.

Se dice que un Splin es de grado n , cuando los polinomios definidos en cada uno de los intervalos es de grado menor o igual a n y la función $S(x)$ definida es de clase C^{n-1} en el intervalo dado. La práctica ha demostrado que los Splines de tercer grado son los más adecuados.

Para funciones de una variable, planteado el problema de interpolar una función f , de la que se conoce su valor en $n+1$ puntos ordenados: $\{x_0, x_1, \dots, x_n\}$, de forma que $f(x_i) = y_i \quad \forall i = 0, 1, \dots, n$, se buscan las funciones Splin definidas en el intervalo $[a = x_0, x_n = b]$, que coincidan con la f en los $n+1$ puntos es decir $S(x_i) = f(x_i) = y_i \quad \forall i = 0, 1, \dots, n$.

A continuación se describen brevemente las funciones Splin de grado 0, las de grado 1 y grado 2. Se definen genéricamente Splines de grado m y se desarrollan con detalle los de grado 3.

1.2.3.1 Función Splin de grado 0

Las funciones Splin de grado 0 son constantes a trozos, ya que según la definición anterior de grado de un Splin, los polinomios correspondientes,

definidos en cada intervalo en el que se divide la zona de estudio, son grado 0. Se pueden representar de forma explícita como sigue:

$$S(x) = \begin{cases} S_0(x) = C_0 & x \in (x_0, x_1) \\ S_1(x) = C_1 & x \in (x_1, x_2) \\ \dots \dots \dots \dots \dots \dots \\ S_{n-1}(x) = C_{n-1} & x \in (x_{n-1}, x_n) \\ S(x_i) = f(x_i) = y_i & \forall i = 0, 1, \dots, n \end{cases} \quad (42)$$

Aunque tiene sentido desde un punto de vista formal hablar de Splines de grado 0, no tiene sentido considerarlos como herramienta de interpolación, ya que no cumple la condición de continuidad, salvo que la función f fuera constante en los puntos conocidos.

1.2.3.2 Función Splin de grado 1

En este caso los Splines vienen definidos mediante funciones polinómicas de grado 1 a lo sumo, es decir su gráfica está formada por segmentos de rectas que unen cada par de puntos. Su expresión es:

$$S(x) = \begin{cases} S_0(x) = a_0x + b_0 & x \in (x_0, x_1) \\ S_1(x) = a_1x + b_1 & x \in (x_1, x_2) \\ \dots \\ S_{n-1}(x) = a_{n-1}x + b_{n-1} & x \in (x_{n-1}, x_n) \end{cases} \quad (43)$$

además $S_{j-1}(x_j) = S_j(x_j)$ para $j = 1, \dots, n$
y $S_j(x_j) = y_j$ para $j = 0, 1, \dots, n$

Por lo que $S(x)$ es continua en todo el intervalo y derivable a trozos. Las funciones en cada subintervalo son de la forma: $ax + b$. En este caso si hay $n+1$ puntos, se deben resolver $2n+2$ ecuaciones en $2n+2$ incógnitas.

Otra forma de escribir el Splin utilizando diferencias divididas es:

$$S(x) = \begin{cases} y_0 + f[x_1, x_2](x - x_0) & \text{si } x \in [x_0, x_1] \\ y_1 + f[x_2, x_1](x - x_1) & \text{si } x \in [x_1, x_2] \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ y_{n-1} + f[x_n, x_{n-1}](x - x_{n-1}) & \text{si } x \in [x_{n-1}, x_n] \end{cases} \quad (44)$$

Siendo $f[x_i, x_j]$ la diferencia dividida de Newton:

$$f[x_i, x_j] = \frac{f(x_j) - f(x_i)}{x_j - x_i}$$

1.2.3.3 Función Splin de grado 2

Los Splines de grado 2 unen cada par de puntos con polinomios de orden 2, es decir funciones de la forma: $S_i(x) = a_i x^2 + b_i x + c_i$. En este caso si hay $n+1$ puntos, se define $S(x)$ mediante n polinomios de grado 2:

$$S(x) = \begin{cases} S_0(x) = a_0 x^2 + b_0 x + c_0 & x \in (x_0, x_1) \\ S_1(x) = a_1 x^2 + b_1 x + c_1 & \text{si } x \in (x_1, x_2) \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ S_{n-1}(x) = a_{n-1} x^{n-1} + b_{n-1} x + c_{n-1} & x \in (x_{n-1}, x_n) \end{cases} \quad (45)$$

Con lo que es necesario encontrar $3n$ incógnitas. Para determinarlas se impone la condición de que $S(x)$ sea continua en el intervalo $[a = x_0, x_n = b]$ y tome el mismo valor que la función f en los $n+1$ puntos $\{x_0, x_1, \dots \dots \dots x_n\}$. Estas condiciones implican resolver el siguiente sistema:

$$\begin{cases} S_0(x_0) = f(x_0) \\ S_0(x_1) = S_1(x_1) = f(x_1) \\ S_1(x_2) = S_2(x_2) = f(x_2) \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ S_{n-3}(x_{n-2}) = S_{n-2}(x_{n-2}) = f(x_{n-2}) \\ S_{n-2}(x_{n-1}) = S_{n-1}(x_{n-1}) = f(x_{n-1}) \\ S_{n-1}(x_n) = f(x_n) \end{cases} \quad (46)$$

Con esto se tienen $2n$ ecuaciones.

Para tener las $3n$ ecuaciones necesarias para resolver el sistema se aplica la condición de que la derivada $S'(x)$ sea continua en el intervalo (a, b) , lo que implica que las derivadas laterales sean iguales en los nodos interiores. La expresión de $S'(x)$ es:

$$S'(x) = \begin{cases} S'_0(x) = 2a_0x + b_0 & \text{si } x \in (x_0, x_1) \\ S'_1(x) = 2a_1x + b_1 & \text{si } x \in (x_1, x_2) \\ \dots \dots \dots \dots \dots \dots \dots \\ S'_{n-1}(x) = 2a_{n-1}x + b_{n-1} & \text{si } x \in (x_{n-1}, x_n) \end{cases} \quad (47)$$

La continuidad de $S'(x)$ supone $n-1$ condiciones adicionales:

$$\begin{cases} S'_0(x_1) = S'_1(x_1) \\ \dots \dots \dots \dots \dots \dots \dots \\ S'_{n-2}(x_{n-1}) = S'_{n-1}(x_{n-1}) \end{cases} \quad (48)$$

por lo que falta una condición supletoria, que hay que determinar. Se puede elegir $S'(a) = f'(a)$ o $S'(b) = f'(b)$ siempre que se conozca el valor de f' en uno de los extremos del intervalo, lo que no es habitual si no se conoce la expresión de f . En este caso se impone que $S'(a) = 0$ o $S'(b) = 0$.

Elegida la condición adicional, la solución se puede conseguir mediante cálculo matricial. Una vez resuelto el sistema de ecuaciones, los coeficientes se sustituyen en cada polinomio y se obtiene el Splin de grado 2.

1.2.3.4. Función Splin de grado 3. Splines cúbicos

De forma general para cualquier grado se definen Splines de grado m de la siguiente manera:

Dado un conjunto de puntos: $x_0, x_1, \dots \dots \dots x_n$, se establece la partición de dichos puntos de la siguiente forma:

$$\Delta = \{a = x_0 < x_1 < \dots \dots \dots < x_n = b\}$$

se suele llamar $h_i = x_{i+1} - x_i \quad i = 0, 1, \dots \dots, n - 1$ (49)

Se definen los polinomios a lo sumo de grado m , $P_i(x)$:

$$P_i(x) = a_{i0} + a_{i1}x + \dots + a_{i, m-1}x^{m-1} + a_{i, m}x^m$$

Se define el Splin $S(x)$ de grado m en $[a, b]$ referido a la partición Δ como:

$$\begin{cases} S(x) \in C^{m-1}([a, b]) \\ S(x)|_{[x_i, x_{i+1}]} = P_i(x) \quad i = 0, 1, \dots, n-1 \end{cases} \quad (50)$$

Siendo las condiciones de interpolación:

$$f(x_0) = y_0, \dots, f(x_n) = y_n \quad (51)$$

Observaciones generales:

- a) $S'(x)$: es un Splin de orden $m-1$ en $[a, b]$ referido a la partición Δ
- b) número de parámetros $(m+1)n$
- c) número de condiciones de continuidad $m(n-1)$
- d) dimensión esperable $n+m$
- e) condiciones de interpolación $n+1$
- f) hay que añadir $m-1$ condiciones adicionales

Se va a estudiar en particular el caso más empleado que es el Splin cúbico, es decir $m = 3$. Se denomina también *Splin por curvas suaves*. $S(x)$ es un Splin cúbico referido a la partición Δ si:

$$\begin{cases} S(x) \in C^2([a, b]) \\ S(x)|_{[x_i, x_{i+1}]} = P_i(x) \quad i = 0, 1, \dots, n-1 \end{cases} \quad (52)$$

siendo $P_i(x) = a_{i0} + a_{i1}x + a_{i2}x^2 + a_{i3}x^3$

En este caso:

- a) $S'(x)$: es un Splin de orden 2 en $[a, b]$ referido a la partición Δ
- b) $S''(x)$: es un Splin de orden 1 en $[a, b]$ referido a la partición Δ
- c) número de parámetros $4n$
- d) número de condiciones de continuidad $3(n-1)$

- e) dimensión esperable $n+3$
- f) condiciones de interpolación $n+1$
- g) hay que añadir dos condiciones. Hay tres opciones de elección, cada una de las cuales da nombre al Splin y son:
 - a. $S''(a) = S''(b) = 0$; Splin natural
 - b. $S'(a) = f'(a)$ y $S'(b) = f'(b)$; Splin de frontera sujeta
 - c. Si $f(a) = f(b)$ se imponen $S'(a) = S'(b)$ y $S''(a) = S''(b)$; Splin periódico

Se puede abordar la resolución del problema planteado directamente, buscando los coeficientes de los polinomios definidos en cada subintervalo e imponiendo las condiciones exigidas, o aplicar que $S''(x)$ es un Splin de grado 1 y obtener $S(x)$ mediante integración.

Como $S''(x)$ es un Splin de orden 1 en un intervalo $[a, b]$ referido a la partición Δ , se conoce su expresión en función de los valores que toma en los nodos:

$$S''(x)|_{[x_i, x_{i+1}]} = -M_i \frac{x - x_{i+1}}{h_i} + M_{i+1} \frac{x - x_i}{h_i} \quad (53)$$

$$i = 0, 1, \dots, n - 1$$

donde: $M_i = S''(x_i)$ $i = 0, 1, \dots, n$ son los coeficientes de interpolación lineal, son valores desconocidos que hay que determinar.

La expresión (53) corresponde a la forma de Lagrange de un polinomio interpolador de grado 1 en un intervalo $[x_i, x_{i+1}]$.²

Se integra dos veces en cada subintervalo añadiendo las constantes de integración:

$$S'(x)|_{[x_i, x_{i+1}]} = -M_i \frac{(x - x_{i+1})^2}{2h_i} + M_{i+1} \frac{(x - x_i)^2}{2h_i} + A_i \quad (54)$$

$$i = 0, \dots, n - 1$$

² El polinomio lineal que pasa por los puntos $(x_0, g(x_0)), (x_1, g(x_1))$ es de la forma :
 $P(x) = g(x_0) \frac{x-x_1}{x_0-x_1} + g(x_1) \frac{x-x_0}{x_1-x_0} = -g(x_0) \frac{x-x_1}{x_1-x_0} + g(x_1) \frac{x-x_0}{x_1-x_0}$ (Fórmula de Lagrange para el polinomio interpolador de grado 1)

$$S(x)|_{[x_i, x_{i+1}]} = -M_i \frac{(x - x_{i+1})^3}{6h_i} + M_{i+1} \frac{(x - x_i)^3}{6h_i} + A_i(x - x_i) + B_i, \quad i = 0, 1, \dots, n-1, \quad (55)$$

Las condiciones de interpolación son:

$$S(x_i^+) = y_i \quad i = 0, 1, \dots, n-1 \quad \text{y} \quad S(x_i^-) = y_i \quad i = 1, \dots, n \quad (56)$$

$$S(x)|_{[x_i, x_{i+1}]} = -M_i \frac{(x - x_{i+1})^3}{6h_i} + M_{i+1} \frac{(x - x_i)^3}{6h_i} + A_i(x - x_i) + B_i, \quad i = 0, \dots, n-1 \quad (57)$$

$$S(x)|_{[x_{i-1}, x_i]} = -M_i \frac{(x - x_i)^3}{6h_{i-1}} + M_i \frac{(x - x_{i-1})^3}{6h_{i-1}} + A_{i-1}(x - x_{i-1}) + B_{i-1}, \quad i = 1, \dots, n \quad (58)$$

$$B_i = y_i - \frac{M_i h_i^2}{6} \quad i = 0, 1, \dots, n-1 \quad (59)$$

$$A_{i-1} = \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{M_i - M_{i-1}}{6} h_{i-1} \quad i = 1, \dots, n-1 \quad (60)$$

$$A_i = \frac{y_{i+1} - y_i}{h_i} - \frac{M_{i+1} - M_i}{6} h_i \quad i = 0, 1, \dots, n-1 \quad (61)$$

Como se impone la continuidad en los nodos interiores tenemos:

$$S(x_i^+) = S(x_i^-) \quad i = 1, \dots, n-1 \quad (62)$$

Por construcción, son polinomios de grado menor o igual a tres en cada subintervalo:

en $[x_i, x_{i+1}]$:

$$S'(x)|_{[x_i, x_{i+1}]} = -M_i \frac{(x - x_{i+1})^2}{2h_i} + M_{i+1} \frac{(x - x_i)^2}{2h_i} + \frac{y_{i+1} - y_i}{h_i} - \frac{M_{i+1} - M_i}{6} h_i, \quad i = 0, 1, \dots, n-1 \quad (63)$$

en $[x_{i-1}, x_i]$:

$$S'(x)|_{[x_{i-1}, x_i]} = -M_{i-1} \frac{(x - x_i)^2}{2h_{i-1}} + M_i \frac{(x - x_{i-1})^2}{2h_{i-1}} + \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{M_i - M_{i-1}}{6} h_{i-1}, \quad i = 1, \dots, n \quad (64)$$

y como las derivadas también deben ser continuas:

$$S'(x_i^+) = S'(x_i^-) \quad i = 1, \dots, n-1 \quad (65)$$

Aplicando estas condiciones se obtiene:

$$\begin{aligned} -M_i \frac{h_i}{2} + \frac{y_{i+1} - y_i}{h_i} - \frac{M_{i+1} - M_i}{6} h_i &= \\ &= M_i \frac{h_{i-1}}{2} + \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{M_i - M_{i-1}}{6} h_{i-1} \end{aligned} \quad (66)$$

expresión que reordenada tiene la forma:

$$\begin{aligned} h_{i-1}M_{i-1} + 2(h_{i-1} + h_i)M_i + h_iM_{i+1} &= \\ = 6 \left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}} \right) & \quad i = 1, \dots, n-1 \end{aligned} \quad (66 \text{ bis})$$

Se tienen $n-1$ ecuaciones, y $n+1$ incógnitas: $M_i \quad i = 0, \dots, n$. Se deben añadir dos condiciones más.

Si se buscan Splines cúbicos naturales, las condiciones adicionales son $S''(a) = s''(b) = 0$; luego $M_0 = M_n = 0$.

Aplicando estas condiciones el sistema a resolver en forma matricial es:

$$\begin{pmatrix} 2(h_0 + h_1) & h_1 & 0 & \dots & 0 \\ h_1 & 2(h_1 + h_2) & h_2 & \dots & 0 \\ 0 & h_2 & 2(h_2 + h_3) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2(h_{n-2} + h_{n-1}) \end{pmatrix} \times$$

$$\times \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_{n-1} \end{pmatrix} = \begin{pmatrix} 6 \left(\frac{y_2 - y_1}{h_1} - \frac{y_1 - y_0}{h_0} \right) \\ 6 \left(\frac{y_3 - y_2}{h_2} - \frac{y_2 - y_1}{h_1} \right) \\ \vdots \\ 6 \left(\frac{y_n - y_{n-1}}{h_{n-1}} - \frac{y_{n-1} - y_{n-2}}{h_{n-2}} \right) \end{pmatrix} \quad (67)$$

Es una matriz tridiagonal o de diagonal dominante, por lo que existe una única solución. Si se cambian las condiciones adicionales, la solución es muy similar.

El Splin cúbico en cada intervalo $[x_i, x_{i+1}]$ se puede escribir de la siguiente forma:

$$\alpha_i + \beta_i(x - x_i) + \gamma_i(x - x_i)^2 + \delta_i(x - x_i)^3 \quad (68)$$

$$\alpha_i = S(x_i^+) = y_i \quad (69)$$

$$\beta_i = S'(x_i^+) = \frac{y_{i+1} - y_i}{h_i} - \frac{M_i h_i}{3} - \frac{M_{i-1} h_i}{6} \quad (70)$$

$$\gamma_i = \frac{S''(x_i^+)}{2} = \frac{M_i}{2} \quad (71)$$

$$\delta_i = \frac{S'''(x_i^+)}{6} = \frac{M_{i+1} - M_i}{6h_i} \quad (72)$$

Este método de interpolación por Splines cúbicos presenta una buena flexibilidad y velocidad de cálculo, requiere menos memoria y es más estable que por ejemplo el método de Lagrange. Como se ha visto en apartados anteriores son varios los métodos que permiten trabajar con Splines cúbicos

pero con diferente metodología y aplicaciones. A continuación se va a hacer una breve descripción de alguno de estos métodos.

Si se trabaja con Splines cúbicos naturales se requiere que dos secciones curvas adyacentes tengan la primera y segunda derivada igual en el extremo del intervalo común, es decir se exige continuidad C^2 . Estos Splines presentan una desventaja y es que no permiten el control local de la curva: la modificación de un punto de control, afecta a la curva entera, por lo que hay que rehacer los cálculos.

Los **Splines de Hermite** poseen tangente específica en cada punto de control, es decir permiten el control local. Estos suelen ser muy útiles en casos en los que no es difícil aproximar las pendientes de la curva, y se suele hacer en función de las posiciones de los puntos de control.

Los **Splines cardinales** también usan tangentes en los puntos frontera de cada intervalo, pero en este caso no hay que dar el valor de la pendiente, como en el de Hermite, sino que se calcula con los dos puntos adyacentes como coordenadas. Aquí cada sección está definida por cuatro puntos de control consecutivos, los dos centrales son los extremos del intervalo y los otros dos son los que se emplean para calcular la pendiente de los extremos.

Los **B-splines** trabajan de forma diferente. Están definidos a través de segmentos curvos, sobre cada uno de los cuales actúan cuatro puntos de control que son compartidos por los segmentos adyacentes, pero no tienen influencia en los segmentos que no los comparten. La característica fundamental de este modelo es que pasa necesariamente por los extremos del espacio de trabajo, es decir por el primero y el último punto de control.

1.2.3.5 Splines bidimensionales

Si se necesita encontrar una superficie que pase por un número determinado de puntos del espacio real tridimensional, de forma que la superficie no presente aristas ni picos, puede parecer que el problema planteado es análogo al anterior, pero la realidad es que la dificultad aumenta en gran

medida. En este caso es necesario definir funciones de dos variables cuya representación sea la superficie indicada. Se puede intentar buscar una función polinómica definida en toda la región pero se presenta el mismo problema que cuando se quiere definir una función polinómica que pase por los puntos de un plano: salvo en contadas ocasiones sería necesario emplear polinomios de grado tan alto que su comportamiento no es adecuado al problema planteado.

Con similar planteamiento al presentado, al definir Splines de una variable, se debe dividir la región del plano en partes suficientemente pequeñas para que, en ellas, se pueda buscar localmente una función suficientemente cómoda y que pase por los puntos fijados con anterioridad. Además las funciones definidas localmente deben representar una superficie que globalmente no presente rupturas y que varíe con suficiente suavidad para no presentar aristas ni picos.

Para resolver este problema, diversos autores han recurrido a funciones Splin de dos o más dimensiones. *Franke, (1982)*, introduce el problema de hallar funciones de interpolación en dos variables independientes, para conjuntos de puntos dispersos, a partir de un conjunto de puntos

$$(x_k, y_k, f_k), \quad k = 1, \dots, N, \text{ con distintos } (x_k, y_k) \quad (73)$$

buscando construir una función $F(x, y)$ tal que $F(x_k, y_k) = f_k, k = 1, \dots, N$.

En primer lugar define partición de la unidad en R^2 como un conjunto de funciones

$$w_i(x, y), \quad i = 1, \dots, m \quad (74)$$

tal que cada una cumpla:

$$w_i(x, y) \geq 0 \text{ y } \sum_{i=1}^m w_i(x, y) \equiv 1 \quad (75)$$

A estas funciones les llama *funciones peso* y define el soporte de w_i como el conjunto S_i

$$S_i = \text{Clausura}\{(x, y) | w_i(x, y) \neq 0\} \quad (76)$$

siendo $I_i = \{k | (x_k, y_k) \in \text{Interior}(S_i)\}$

Denomina aproximaciones locales a las funciones Q_i tales que si:

$$k \in I_i, \quad Q_i(x_k, y_k) = f_k \quad (77)$$

Define:

$$F(x, y) = \sum_{i=1}^m w_i(x, y) Q_i(x, y) \quad (78)$$

Con esa definición se comprueba que la función $F(x, y)$ tiene las siguientes propiedades:

- $F(x_k, y_k) = f_k, k = 1, \dots, N$
- $F(x, y)$ tiene al menos las mismas propiedades de continuidad y derivabilidad que w_i y Q_i , es decir que si w_i y $Q_i \forall i = 1, \dots, m$ tienen derivadas primeras continuas, también las tiene $F(x, y)$
- Dependencia local de los datos: Sea (x, y) fijo y sea

$$J_{x,y} = \{i | w_i(x, y) \neq 0\}$$

entonces $F(x, y)$ depende solo de los puntos (x_k, y_k, f_k) para los cuales $k \in \left(\bigcup_{i \in J_{x,y}} I_i \right) \cup \{i | \text{algún } Q_j, j \in J_{x,y} \text{ depende de } (x_k, y_k, f_k)\}$.

En particular $F(x, y) = \sum_{i \in J_{x,y}} w_i(x, y) Q_i(x, y)$

Para elegir las funciones peso, considera por comodidad, regiones soporte que sean rectángulos y lo realiza de la siguiente forma:

El polinomio cúbico (polinomio de Hermite):

$$H_3(s) = 1 - 3s^2 + 2s^3 \quad (79)$$

satisface las condiciones: $H_3(0) = 1$, $H_3(1) = H'_3(0) = H'_3(1) = 0$ y su gráfica en el intervalo $[0, 1]$ es la que aparece en la figura 1.

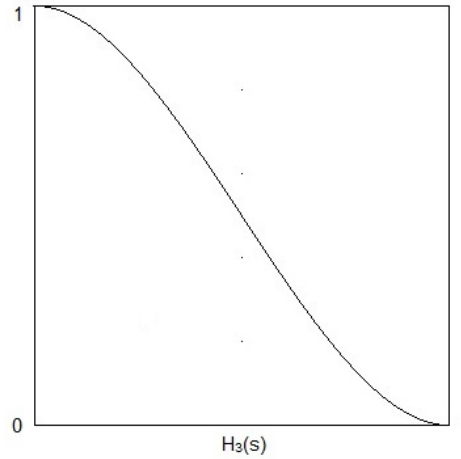


Figura 1.1. Polinomio de Hermite en el intervalo $[0, 1]$

Sean n_x y n_y números naturales distintos de 0 y un número finito de valores dados:

$$\begin{aligned} \widetilde{x}_0 < \widetilde{x}_1 < \widetilde{x}_2 < \dots < \widetilde{x}_{n_x} < \widetilde{x}_{n_x+1} \\ & \text{y} \\ \widetilde{y}_0 < \widetilde{y}_1 < \widetilde{y}_2 < \dots < \widetilde{y}_{n_y} < \widetilde{y}_{n_y+1} \end{aligned} \quad (80)$$

Para cada $i = 1, 2, \dots, n_x$ y cada $j = 1, 2, \dots, n_y$ se considera el rectángulo

$$r_{i,j} = [\widetilde{x}_{i-1}, \widetilde{x}_{i+1}] \times [\widetilde{y}_{j-1}, \widetilde{y}_{j+1}] \quad (81)$$

Define la cúbica a trozos con derivadas primeras continuas que son nulas salvo en dos intervalos adyacentes y que satisfacen

$$\begin{cases} v_i(x_j) = \delta_{ij} & i, j = 1, 2, \dots, n_x \\ u_j(y_i) = \delta_{ij} & i, j = 1, 2, \dots, n_y \end{cases} \quad (82)$$

En particular se definen:

$$\begin{aligned}
 v_1(x) &= \begin{cases} 1 & x < \widetilde{x}_1 \\ H_3\left(\frac{x - \widetilde{x}_1}{\widetilde{x}_2 - \widetilde{x}_1}\right) & \widetilde{x}_1 \leq x < \widetilde{x}_2 \\ 0 & x \geq \widetilde{x}_2 \end{cases} \\
 v_i(x) &= \begin{cases} 0 & x < \widetilde{x}_{i-1} \\ 1 - v_{i-1}(x) & \widetilde{x}_{i-1} \leq x < \widetilde{x}_i \\ H_3\left(\frac{x - \widetilde{x}_1}{\widetilde{x}_2 - \widetilde{x}_1}\right) & \widetilde{x}_i \leq x < \widetilde{x}_{i+1} \\ 0 & x \geq \widetilde{x}_{i+1} \end{cases} \quad \forall i = 2, \dots, n_x - 1 \\
 v_{n_x}(x) &= \begin{cases} 0 & x < \widetilde{x}_{n_x-1} \\ 1 - v_{n_x-1}(x) & \widetilde{x}_{n_x-1} \leq x < \widetilde{x}_{n_x} \\ 1 & x \geq \widetilde{x}_{n_x} \end{cases}
 \end{aligned} \tag{83}$$

De forma análoga se define $u_i(y)$.

Si consideramos

$$w_{i,j}(x, y) = v_i(x)u_j(y), \quad i = 1, 2, \dots, n_x, \quad j = 1, 2, \dots, n_y \tag{84}$$

el soporte de $w_{i,j}$ es el rectángulo $r_{i,j}$ salvo para $i = 1$, $i = n_x$, $j = 1$, $j = n_y$ en el que el soporte se extiende a ∞ en una o las dos variables. Las funciones $w_{i,j}$ forman una partición de la unidad en R^2 . Los valores \widetilde{x}_i y \widetilde{y}_i generan una cuadrícula en el plano (un grid), por lo que les denomina "grid values". La elección de los puntos \widetilde{x}_i y \widetilde{y}_i dependen de los datos de partida. Para determinar la función interpolante Splin de placa fina (Splín thin plate) como solución al problema de una placa fina a la que se le fuerza a pasar por ciertos puntos toma:

$$Q(x, y) = \sum_{k \in I} A_k d_k^2 \log d_k + a + bx + cy \tag{85}$$

siendo $d_k^2 = (x - x_k)^2 + (y - y_k)^2$ el cuadrado de la distancia entre los puntos (x, y) y (x_k, y_k) y siendo el conjunto de índices $I = \{k | Q(x_k, y_k) = f_k\}$

Los coeficientes A_k, a, b , y c se determinan mediante el sistema de ecuaciones lineales:

$$\begin{aligned}
 \sum_{k \in I} A_k d_k^2 \log d_k + a + bx + cy \Big|_{(x,y)=(x_i,y_i)} &= f_i \quad i \in I \\
 \sum_{k \in I} A_k &= 0 \\
 \sum_{k \in I} A_k x_k &= 0 \\
 \sum_{k \in I} A_k y_k &= 0
 \end{aligned} \tag{86}$$

En trabajos posteriores, *Franke, (1985)*, estudia los Splines de placa fina con tensión incidiendo en los problemas físicos que llevan a plantear estas funciones como solución a los mismos. Un enfoque semejante se realiza en *Mitáš y Mitášová, (1988)* incluyendo en su estudio los modelos físicos para la interpolación suave.

Teniendo en cuenta que, en el algoritmo para la interpolación para los Splines de placa delgada aparece la distancia euclídea entre los nodos y el punto en que se busca el valor, numerosos autores como *Beatson et al., (2014)* inician su estudio definiendo funciones de base radial y estudiando después los Splines como un caso particular de ellas.

1.2.4 MÉTODOS DE INTERPOLACIÓN EN ARCGIS

Se ha optado en este trabajo resolver el problema planteado utilizando técnicas de interpolación mediante el software **ArcGis 10.3**. En este sistema de trabajo se define la técnica de interpolación de la siguiente forma: “*la interpolación busca valores para las celdas de un ráster a partir de una cantidad limitada de puntos de datos de la muestra*”. En este apartado se comentarán las principales características de cada uno de los métodos de interpolación anteriormente descrito cuando se implementan en ArcGis. Para cada método se indica su funcionamiento, forma de utilización y principales parámetros de control.

Como se ha indicado a lo largo del texto, La interpolación es un procedimiento matemático que supone que los objetos que están distribuidos

espacialmente tienen algún tipo de correlación, o que, por el mero hecho de estar próximos, tienden a tener características similares. Es decir, los puntos próximos tienen más posibilidades de ser parecidos. El objetivo de la interpolación es crear una superficie a partir de los puntos medidos. ArcGis emplea una rejilla rectangular de píxeles o puntos, **matriz**, que se puede visualizar y que denomina **ráster**.

1.2.4.1 IDW en ArcGis

a) Definición: en este caso ArcGis 10.3, estima los valores de la rejilla calculando los promedios de los valores de los puntos de la muestra que están en las celdas vecinas a la de estudio. Cuanto más cerca está más influencia o peso tiene en el proceso de cálculo. Interpola una superficie ráster con el inverso de la distancia ponderada.

b) Uso: *Watson y Philip, (1985)*, asegura que el promedio no puede ser mayor que el valor máximo ni menor que el más pequeño, es decir no se pueden crear crestas o valles si los extremos no se usan para el muestreo. El IDW da buenos resultados cuando la muestra es lo suficientemente grande en el espacio de trabajo, si no es suficientemente densa la superficie, no representa correctamente la interpolación. Hay que tener en cuenta que la influencia de los puntos sobre un valor interpolado es isotrópica. Es bueno preparar los datos quitando puntos coincidentes para que no haya problemas de interpretación, ya que esto puede causar que algunos valores de entrada tengan distintos valores a los que se esperaban en algunas ubicaciones del ráster. Se puede emplear una barrera para delimitar una zona que no entra en la superficie interpolada. Por ejemplo en una costa, en una montaña o en una interrupción del terreno. Una barrera es un conjunto de puntos que determinan una poli-línea utilizada como línea de corte. Lo que hace es limitar los puntos de la muestra de entrada, considerando sólo los puntos que estén del mismo lado de la barrera que la celda de procesamiento actual. Hay que tener en cuenta que el uso de barrera amplía el tiempo de procesamiento. La herramienta de ArcGis 10.3 tiene un límite de aproximadamente unos 45 millones de puntos de entrada. Trabajar con

más puntos puede dar un fallo de respuesta. Este problema se puede solventar interpolando por partes la zona de estudio.

c) Funcionamiento: la ponderación en este procedimiento es una función del inverso de la distancia, lo que supone que la variable de entrada disminuye su influencia a medida que aumenta la distancia. Se controla con los siguientes parámetros:

1) **Control de la influencia con el parámetro “potencia”:** Es el parámetro fundamental en el método y es el encargado de controlar la influencia de los puntos en estudio. Se trata del exponente al que hay que elevar el inverso de la distancia. Este parámetro controla la influencia de los valores conocidos (medidos), en los valores obtenidos (interpolados), mediante la distancia entre ellos. Es un número real positivo y de forma generalizada se emplea el 2. Si se usa una “potencia” mayor de 2, se está enfatizando los puntos más cercanos, por lo que tendrán más influencia en el valor interpolado. La superficie así obtenida tendrá mucha variación en elementos de área pequeña, por lo que no será suave. Si por el contrario se emplea una “potencia” menor, se da mayor influencia a los puntos más lejanos, la superficie será más suave. El valor óptimo habrá que calcularlo para cada caso, en función de las necesidades y eso se puede hacer a través de **ArcGis Geostatistical Analyst**.

2) **Número de puntos utilizados para la interpolación:** Se limitan los valores de entrada en cada celda de trabajo, para un valor de la celda de salida. Es una forma de controlar la superficie interpolada. Para ello hay que especificar la cantidad de puntos de entrada por celda y el radio, dentro del cual se incluirán los puntos en la superficie de interpolación. El radio puede ser variable o fijo. Si es variable, el modelo busca los puntos necesarios para tener la cantidad especificada en cada celda. De esta forma algunas vecindades serán pequeñas y otras grandes. Se puede especificar la distancia máxima que no puede sobrepasar el radio, aunque no

haya alcanzado el número de puntos establecido. Si el radio es fijo, se requiere un valor fijo y una cantidad de puntos mínima. Si en una celda con un radio fijo no hay puntos suficientes para llegar al mínimo, se aumenta el radio hasta completar el número mínimo predeterminado de puntos en la celda. Generalmente los datos no están distribuidos uniformemente si se mantiene el radio fijo, habrá diferente número de puntos en cada celda.

- 3) **Señalización de las barreras:** una barrera no es más que un dataset de poli-línea. La barrera se utiliza como línea de corte, en los casos en los que los puntos en los que se quiere estudiar la interpolación estén fuera de la red de puntos medidos. Un caso representativo podría ser un acantilado como se ha comentado anteriormente.

1.2.4.2 Kriging en ArcGis

a) Definición: se genera una superficie a partir de un conjunto de puntos dispersos con valores z . Antes de seleccionar el método de estimación a emplear para generar la superficie de estimación se requiere un estudio profundo para encontrar la relación espacial del fenómeno en estudio y representado por los valores z ,

b) Uso: en este caso la velocidad de respuesta del método depende del número de datos con el que se quiera trabajar y de la rejilla que se establezca para calcular puntos interpolados. Es preciso recordar que el Kriging es un proceso intensivo de cálculo. Si la respuesta tiene una varianza baja, se interpreta como una respuesta buena o por lo menos de confianza. Si por el contrario la varianza toma valores elevados hay que pensar que el número de puntos que se ha dado no es suficiente y es preciso ampliar el conjunto de datos de entrada. Si se trabaja con Kriging universal hay que tener en cuenta que este método supone que existe una tendencia local variable, es decir hay una componente estructural. Si no se introduce un valor para el tamaño del paso o "lag", ArcGIS utiliza un tamaño por defecto para establecer la rejilla. Lo mismo sucede con el valor "nugget",

al que se adjudica un valor por defecto si no se le asigna de antemano. En este caso, de forma opcional, se puede asignar un valor de varianza para cada celda de la rejilla. Se obtiene la varianza Kriging, es decir una medida de los errores; suponiendo que siguen una distribución normal, la probabilidad de que el valor predicho sea suficientemente próximo al real es del 95.5 %. También aquí es recomendable preparar los datos quitando puntos coincidentes para que no haya problemas de interpretación, ya que esto puede causar que algunos valores de entrada tengan distintos valores a los que se esperaban en algunas ubicaciones del ráster.

c) Funcionamiento: Kriging es un procedimiento geoestadístico que genera una superficie estimada a partir de un conjunto de puntos dispersos con valores z . Utilizar Kriging implica una investigación del comportamiento espacial de los datos antes de seleccionar el método de estimación para generar la superficie de predicción. Este modelo supone que hay una correlación espacial que se usa para la ver la variación de la superficie, es decir, ajusta una función matemática que permite obtener el comportamiento en un punto a partir de los puntos tomados como entradas y su correlación. Este proceso lleva implícitos varios pasos para su desarrollo, análisis estadístico de los datos, modelado del variograma, creación de la superficie y estudio de la superficie de varianza. Obviamente este método es adecuado cuando se conoce la influencia direccional o correlacionada de los datos.

- 1) **Formula del Kriging:** igual que en el IDW hay que realizar una ponderación de los valores de entrada:

$$\hat{Z}(S_0) = \sum_{i=1}^N \lambda_i Z(S_i) \quad (87)$$

donde: $Z(S_i)$ es el valor medido en la ubicación i -ésima;

λ_i es una ponderación desconocida para el valor medido en la ubicación i -ésima;

S_0 es la ubicación de la predicción;

N es la cantidad de valores medidos.

Las ponderaciones se basan en la distancia entre los puntos medidos y el predicho y la disposición espacial de los puntos medidos. Es preciso cuantificar la correlación entre los puntos, por ello, la ponderación λ_i depende del modelo ajustado a los puntos, a la distancia y a la correlación espacial entre los valores medidos en la proximidad del que se va a realizar la predicción.

- 2) **Mapa de superficie de predicción con el método Kriging.** La realización de la predicción es preciso conocer las reglas de dependencia y después realizar las predicciones. Primero hay que crear los variogramas y las funciones de covarianza para calcular los valores de la dependencia estadística (autocorrelación espacial), que depende del modelo de autocorrelación (ajuste de modelo) y segundo prever los valores desconocidos (hacer una predicción). Se dice que los datos se emplean dos veces, una para la autocorrelación y otra para las predicciones.
- 3) **Variografía:** es un análisis estructural, es decir ajustar un modelo espacial. Lo primero que se hace es un grafico del semivariograma empírico:

$$\begin{aligned} \text{Semivariograma (distancia } h) \\ = 0.5 * \text{media } ((\text{valor } i - \text{valor } j)^2) \end{aligned} \quad (88)$$

- 4) Cada par de ubicaciones tiene una distancia única y en general habrá numerosos pares de puntos. Si se representan todos en un diagrama se obtiene una imagen imposible de administrar, por lo que se agrupan los pares en bins de intervalo.
- 5) **Ajustar un modelo al semivariograma empírico:** Hay que ajustar un modelo a los puntos que forman el variograma empírico, el cual proporciona información sobre la

autocorrelación espacial en los datasets, pero no da información para todas las direcciones y distancias posibles. Para asegurar que el Kriging tenga las varianzas mínimas hay que ajustar un modelo, es decir una función o curva continua, al semivariograma empírico. Para ello se selecciona una función que sirva como modelo. Habrá desviaciones de los puntos en el semivariograma respecto al modelo, unos por encima y otros por debajo, pero la suma de las distancias por encima y por debajo debe ser similar, lo que implica la selección del modelo más adecuado.

- 6) **Seleccionar los modelos de semivariograma:** en este caso la herramienta Kriging proporciona varias funciones para elegir aquel que mejor se ajuste a nuestros datos, es decir, el modelado del semivariograma empírico: circular, esférico, exponencial, gaussiana y lineal. Es bueno tener presente que el modelo influye en la predicción de los valores, en particular cuando la forma de la curva en el origen difiere de forma significativa. Cuanto más difiere, más influyen los vecinos cercanos a la predicción. El resultado es que la superficie de predicción no es suave. Como es natural, cada modelo se ajusta a un tipo de datos y correlación entre ellos.
- 7) **Comprender un semivariograma, rango, meseta y nugget:** un semivariograma representa la autocorrelación espacial entre los datos medidos, más parecido cuanto más cerca. Pero también la distancia al cuadrado será menor. Cuando se hayan diagramado todos los pares y se hayan colocado en un “bin” se ajusta un modelo para estas ubicaciones, empleando para esta labor el rango, la meseta y el nugget. El **rango** es la distancia a la que el modelo comienza a aplanarse, los puntos con distancia menor que el rango están autocorrelacionados espacialmente, los que tienen más distancia que el rango no. El valor de la Y para el cual el modelo alcanza el rango se llama **meseta**. El **nugget** es el corte del modelo con el eje Y; el efecto nugget son errores de

medida o variaciones para puntos cuyas distancias sean menores que el intervalo de muestreo.

- 8) **Realizar una predicción:** una vez que se conoce la autocorrelación de los datos y se haya usado esa información para calcular las distancias y para modelar la autocorrelación espacial, se puede hacer la predicción utilizando el modelo ajustado. Hay que realizar predicciones para cada celda de la rejilla en el centro de cada una.
- 9) **Los métodos Kriging** son básicamente dos: el universal y el ordinario. El ordinario, el más general y por lo tanto el más utilizado, supone que el valor medio constante es desconocido. Esta suposición es bastante razonable a menos que se disponga de alguna otra información o haya alguna razón que obligue a rechazarla. El universal supone que los datos tienen una tendencia de invalidación (p. ej. viento prevalente en una dirección). Esta función polinómica se resta de los valores medidos y la autocorrelación se modela a partir de los errores aleatorios. Cuando se ha ajustado el modelo a los errores aleatorios y antes de realizar la predicción se suma la función a las predicciones para obtener unos resultados satisfactorios. Este modelo universal sólo debe usarse si se conoce la tendencia de los datos y se puede justificar científicamente su descripción.
- 10) **Gráficos de semivariograma:** el modelo Kriging es complejo y requiere que se conozca la estadística espacial. Es necesario conocer profundamente los principios básicos y evaluar la adecuación de los datos para el modelado. Kriging supone que la variación espacial de los valores z es estadísticamente homogénea a lo largo de toda la superficie, lo que es fundamental para la teoría de variables regionalizadas.
- 11) **Modelos matemáticos:** los modelos matemáticos utilizados por ArcGis son los siguientes:

1.- Esférico:

$$\begin{aligned}\gamma(h) &= C_0 + C \left(\frac{3h}{2\alpha} - \frac{1}{2} \left(\frac{h}{\alpha} \right)^3 \right) & 0 < h \leq \alpha \\ \gamma(h) &= C_0 + C & h > \alpha \\ \gamma(0) &= 0\end{aligned}\tag{89}$$

2.- Circular:

$$\begin{aligned}\gamma(h) &= C_0 + C \left(1 - \frac{2}{\pi} \cos^{-1} \left(\frac{h}{a} \right) + \sqrt{1 - \frac{h^2}{a^2}} \right) & 0 < h < \alpha \\ \gamma(h) &= C_0 + C & h > \alpha \\ \gamma(0) &= 0\end{aligned}\tag{90}$$

3.- Exponencial:

$$\begin{aligned}\gamma(h) &= C_0 + C \left(1 - \exp \left(\frac{-h}{r} \right) \right) & h > 0 \\ \gamma(0) &= 0\end{aligned}\tag{91}$$

4.- Gaussiano:

$$\begin{aligned}\gamma(h) &= C_0 + C \left(1 - \exp \left(-\frac{h^2}{r^2} \right) \right) & h > 0 \\ \gamma(0) &= 0\end{aligned}\tag{92}$$

5.- Lineal:

$$\begin{aligned}\gamma(h) &= C_0 + C \left(\frac{h}{\alpha} \right) & 0 < h \leq \alpha \\ \gamma(h) &= C_0 + C & h > \alpha \\ \gamma(0) &= 0\end{aligned}\tag{93}$$

1.2.4.3 Natural Neighbor (vecino natural) en ArcGis

a) Definición: se busca el subconjunto de muestras de entrada más cercano al punto de consulta y se hace la correspondiente ponderación basándose en áreas proporcionales para encontrar el valor interpolado. A esta forma de trabajar se le conoce como interpolación de Sibson o de robo de área (*Sibson, (1981)*)

b) Uso: a la célula se le asigna un valor *NoData*, tanto en caso de que el centro de las células del perímetro caiga fuera de la zona convexa (definida por los puntos de entrada), como si el punto de entrada está dentro de las células perimetrales y cae fuera de la zona convexa. También en este caso hay posibilidad de que los valores sean distintos a los esperados. Para evitarlo hay que quitar los puntos coincidentes de los datos de entrada. En este caso hay un límite de aproximadamente 15 millones de puntos de entrada. Los resultados pueden no ser adecuados, lo que se evita dividiendo el área de estudio en varias secciones asegurando que haya solapamientos entre las secciones. En este caso es mejor que los datos de entrada sean un sistema de coordenadas proyectado y no coordenadas geográficas simplemente. Si se dispone de un conjunto de datos repartidos muy irregularmente o si hay líneas de ruptura en el conjunto de datos, es útil disponer de ArcGis 3D Analyst, el cual crea un TIN (*Triangular Irregular Networks*) a partir de los datos origen, luego se convierte TIN en una trama TIN utilizando la opción de Vecinos Natural.

c) Funcionamiento: el algoritmo utilizado en este procedimiento trabaja con un subconjunto de las muestras de entrada más próximas al punto de interpolación y aplica el cálculo de ponderaciones basándose en áreas proporcionales, según *Sibson, (1981)*. Las propiedades de este procedimiento residen en que es un modelo local y que las altitudes interpoladas están dentro del rango de las muestras empleadas. No tiene interferencias ni picos, depresiones, crestas o valles que no se muestren con los datos de entrada. La superficie generada pasa por todos los puntos de

entrada y es suave en todas las zonas, excepto en los puntos de los valores de entrada.

Si se dispone de 3D Analyst se puede utilizar la herramienta “*De TIN a ráster*” que sirve para incorporar líneas de corte. Este procedimiento aumenta los resultados de la interpolación y crea discontinuidades lineales. Se adapta localmente a la estructura de los datos de entrada, sin necesidad de introducción por el usuario del radio de la búsqueda. Funciona bien con datos distribuidos tanto aleatoriamente, como regularmente.

Los vecinos naturales de un punto son los que están asociados con polígonos de Thiessen adyacentes. Lo primero que hace es construir un diagrama de Voronoi de todos los puntos dados. A continuación se forman los nuevos polígonos alrededor del punto de interpolación y la proporción de superposición entre ambos polígonos es la ponderación.

1.2.4.4 Splines en ArcGis:

a) Definición: en este procedimiento se estiman los valores usando una función matemática que minimiza la curvatura general de la superficie, dando lugar a una superficie suave que pasa exactamente por los puntos de entrada.

b) Uso: es mejor usar la opción *REGULARIZED* que la *TENSION* ya que genera superficies más suaves. Con la primera, los valores más altos son los que se emplean para el peso, por eso suministran superficies más suaves. Los valores introducidos para estos datos tienen que ser iguales o mayores que cero siendo los valores típicos 0; 0.001; 0.01; 0.1 y 0.5. El peso es el cuadrado del parámetro, que se denomina (t). Con *TENSION* los valores que se suelen introducir son 0; 1.5 y 10. En este caso, el cuadrado del peso se representa con (ϕ). Cuantos más valores haya más suave será la superficie del ráster. Aquí también podemos encontrar valores interpolados diferentes de los que se esperan, la forma de solucionarlo está en eliminar los valores de puntos coincidentes.

c) Funcionamiento: este procedimiento estima valores usando una función matemática que minimiza la curvatura de la superficie, dando origen a una superficie suave y que pasa exactamente por todos los puntos. Este método es bueno para generar superficies que varíen levemente, como la altitud. Esa forma básica de curvatura mínima impone unas condiciones debiendo pasar la superficie por todos los valores dados y con una curvatura mínima, es decir, la suma acumulada de los cuadrados de la derivada segunda tomada sobre cada punto debe ser mínima. Esta técnica, también denominada técnica de lámina delgada, ofrece una superficie suave, continua y derivable.

- 1) **Tipos de Splines:** son dos los tipos de Splin disponibles en ArcGis, regularizado y de tensión. El primero crea una superficie suave que cambia regularmente, con valores que pueden estar fuera de rango respecto a los valores medidos. La opción de trabajo es *REGULARIZED*, modifica el criterio de minimizado para que los términos de la derivada tercera se añadan a los criterios de minimización. El **Peso**, especifica el peso adjunto a los términos de la derivada tercera durante la minimización y se denomina (τ). Cuanto más alto sea este término, la superficie de respuesta será más suave. Los valores que se emplean están entre 0 y 0.5. El Splin de tensión controla la rigidez de la superficie según el fenómeno que se esté modelando. Por lo general se crea una superficie menos suave con valores más restringidos por el rango de los datos de entrada. Modifica el criterio de minimización para que los términos de la derivada primera se incorporen a los criterios de minimización. El **Peso**, que especifica el peso adjunto a los términos de la derivada primera durante la minimización, se denomina (ϕ). Cuanto mayor sea el peso, menos rígida es la superficie y por lo tanto se aproxima más a la forma de una membrana, que pasa por todos los puntos. La superficie es suave, pero las derivadas primeras son continuas pero no suaves.

2) **Parámetros adicionales de Splin:** si se quiere un mayor control de la superficie de salida se necesitan dos parámetros adicionales, el **Peso** y la **Cantidad de puntos**. El **Peso** define el peso de las derivadas terceras de la superficie en la expresión de minimización, cuanto más alto sea, más suave es la superficie de salida. Los valores válidos para este parámetro son: 0; 0.001; 0.01; 0.1 y 0.5. El segundo parámetro, **Cantidad de puntos**, identifica el número de puntos que se emplean para el cálculo de cada celda interpolada. Cuanto mayor sea este número, mayor será la influencia de los puntos distantes sobre cada celda y más suave será la superficie de salida. Hay que tener en cuenta también que cuantos más puntos de entrada se tienen, más tiempo tarda el sistema en procesar y obtener los resultados.

3) **La ecuación de Splin:** el algoritmo que emplea en este procedimiento es:

$$S(x, y) = T(x, y) + \sum_{j=1}^N \lambda_j R(r_j) \quad (94)$$

donde: $j = 1, 2, \dots, N$, siendo N el número de puntos;

λ_j son los coeficientes determinados por la solución de un sistema de ecuaciones lineales;

r_j es la distancia del punto (x, y) al punto j .

$S(x, y)$ y $T(x, y)$ se definen de forma diferente según se emplee la opción *REGULARIZED* o la *TENSION*. El ráster de salida está dividido en bloques o regiones de igual tamaño. La cantidad de regiones en las direcciones x e y son las mismas y tienen forma rectangular. El número de regiones se obtienen dividiendo el número total de valores de entrada entre el valor especificado para la cantidad de puntos. Hay ocasiones en los que las regiones pueden tener diferente número de puntos, en regiones con distribución muy asimétrica, empleándose en este caso un valor para el número de puntos que es el promedio aproximado. Si cada

región tuviera menos de ocho puntos, hay que agrandarla para tener el mínimo de ocho puntos.

Si se elige la opción *REGULARIZED*, la expresión a emplear es:

$$T(x, y) = a_1 + a_2x + a_3y \quad (95)$$

donde: a_i son los coeficiente determinados por la solución de un sistema de ecuaciones lineales.

y

$$R(r) = \frac{1}{2\pi} \left\{ \frac{r^2}{4} \left[\text{Ln} \left(\frac{r}{2\tau} \right) + c - 1 \right] + \tau^2 \left[K_0 \left(\frac{r}{\tau} \right) + c + \text{Ln} \left(\frac{r}{2\pi} \right) \right] \right\} \quad (96)$$

donde: r es la distancia entre el punto y el valor de entrada;

τ^2 es el parámetro de **Peso**;

K_0 es la función de Bessel modificada;

c es la constante de Euler, que vale 0.577215.

Si se trabaja con *TENSION*, la expresión que rige los cálculos es:

$$T(x, y) = a_1 \quad (97)$$

donde: a_1 es un coeficiente determinado por la solución de un sistema de ecuaciones lineales

y:

$$R(r) = -\frac{1}{2\pi\varphi^2} \left[\text{Ln} \left(\frac{r\varphi}{2} \right) + c + K_0(r\varphi) \right] \quad (98)$$

donde: r es la distancia entre el punto y el valor de entrada;

φ^2 es el parámetro de **Peso**;

K_0 es la función de Bessel modificada;

c como en el caso anterior es una constante igual a 0.577215.

- 4) **Procesamiento regional de la salida:** toda la región está dividida en regiones rectangulares de igual tamaño; Las direcciones x e y tienen el mismo número de regiones. La cantidad de regiones se calcula dividiendo el número total de valores de entrada por el valor especificado de puntos. Hay regiones con un número de valores significativamente menor cuando se trabaja con valores distribuidos con poca uniformidad. El número mínimo de puntos de una región es ocho. Si no los tienen hay que ampliar la región.

1.2.4.5 Splines con barreras en ArcGis

a) Definición: se estiman valores usando una función matemática que minimiza la curvatura general de la superficie, dando lugar a una superficie suave que pasa exactamente por los puntos de entrada.

b) Uso: se requiere el entorno de ejecución Java v.6 o superior. La superficie que se obtiene está limitada por la característica de la barrera de entrada. Si hay algunos puntos duplicados, pueden afectar a la salida, obteniendo resultados no esperados. El problema se evita eliminando los datos coincidentes mediante la herramienta “*Adquirir eventos*”. La celda menor con la que se trabaja será el tamaño más pequeño de ancho o de largo de los valores de entrada dividido por 250, aunque se haya introducido el valor 0. Las características de la barrera son datos de la trama y el centro de la celda se usa para decidir si la celda está dentro del polígono o en la barrera.

c) Funcionamiento: se aplica un método de barrera mínima, comenzando por la cuadrícula vasta que se calcula con el promedio de los datos de entrada. Luego se van afinando las cuadrículas para hacer la aproximación de la superficie de curvatura mínima en el espacio de fila y columna que se quiere. En cada nivel de refinamiento, el modelo de superficie se trata como una membrana elástica aplicando un operador de deformación iterativa, lineal y convergente, repetitiva en cada nodo para poder alcanzar la aproximación deseada. Este sistema se aplica a los puntos de entrada y a las discontinuidades en las barreras. La deformación con la que se trabaja

en cada celda se obtiene como una suma molecular que compara la suma ponderada de las 12 celdas vecinas con el valor actual de una celda de destino (*Terzopoulos, (1988)*).

1.2.4.6. Top to ráster

a) Definición: es una técnica específica que también genera una superficie, de mayor precisión que el caso anterior. El sistema es capaz de preservar las redes x de transmisión de los datos de las curvas de nivel. Emplea el algoritmo de ANUDEM (Hutchinson, 2000)

b) Uso: los mejores resultados se obtienen almacenando todos los datos de entrada en un mismo sistema de coordenadas planas, *ZUnit*. Si los datos no son proyectados, los resultados no son exactos. En este caso solo se usan cuatro valores de entrada para cada celda de interpolación y los demás puntos se ignoran. Si hay demasiados puntos, se puede producir un error que indica que el conjunto de datos tiene demasiados puntos. La cantidad máxima de puntos que se puede tener es: $NRows * NCols$ (número de filas por número de columnas del ráster de salida). El algoritmo crea una morfología generalizada y luego implementa las curvas de nivel como fuentes de información. Las curvas de nivel se adaptan mejor a los datos de gran escala donde éstas, junto con las esquinas, son indicadores fiables, por ejemplo de ríos y montañas. Los valores típicos de esta configuración en las dos vertientes *Tolerancia 1* y *Tolerancia 2* son: para datos de punto a una escala 1:100.000, hay que utilizar 5 y 200; para datos con menos densidad a una escala de 1:500.000 utilizar 10 y 400 y por ultimo para datos de curvas de nivel con un espaciado de 10 utilizar 5 y 100.

c) Funcionamiento: es específico para crear modelos digitales de elevación (DEM, *Digital Elevation Model*). Está basado en el programa ANUDEM v.4.6.3. desarrollado por Hutchinson (1989), La interpolación se lleva a cabo con una serie de restricciones que aseguran una estructura de drenaje conectado y la correcta representación de crestas. Es el único interpolador de ArcGis diseñado exclusivamente para trabajar con curvas de nivel.

- 1) **Proceso de interpolación:** este proceso se ha diseñado para aprovechar los datos disponibles junto con las características conocidas de las superficies.
- 2) **Procesar datos de arroyos:** requiere que los datos tengan todos los arcos apuntando hacia abajo y que no haya polígonos trenzados en la red.
- 3) **Crear rásteres adyacentes y realizar mosaicos con ellos:** a veces es necesario crear DEM a partir de las teselas adyacentes de los datos de entrada, debido a usar mapas como entradas o a limitaciones de memoria.
- 4) **Evaluar la salida:** hay que evaluar cada superficie creada para asegurarse que los datos y parámetros suministrados por el programa sean representaciones realistas de superficie.
- 5) **Causas posibles de problemas:** en caso de que aparezca algún problema hay que revisar los siguientes puntos:
 - no hay recursos en el sistema disponibles
 - la curva de nivel o el punto de entrada es demasiado denso
 - la aplicación del interpolador puede no ser consistente

1.2.4.7. Tendencia en ArcGis:

a) Definición: es un método de interpolación polinómica global que a partir de los valores de entrada y que por medio de una función matemática ajusta una superficie suave. Ésta, captura las variaciones de los puntos de entrada y va modificando su tendencia gradualmente.

b) Uso: cuanto mayor es el grado del polinomio, más compleja es la superficie y no necesariamente más precisa. Opcionalmente se puede tener el fichero que contiene los errores RMS. Esta información se puede utilizar para determinar el grado del polinomio, cambiando el orden hasta obtener el error más bajo. Si se trabaja con la opción *LOGISTIC* (tipo de regresión), los valores de entrada debe tener como código un (0) y un (1). También en este

caso existe la posibilidad de puntos coincidentes que pueden dar resultados no esperados. Se evita rechazando los puntos coincidentes a través de la opción Adquirir eventos.

c) Funcionamiento: en este caso la interpolación ajusta a los puntos de entrada una superficie suave definida por un polinomio que cambia gradualmente y captura patrones de escala. Es como si se ajustara un trozo de papel a los puntos elevados. Algunos puntos quedaran por encima y otros por debajo, pero las sumas de las distancias debe ser similar. Cuanto menor sea el error RMS más precisa es la representación de puntos en la superficie interpolada. Los polinomios más frecuentes para este tipo de interpolación son de grado uno a tres. En este tipo de interpolación se crean superficies lisas.

- 1) **Uso:** para ajustar una superficie de puntos que varía gradualmente en una región a otra área de interés, o en un análisis de superficie de tendencia, es decir cuando se quiere examinar o quitar los efectos de las tendencias de largo alcance o globales. Este tipo de interpolación crea una superficie que varía en forma gradual utilizando polinomios de bajo grado que describen un proceso físico. A mayor complejidad de la función, más difícil es atribuirle un significado físico. Las superficies son muy sensibles a los valores atípicos, altos o bajos y especialmente a los extremos.
- 2) **Tipos de interpolación por tendencia:** lineal y logísticas. La primera, *LINEAR*, crea un ráster de punto flotante, mediante una regresión polinómica para ajustar una superficie de mínimos cuadrados a los puntos de entrada. Este tipo permite controlar el grado del polinomio utilizado para el ajuste. La opción *LOGISTIC* es adecuada para la predicción de un determinado fenómeno, en forma de probabilidad.
- 3) **Archivos RMS de salida:** en el archivo de error se guarda el error cuadrático medio de la interpolación comparando el valor de entrada con el valor de las mismas ubicaciones en la superficie de

ráster interpolada. Este valor del RMS se puede usar para cambiar el valor del grado del polinomio de interpolación hasta obtener el menor error con el parámetro. En este archivo también se guarda el valor de chi cuadrado.

1.3 MODELOS DE RADIACIÓN

El objetivo fundamental del trabajo es la obtención de un procedimiento óptimo para el cálculo de la radiación global diaria media anual en una superficie vertical orientada en las direcciones cardinales (norte, sur, este y oeste). El problema se aborda en dos pasos: primero hay que calcular el valor de la GHI sobre cualquier punto del plano horizontal a partir de métodos de interpolación, que se han introducido en el apartado anterior, y posteriormente hay que transformar ese valor a radiación global sobre superficie vertical. Este segundo paso es el que se explica en este apartado.

La aplicación fundamental de este procedimiento, que permite obtener el valor representativo del recurso solar en superficies verticales en cualquier orientación, radica en la instalación de colectores solares, fotovoltaicos y/o térmicos, sobre fachadas, para aumentar las posibilidades de generación energética distribuida, así como otras aplicaciones, como el aprovechamiento de la iluminación natural, el cálculo de sistemas de sombra o el cálculo de cargas térmicas sobre el edificio. Las importantes aplicaciones de estos procedimientos explican la existencia en la bibliografía de diferentes y variados modelos matemáticos para la transformación de datos de radiación global diaria media anual horizontal a vertical y con cualquier orientación. De entre los modelos encontrados en la bibliografía se seleccionan para este estudio cuatro característicos. Aplicados estos modelos a los datos experimentales disponibles, se escoge para la aplicación aquel que resulta más indicado para los datos disponibles. Es conveniente señalar que los valores de la radiación solar tienen una fuerte componente local, por lo que los resultados obtenidos en este trabajo no son directamente extrapolables fuera de la región en estudio, Castilla y León, y

para datos de estaciones diferentes a las utilizadas en el trabajo (red SIAR y estación propia del grupo de investigación).

1.3.1 REVISIÓN BIBLIOGRÁFICA

Como se ha comentado previamente, la dependencia de la radiación solar con el ángulo de inclinación de la superficie de incidencia es una cuestión de gran interés para los estudios de eficiencia energética de sistemas solares térmicos y fotovoltaicos. *Saraf y Hamad, (1988)* desarrolla un modelo de cálculo del ángulo de incidencia óptimo para paneles solares térmicos en Basrah, (Irak) mediante optimización de la energía producida por el colector térmico. El trabajo se desarrolla a partir de datos diarios de radiación e incluye parámetros tales como la temperatura. Se calcula la radiación global a través de sus componentes difusa, directa y reflejada, mediante una expresión en la que aparecen el ángulo de incidencia de la radiación sobre la superficie del panel, el ángulo de inclinación de la superficie respecto a la horizontal y el coeficiente de reflexión. El ángulo de inclinación óptimo varía desde 15° en el mes de junio hasta 63° en diciembre, siendo el ángulo medio óptimo, válido para todo el año, de 38.13° superior a la latitud del lugar, que tradicionalmente se considera como el ángulo de incidencia óptimo. Esta discrepancia puede deberse al hecho de optimizar la energía producida por el panel y no la radiación máxima incidente en el mismo.

Utrillas et al., (1991) analiza la radiación global en vertical en las direcciones cardinales mediante diferentes modelos: isotrópico (*Liu y Jordan, (1962)*), Temps y Coulson (*Temps y Coulson, (1977)*), Klucher (*Klucher, (1979)*), Hay (*Hay y McKay, (1985)*), Skartveit y Olseth (*Skartveit y Asle Olseth, (1986)*), Gueymard (*Gueymard, (1987)*) y Perez simplificado (*Perez et al., (1983, Perez et al., (1986)*). A partir de datos de radiación global horaria vertical en las cuatro direcciones cardinales, medidos en un periodo de tres meses de invierno en Valencia, y con los parámetros originales propuestos por los autores concluye que el modelo de Perez para 25° es el más adecuado. El modelo se mejora significativamente utilizando los parámetros calculados localmente. No hay discrepancias significativas entre los modelos

utilizados en la orientación sur. Es importante señalar que los datos corresponden a época invernal, en la que la componente difusa de la radiación es importante.

Li et al., (2002) estima datos de radiación global vertical a partir de global y directa sobre superficie horizontal. Utiliza datos horarios de tres años completos recogidos en Hong Kong. Los autores comparan el modelo propuesto con dos modelos anisotrópicos clásicos, Muneer (*Muneer, (1997)*) y Perez (*Perez et al., (1986)*) El modelo propuesto da resultados razonables cuando se trabaja con valores de radiación vertical en todas las orientaciones, pero es menos eficaz que el modelo de Perez. Sin embargo la sencillez de cálculo del modelo propuesto lo convierte en una alternativa razonable para estimaciones en el campo de la ingeniería y el diseño.

Cucumo et al., (2007) analiza la validez de modelos clásicos a partir de datos horarios sobre superficie vertical en las orientaciones cardinales. Se utiliza la correlación de Erbs (*Erbs et al., (1982)*) para descomponer la radiación global en sus componentes difusa y directa suponiendo un valor del coeficiente de reflexión de 0.2. Los resultados no presentan variación importante en función de la orientación, siendo todos los modelos igualmente efectivos, excepto en la orientación norte. En la orientación sur el modelo isotrópico es el que presenta menor error. En las orientaciones este y oeste el modelo de Hay y Davies (*Hay y Davies, (1980)*) es el más adecuado y en la dirección norte, Perez (*Perez et al., (1986)*) es el más recomendable. Otra conclusión importante del trabajo es la importancia del albedo en los cálculos. Modificando este valor a 0.15, el modelo de Perez presenta los mejores resultados en todas las direcciones. En general el modelo empleado no es determinante cuando los datos son de poca calidad, pero el conocimiento del factor de reflexión es fundamental para analizar la precisión de las medidas. Por el contrario, si se quiere precisión en la predicción, es necesario emplear el modelo mejor y el conocimiento del albedo es fundamental cuando se trata de modelos de composición.

Chirarattananon et al., (2007) determina que el modelo de Perez con coeficientes localizados es el más adecuado para modelizar datos de

iluminancia y radiación vertical en las direcciones cardinales en Thailandia. Los resultados obtenidos demuestran la importancia de las condiciones del cielo y de la posición relativa del sol con respecto al plano de orientación. Los valores de radiación difusa sobre superficies verticales son relativamente elevados y tienen una variación significativa, cuando el cielo está parcialmente cubierto.

La precisión de siete modelos clásicos es analizada en *Loutzenhiser et al., (2007)* para el cálculo de eficiencia energética con software comercial (EnergyPlus, DOE-2.1 E, TRANSYS-TUD y EST-r.). Los modelos que analiza son: isótropo (*Duffie y Beckman, (2013)*), Klucher (*Klucher, (1979)*), Hay-Davies (*Hay y McKay, (1985)*), Reindl (*Reindl et al., (1990b)*), Munner (*Munneer, (1997)*) y Perez (*Perez et al., (1987)*, *Perez et al., (1990)*). Se emplean en el estudio datos horarios de radiación global horizontal, difusa y directa, en diferentes características climatológicas y altitudes solares en Duebendorf, Suiza.

Orehounig et al., (2014) compara seis posibilidades de combinación de valores de radiación (global y difusa horizontal, o solamente global) con los modelos de Reindl (*Reindl et al., (1990b)*) y Perez (*Perez et al., (1991)*), para obtener los valores sobre superficies inclinadas en Austria. Los resultados cuando emplean la opción de utilizar como datos de entrada, valores de radiación global y difusa, son mejores que los que solo emplean global horizontal, pero el rango de diferencia es muy pequeño, menor que el 20 %. De las seis opciones empleadas, dos requieren datos de radiación global y difusa y emplean los sistemas de Energy-Plus y RADIANCE, las otras cuatro solo usan datos de radiación global, y emplean los modelos de Perez y Reindl, para el cálculo de la fracción directa o difusa y trabajan con los dos programas de simulación Energy-Plus y RADIANCE.

Los modelos de composición (trasposición) y descomposición son analizados en (*Yang et al., (2013)*). Calcula el valor de la radiación solar sobre superficies inclinadas a partir de medidas de radiación horizontal. A través de valores experimentales horarios tomados en una región tropical (Singapur), con diferentes inclinaciones y acimuts, se calcula la radiación

horizontal utilizando cuatro modelos de composición isótropos (Liu y Jordan (*Liu y Jordan, (1962)*), Badescu (*Badescu, (2002)*), Tian (*Tian et al., (2001)*) y Koronakis (*Koronakis, (1986)*) y seis anisótropos (Hay (*Hay y McKay, (1985)*), Reindl (*Reindl et al., (1990b)*), Temps y Coulson (*Temps y Coulson, (1977)*), Klucher (*Klucher, (1979)*), Perez (*Perez et al., (1987)*) y Skarveit y Olseth (*Skartveit y Asle Olseth, (1986)*). Como modelos de descomposición emplean cinco modelos diferentes: (Erbs (*Erbs et al., (1982)*), Orgill (*Orgill y Hollands, (1977)*), Reindl univariante y bivariante (*Reindl et al., (1990b)*, *Reindl et al., (1990a)*) y Maxwell (*Maxwell, (1987)*). En este trabajo a partir de valores horarios medidos en diferentes inclinaciones y acimuts en una región tropical (Singapur), calculan la radiación horizontal utilizando los dos procedimientos: composición y descomposición. Como resultado los modelos anisótropos de trasposición son más precisos, pero las diferencias no son significativas para inclinaciones inferiores a 20°. La conversión de radiación inclinada a horizontal es precisa pero, para ángulos acimutales pequeños, los errores aumentan de forma exponencial, con el ángulo cenital. Los modelos más complejos no presentan mejoras significativas con respecto a los modelos más simples. Cuando se combinan modelos de descomposición con modelos de trasposición, el error al convertir los datos de radiación inclinada a horizontal aumenta.

Raichijk y Taddei, (2012) hacen un estudio comparativo con 12 modelos empíricos de descomposición para el cálculo de la radiación solar directa horaria. Se utilizan datos de radiación global horizontal y normal directa en Argentina recogidos durante año y medio. Se trabaja con ocho modelos para la fracción difusa, uno para la transmitancia directa, otro para el índice de claridad de fracción difusa y otros dos para radiación normal directa. Una característica fundamental de este trabajo es la utilización siempre de valores correspondientes a altura solar superior a 5°. Utilizando los estimadores estadísticos habituales (RMSE, MBE y el t-estadístico), los modelos de mejor respuesta fueron el modelo de Perez (*Perez et al., (1987)*) con un 15.2 % y Skartveit (*Skartveit y Asle Olseth, (1986)*) con un 16.3 %, concluyendo que los modelo más modernos no introducen mejoras significativas en la predicción.

Gueymard y Myers, (2009) comparan 10 modelos de trasposición con valores de entrada óptimos y subóptimos (habituales en aplicaciones de ingeniería cuando no se dispone de los datos óptimos) para superficies inclinadas 40° y orientación vertical sur, obteniendo el modelo propio resultados comparables con el modelo de Perez. Los modelos de descomposición para radiación directa y difusa ofrecen predicciones sesgadas en función del modelo. Comparando en las mismas condiciones los modelos de descomposición y de trasposición, obtiene que en caso de conocer únicamente valores de radiación local, la bondad del modelo viene determinada por la representatividad local del modelo de difusa y el valor seleccionado para el albedo. En España, *López et al., (2000)* trabaja también con modelos de descomposición utilizando datos del Instituto Nacional de Meteorología y de estaciones meteorológicas propias. Proponen tres modelos diferentes en función de las condiciones del cielo a través del índice de claridad, la transmitancia y el $\cos\theta_z$ y los comparan con modelos clásicos de la bibliografía. Concluyen que las nuevas correlaciones de transmitancia para el cálculo de la radiación directa, en las que se incluye información geométrica (como el ángulo acimutal), mejoran las predicciones.

El análisis de optimización de la radiación en función del ángulo de inclinación de un panel, también se puede hacer por medio de correlaciones numéricas. *Elsayed, (1989)* trabaja con largas series de medidas y emplea dos correlaciones numéricas para predecir el ángulo óptimo mensual de inclinación de un panel, orientado al sur, para cualquier periodo de tiempo. Las dos correlaciones las aplica con albedos diferentes, 0.2 y 0.7. También analiza el ángulo óptimo de inclinación y el ángulo acimutal de superficies expuestas a sombras por objetos próximos. Generalmente el ángulo óptimo acimutal es 0° , es decir orientada al ecuador, pero variaciones del ángulo óptimo, inferiores a 10° , no produce prácticamente modificación de la radiación incidente.

Şen et al., (2004) realiza una completa revisión y clasificación de los modelos de radiación para diferentes aplicaciones e implementa un modelo lineal de cálculo de radiación, a partir de las horas de insolación. *El-Sebaei et*

al., (2010) analiza la influencia de otras variables, como la temperatura, humedad, nubosidad y número de horas de sol, sobre la radiación global y difusa horizontal en Arabia Saudí. Propone un método empírico para calcular la radiación difusa horizontal y calcula la radiación global sobre superficie con distintos ángulos de inclinación utilizando los modelos de Liu y Jordan *Liu y Jordan, (1962)* y Klucher (*Klucher, (1979)*). Los resultados obtenidos concuerdan con los datos proporcionados por NASA SEE. El ángulo de inclinación óptima en la localización estudiada coincide con la latitud para una superficie orientada al sur.

Stanciu y Stanciu, (2014) intentan encontrar la inclinación óptima para colectores situados en distintas localizaciones con distinta latitud y en diferente época del año a través de tres modelos (*Hottel y Woertz, (1942)*, *Liu y Jordan, (1962)* y el modelo Hay-Davis-Klucher-Reindl (*Duffie y Beckman, (2013)*, *Reindl et al., (1990b)*). Cada modelo da como inclinación óptima una distinta, pero no superior a 10° de diferencia entre el modelo más sencillo y el más complejo.

Análisis semejantes en diferentes localizaciones pueden encontrarse en *Okundamiya y Nzeako, (2011)*, en Egipto, *Baklouti et al., (2012)* en Túnez o *Pérez-Higueras et al., (2012)* en España. Todos ellos utilizan diferentes bases de datos y ponen a prueba modelos clásicos para calcular las componentes de la radiación solar (difusa, directa) en función de datos meteorológicos y/ geográficos.

La clasificación de los cielos en función de distintos tipos de índices de claridad, fracción directa y/o difusa también es un concepto ampliamente utilizado. En este sentido, los diferentes criterios de clasificación del cielo y el solape de los intervalos de diferenciación entre cielos claros, intermedios y cubiertos es un hecho que complica y mucho la aplicación de los modelos y la comparación entre ellos. *Okundamiya y Nzeako, (2011)* o *Bakirci, (2012)* trabajan fundamentalmente con el índice de claridad clásico. *Demain et al., (2013)* demuestra que el ajuste de los modelos varía según las condiciones de cielo existiendo grandes discrepancias tanto para cielos intermedios como para cielos claros o cubiertos.

Enriquez et al., (2012) estima la radiación reflejada a partir de la global horizontal y vertical en zonas urbanas. Demuestra la necesidad de incorporar procedimientos de regularización para poder aplicar con la misma resolución la técnica directa (a partir de la radiación global horizontal y vertical) y la inversa (aplicando correlaciones de difusa y modelos físicos).

La forma de trabajar con los datos disponibles y la influencia de la zona en estudio se ha revelado como fundamental a la hora de analizar el comportamiento de los diferentes modelos. Muchos autores tratan de ajustar localmente los modelos propuestos. *Notton et al., (2006)* intenta el cálculo de radiación solar global sobre superficies inclinadas a través de valores horizontales. Analiza la precisión de modelos clásicos para el cálculo de difusa a partir de global y de global inclinada a partir de difusa y global horizontal. Trabaja con 94 combinaciones de modelos conocidos para obtener radiación global sobre 45° y 60° de inclinación. Varias combinaciones de las empleadas dieron resultados satisfactorios concluyen poniendo de manifiesto la dificultad de señalar un método óptimo de cálculo. *Badescu et al., (2012)*, *Badescu et al., (2013)* analiza 54 modelos de radiación global y difusa en Rumanía y destacan la dependencia de los resultados de buen número de factores atmosféricos y de los datos de entrada disponibles. Otros modelos no lineales, los de lógica difusa, o las redes neuronales (ANN, *Artificial Neuronal Network*) también se han aplicado al problema que nos ocupa (*Khatib et al., (2012)*, *Tolabi et al., (2014)*) Los métodos ANN resultan más precisos frente a modelos lineales y no lineales.

1.3.2. CONCEPTOS GENERALES DE RADIACIÓN SOLAR

Las relaciones astronómicas entre el Sol y la Tierra, son imprescindibles para la aplicación de los modelos matemáticos del cálculo de radiación global vertical en las direcciones norte, sur, este y oeste. Son parámetros que representan la variabilidad diaria de la relación entre el Sol y la Tierra. La Tierra gira alrededor del Sol describiendo una órbita elíptica, siendo el Sol uno de los focos. La forma de la órbita hace que la distancia entre el Sol y la Tierra no sea uniforme, lo que implica que la cantidad de radiación solar que

incide sobre la superficie de la Tierra no es constante, como tampoco lo es el ángulo de incidencia. Estos dos aspectos hacen que sea imprescindible conocer la distancia entre el Sol y la Tierra cada día del año.

Se denomina factor de corrección de la excentricidad de la órbita terrestre a E_0 y se puede calcular a través de la expresión matemática :

$$E_0 = \left(\frac{r_0}{r}\right)^2 = 1.000110 + 0.034221\cos\Gamma + 0.001280\sen\Gamma + 0.000719\cos 2\Gamma + 0.000077\sen 2\Gamma \quad (98)$$

donde: Γ (radianes) es el ángulo del día que se puede calcular con la expresión:

$$\Gamma = \frac{2\pi(d_n - 1)}{365} \quad (99)$$

siendo d_n el número de día del año, 1 para el 1 de enero y 365 para el 31 de diciembre. Se emplea siempre el mes de febrero con 28 días.

El plano de revolución de la Tierra alrededor del Sol se denomina **plano eclíptico**. Este plano y el eje polar, respecto al que la tierra gira para dar una vuelta sobre sí misma, forman un ángulo, que varía cada día, denominado **declinación**, δ , parámetro que se calcula por medio de la expresión:

$$\begin{aligned} \delta = & (0.006918 - 0.399912\cos\Gamma + 0.070257\sen\Gamma \\ & - 0.006758\cos 2\Gamma + 0.000907\sen 2\Gamma \\ & - 0.002697\cos 3\Gamma + 0.00148\sen 3\Gamma) \left(\frac{180}{\pi}\right) \end{aligned} \quad (100)$$

(δ se expresa en grados).

Otro parámetro característico y necesario para obtener los valores de radiación es la **constante solar**, \dot{I}_{sc} , definida como la energía procedente del Sol por unidad de tiempo y de área perpendicular a los rayos del Sol, que llega a un punto más allá de la atmosfera a una distancia del Sol igual al radio medio de la órbita terrestre. Realmente no es una constante, sino que varía a lo largo del año. Los diferentes estudios realizados unifican su valor casi de forma general como: $\dot{I}_{sc} = 1367 \text{ W}\cdot\text{m}^{-2}$, o $I_{sc} = 4921 \text{ kJ}\cdot\text{m}^{-2}\cdot\text{h}^{-1}$.

La **radiación extraterrestre**, es la radiación que incidiría en una superficie con cualquier orientación y en cualquier momento, en ausencia de atmósfera. Se puede definir la radiación extraterrestre sobre superficie horizontal o sobre superficie inclinada con un ángulo y orientada arbitrariamente, cada uno de ellos para diferentes periodos de tiempo (horaria, diaria, diezminutal...). Las siguientes expresiones dan cuenta de su cálculo:

- Radiación extraterrestre horaria sobre horizontal

$$I_0 = I_{sc}E_0 \left(\text{sen}\delta\text{sen}\phi + \left(\frac{24}{\pi}\right)\text{sen}\left(\frac{\pi}{24}\right)\text{cos}\delta\text{cos}\phi\text{cos}\omega_i \right) \quad (101)$$

- Radiación extraterrestre diaria sobre horizontal

$$H_0 = \left(\frac{24}{\pi}\right) I_{sc}E_0\text{cos}\phi\text{cos}\delta \left[\text{sen}\omega_s - \left(\frac{\pi}{180}\right)\omega_s\text{cos}\omega_s \right] \quad (102)$$

- Radiación extraterrestre horaria sobre una superficie orientada arbitrariamente

$$I_{0,\beta\gamma} = \left(\frac{12}{\pi}\right) I_{sc}E_0 \left[0.26A\text{sen}\omega_i + 0.26B\text{cos}\omega_i + \left(\frac{\pi}{12}\right)C \right] \quad (103)$$

$$A = \text{cos}\delta\text{sen}\beta\text{sen}\gamma$$

$$B = \text{cos}\delta\text{cos}\phi\text{cos}\beta + \text{cos}\delta\text{sen}\phi\text{sen}\beta\text{cos}\gamma$$

$$C = \text{sen}\delta\text{sen}\phi\text{cos}\beta - \text{sen}\delta\text{cos}\phi\text{sen}\beta\text{cos}\gamma$$

- Radiación extraterrestre diaria sobre una superficie orientada arbitrariamente

$$\begin{aligned}
 H_{0,\beta\delta} = & \left(\frac{12}{\pi}\right) I_{sc} E_0 \left(\cos\beta \operatorname{sen}\delta \operatorname{sen}\phi \frac{|\omega_{ss} - \omega_{sr}|\pi}{180} \right. \\
 & - \operatorname{sen}\delta \cos\phi \operatorname{sen}\beta \cos\gamma \frac{|\omega_{ss} - \omega_{sr}|\pi}{180} \\
 & + \cos\phi \cos\delta \cos\beta |\operatorname{sen}\omega_{ss} - \operatorname{sen}\omega_{sr}| \\
 & + \cos\delta \cos\gamma \operatorname{sen}\phi \operatorname{sen}\beta |\operatorname{sen}\omega_{ss} - \operatorname{sen}\omega_{sr}| \\
 & \left. + \cos\delta \operatorname{sen}\beta \operatorname{sen}\gamma |\cos\omega_{ss} - \cos\omega_{sr}| \right)
 \end{aligned} \tag{104}$$

donde: β es el ángulo de inclinación de la superficie respecto a la horizontal (grados);

γ es el ángulo acimutal de la superficie (grados);

I_{sc} es el constante solar ($\text{W}\cdot\text{m}^{-2}$) ò ($\text{kJ}\cdot\text{m}^{-2}\cdot\text{h}^{-1}$);

E_0 es el factor de corrección de la excentricidad de la órbita terrestre;

δ es la declinación del lugar;

ϕ es la latitud del lugar;

ω_i es el ángulo horario (se emplea el del valor medio de la hora en estudio) (grados);

ω_s es el ángulo de la puesta del sol para superficie horizontal (grados);

ω_{sr} es el ángulo salida del sol para una superficie arbitrariamente orientada (grados);

ω_{ss} es el ángulo puesta del sol para una superficie arbitrariamente orientada (grados).

- Índice de claridad

Para ciertas aplicaciones es necesario determinar la distribución estadística de la radiación global diaria en un cierto lugar (1983) define el índice de claridad diario K_T

$$K_T = \frac{H}{H_0} \tag{105}$$

donde: H_0 es radiación extraterrestre diaria sobre superficie horizontal

($\text{MJ}\cdot\text{m}^{-2}\cdot\text{dia}^{-1}$);

H es la radiación diaria sobre superficie horizontal ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{d}^{-1}$) en un lugar.

Según el índice de claridad los días se clasifican en cubiertos, parcialmente cubiertos y días claros.

1.3.3 CLASIFICACIÓN DE LOS MODELOS DE RADIACIÓN

La clasificación de los modelos se hace en base a la componente difusa, es esta la que más dependencia tienen de las condiciones de la atmósfera y/o del cielo. A la vista de los modelos encontrados en la revisión bibliográfica se ha pensado en hacer una clasificación de los mismos, entendiendo que no es una tarea fácil, ya que son muchos los puntos respecto a los cuales los autores de los modelos establecen las condiciones del mismo. Se trata de agruparlos en función de algún término o parámetro que los identifique.

Se puede hablar de modelos *paramétricos*, caracterizados por un conjunto de “parámetros” que definen las condiciones de cielo y atmósfera y que hay que ajustar en función de unas determinadas características. Para complementar están los *no paramétricos*, por lo general dependen de la influencia que una o varias variables puede tener en el cálculo de la irradiación solar, son habitualmente mucho más complejos y se resuelven con informática científica, redes neuronales, lógica difusa, etc.

Dentro de los paramétricos, la historia siempre ha clasificado los modelos como *isótropos* y *anisótropos*. Esta clasificación se hace en función de las condiciones de cielo o de la atmósfera. Si se considera que la distribución de la radiación solar difusa es homogénea en toda la bóveda celeste, se dice que el cielo es isótropo. En caso contrario estamos en condiciones de cielo anisótropo, esta variación de distribución de radiación, se estudia en la región circunsolar (zona próxima al Sol), o en la límite del horizonte.

Si nos centramos en los modelos isotrópicos, nos encontramos con el primero establecido en 1961 y que se denomina Liu y Jordan, es el más sencillo, pero también el que mejor se ajusta en la mayoría de los estudios.

Koronakis en 1986, presenta un modelo isotrópico que no es más que una modificación de su antecesor Liu y Jordan. Olmo et al. en 1999, desarrolla un modelo para el cálculo de irradiación global en superficies inclinadas con datos tomados en Granada. En 2001 Tian et al. hace una consideración en la distribución isótropa de la radiación que es proporcional a la superficie esférica y no a la proyección como en los anteriores. Y en 2002, Badescu trata de mejorar el modelo Liu y Jordan.

Dentro de los modelos anisótropos podemos encontrar dos grandes grupos, los de superposición y los de gradiente de radiancia. Los de superposición son los más frecuentes y también los más testeados en diferentes partes del mundo. En estos se puede hacer una nueva clasificación atendiendo a la radiación, por una parte la circunsolar y por otra la del horizonte. En general los modelos emplean ambas componentes. La relación de los más conocidos es la siguiente:

Gueymard (1987 y 1988): genera lo que se puede denominar, modelo para todo tipo de cielo. La expresión tiene dos sumandos, uno para cielo totalmente despejado, situación anisótropa donde incluye la radiación de la región circunsolar y la semiesférica de la componente difusa con el ángulo cenital y otro para cielo totalmente cubierto, que se corresponde con la proporción de radiación de onda corta que incide sobre una superficie inclinada respecto al plano horizontal. El autor también analiza la componente reflejada o de albedo introduciendo un nuevo concepto, el modelo anisotrópico semiesférico.

Hay y Davies (1979, 1980 y 1988): el primer modelo de 1979 se modifica con el posterior de 1980 y amplía con el de 1988, modelando la componente difusa como composición de la difusa isótropa de fondo y la de la región circunsolar. Define también una función de modulación o índice de anisotropía. Lo más característico de este modelo es que no considera el brillo de la banda del horizonte y la componente reflejada se considera isótropa.

Klucher (1979): es una modificación del modelo de Temps y Coulson, lo que hace es generalizarlo para todas las posibles condiciones de cielo, empleando una función de modulación que coincide con la expresión del modelo de Liu y Jordan para cielos cubiertos y la de Temps y Coulson para cielos parcialmente cubiertos o despejados. Para la radiación reflejada emplea el modelo isotrópico.

Ma y Iqbal (1983): los autores plantean una modificación del modelo de Hay y Davies, en base a un estudio del mismo con valores recogidos en Canadá y para inclinaciones de 30° , 60° y 90° . Combinan la atmosfera isótropa de fondo como el modelo Hay y Davies y una componente de radiación de la región circunsolar, introducen un factor que es el índice de claridad que relaciona la radiación global en el plano horizontal con la extraterrestre en el mismo plano y para evaluar la reflejada emplean el modelo de Liu y Jordan.

Muneer (1987): presenta diferentes trabajos en los que va modificando las condiciones de su primer modelo, pretendiendo mejorar la estimación de la radiación difusa. El modelo contempla la distribución de la radiación y una función de modulación o función de claridad del cielo y distingue entre cielos claros y cubiertos.

Perez (1983, 1987 y 1990): el primer modelo de este autor contempla en un conjunto las tres componentes de radiación difusa, isótropa, circunsolar y brillo de la banda del horizonte, junto con la luminancia. Posteriormente modifica y fundamentalmente simplifica la expresión del primer modelo.

Reindl (1990): a partir de diferentes comprobaciones con varios modelos y orientaciones, establece que es fundamental emplear un término que denomina índice de anisotropía, que represente la parte de difusa, tratada como circunsolar y el resto lo considera como condiciones isótropas. También añade un término para considerar el brillo de la banda del horizonte similar a la de Temps y Coulson.

Skartveitand Olseth (1986): modifica el modelo de Hay y Davies, validando un modelo de radiación difusa para planos inclinados en lugares de latitudes elevadas. Este modelo considera diferentes componentes para cuantificar,

radiación circunsolar, difusa cenital, difusa de fondo, difusa apantallada y la apantallada por los obstáculos del horizonte.

Temps y Coulson (1977): en este modelo se incorporan dos factores, uno para considerar el incremento de la radiación en la zona circunsolar y el otro para cuantificar la luminosidad de las zonas próximas al horizonte. En primera aproximación se establece como una modificación del modelos de Liu y Jordan.

Willmot (1982): está caracterizado por calcular la radiación difusa sobre superficies inclinadas, como suma de tres componentes, la difusa de fondo, la proporción de la reflejada por el terreno y la proporción de reflejada por obstáculos horizontales.

Los modelos anisótropos de gradiente de radiancia evalúan la componente difusa sobre cualquier inclinación como una integración de la radiancia y la difusión en función de los ángulos acimutal y cenital en cada región diferencial del cielo. La característica de estos modelos es la gran dependencia de la radiancia del tipo de cielo que se esté estudiando, el problema es que no hay ninguna clasificación normalizada del cielo. La gran dificultad estriba en la resolución de las integrales planteadas.

También están los modelos no paramétricos, entendiéndolos como aquellos que no utilizan ningún criterio físico o geométrico en su formulación, simplemente son una estructura obtenida como puntos de salida a partir de un conjunto de puntos de entrada, es decir son los que proporcionan mapas. Entre ellos podemos citar los de inferencia borrosa, redes neuronales artificiales y redes neuro-difusas.

1.3.4 MODELOS DE RADIACIÓN SOLAR EMPLEADOS EN ESTE TRABAJO

Dado que la bibliografía nos muestra una gran cantidad de modelos y muchos autores coinciden en señalar que la bondad del modelo depende fundamentalmente de las condiciones atmosféricas y no tanto del modelo en

sí, se han seleccionado cuatro modelos clásicos y por lo tanto muy probados en distintas zonas de la geografía mundial. Los modelos seleccionados son: Circunsolar (*Iqbal, (1983)*), Isótropo (*Iqbal, (1983)*), Klucher (*Klucher, (1979)*) y Hay (*Hay y Davies, (1980)*). Estos cuatro modelos no son ni los más sencillos ni los más complicados, simplemente abarcan todas las posibles condiciones del cielo, claro, cubierto y parcialmente cubierto.

La característica común a la mayoría de los modelos matemáticos que permiten conocer el valor de la radiación global anual media diaria en un punto, es que el valor de esa radiación es suma de tres componentes, la directa, la difusa y la reflejada. La diferencia entre modelos reside en las expresiones utilizadas por cada modelo para el cálculo de estas componentes.

La radiación directa es aquella que recibe la superficie inclinada y orientada arbitrariamente directamente del sol. La radiación difusa proviene del resto del cielo visible por el sensor y la radiación reflejada es la que recoge después de reflejarse en cualquier superficie próxima al sensor, como suelo, u obstáculos cercanos. La expresión matemática que permite el cálculo de la radiación global que incide en una superficie orientada un ángulo γ e inclinada un ángulo β , es:

$$H_{\beta\gamma} = H_{b,\beta\gamma} + H_{d,\beta\gamma} + H_{r,\beta\gamma} \quad (106)$$

donde: $H_{\beta\gamma}$ es la radiación global diaria sobre superficie orientada

arbitrariamente;

$H_{b,\beta\gamma}$ es la radiación directa diaria sobre superficie orientada arbitrariamente;

$H_{d,\beta\gamma}$ es la radiación difusa diaria sobre superficie orientada arbitrariamente;

$H_{r,\beta\gamma}$ es la radiación reflejada diaria sobre superficie orientada arbitrariamente;

Es conveniente aclarar que cuando se habla de radiación difusa, se entiende que es la difusa en el cielo y cuando se habla de radiación reflejada nos referimos a la reflejada sobre la Tierra.

En este trabajo se va a tratar con valores diarios, ya que como se ha comentado en ocasiones anteriores, entendemos que para el cálculo de la producción de una instalación fotovoltaica, no es fundamental en conocimiento de la energía generada en intervalos menores y es suficiente conocer una media diaria que permita conocer la producción anual media. Los cálculos se realizarán para una superficie orientada arbitrariamente (en este caso, orientación norte, sur, este y oeste) y con ángulo de inclinación de 90° respecto a la horizontal del lugar.

1.3.4.1 Radiación directa

La radiación directa es aquella que se recibe en un punto cualquiera procedente del Sol, de forma que no sufra ninguna desviación en su trayectoria desde el Sol a la superficie de la Tierra. Se mide con un pirheliómetro. Este sensor emplea unos obturadores, de forma que únicamente mide la radiación procedente del Sol y de una región anular de cielo próxima al Sol. Esta región, en la actualidad, abarca aproximadamente unos 2,5° medidos desde el centro del Sol. El pirheliómetro tiene un visor con un pequeño punto luminoso que coincide con una marca en el centro, lo que permite saber si se halla en la posición adecuada (cuando esté exactamente en posición perpendicular al haz solar). Por esta razón el pirheliómetro necesita un mecanismo que desplace el sensor continuamente realizando un seguimiento de la posición del Sol en todo momento.

En este caso se va a calcular la radiación directa sobre una superficie orientada un ángulo γ e inclinada un ángulo β respecto de la horizontal, caso general que particularizaremos al caso de estudio. Se puede expresar por medio de la ecuación (107)

$$H_{b,\beta\gamma} = \frac{H_b \int_{\omega_{sr}}^{-\omega_{ss}} \tau_b(\omega) \dot{I}_{0,\beta\gamma} d\omega}{\int_{\omega_s}^{-\omega_s} \tau_b(\omega) \dot{I}_0 d\omega} \quad (107)$$

donde: H_b es la radiación directa diaria sobre una superficie horizontal ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{dia}^{-1}$);

$\tau_b(\omega)$ es la transmitancia atmosférica para radiación directa, variable con el ángulo horario (adimensional);

$\dot{I}_{0,\beta\gamma}$ es la radiación extraterrestre horaria sobre plano inclinado un ángulo β y orientado un ángulo γ ($\text{kJ}\cdot\text{m}^{-2}\cdot\text{h}^{-1}$);

ω_{sr} y ω_{ss} son el ángulo horario de amanecer y de anochecer respectivamente, para una superficie inclinada y orientada (grados);

ω_s es ángulo horario para la superficie horizontal (grados);

\dot{I}_0 es la radiación extraterrestre horaria sobre superficie horizontal ($\text{kJ}\cdot\text{m}^{-2}\cdot\text{h}^{-1}$).

Suponiendo que la transmitancia atmosférica, $\tau_b(\omega)$, es constante a lo largo de un día cualquiera, la expresión se puede simplificar y expresar por medio de la ecuación:

$$H_{b,\beta\gamma} = \frac{H_b \int_{\omega_s}^{-\omega_{ss}} \dot{I}_{0,\beta\gamma} d\omega}{\int_{\omega_s}^{-\omega_s} \dot{I}_0 d\omega} \quad (108)$$

O bien a través de la expresión:

$$H_{b,\beta\gamma} = \frac{H_b H_{0,\beta\gamma}}{H_0} \quad (109)$$

donde: $H_{0,\beta\gamma}$ es la radiación extraterrestre diaria sobre superficie inclinada un ángulo β y orientada γ ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{dia}^{-1}$);

H_0 es la radiación extraterrestre diaria sobre superficie horizontal ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{dia}^{-1}$).

Normalmente al cociente entre las radiaciones extraterrestres se le denomina, R_b , conocido como **factor de conversión para la radiación directa diaria**, por lo que la expresión de la radiación directa se puede reducir a la ecuación:

$$H_{b,\beta\gamma} = H_b R_b \quad (110)$$

R_b se puede poner también en función de los parámetros característicos del lugar, del día, del ángulo de inclinación y del ángulo de orientación según la expresión:

$$\begin{aligned}
 R_b = & \left[\cos\beta \operatorname{sen}\delta \operatorname{sen}\phi \left(|\omega_{ss} - \omega_{sr}| \pi / 180 \right) \right. \\
 & - \operatorname{sen}\delta \cos\phi \operatorname{sen}\beta \cos\gamma \left(|\omega_{ss} - \omega_{sr}| \pi / 180 \right) \\
 & + \cos\phi \cos\delta \cos\beta |\operatorname{sen}\omega_{ss} - \operatorname{sen}\omega_{sr}| \\
 & + \cos\delta \cos\gamma \operatorname{sen}\phi \operatorname{sen}\beta |\operatorname{sen}\omega_{ss} - \operatorname{sen}\omega_{sr}| \\
 & \left. + \cos\delta \operatorname{sen}\beta \operatorname{sen}\gamma |\cos\omega_{ss} - \cos\omega_{sr}| \right] \times \\
 & \times \frac{1}{2[\cos\phi \cos\delta \operatorname{sen}\omega_s + (\omega_s \pi / 180) \operatorname{sen}\phi \operatorname{sen}\delta]} \quad (111)
 \end{aligned}$$

donde: β , γ , δ , ϕ , ω_s , ω_{sr} y ω_{ss} son ángulos determinados por el día y el lugar, según se han definido en apartados anteriores en este capítulo.

1.3.4.2. Radiación reflejada

En este caso se entiende por reflejada, aquella radiación que “llega” a la superficie inclinada y orientada después de ser reflejada por la parte de tierra que ve la superficie sobre la que incide. En cierta manera se puede decir que lo que se mide con este término es la parte difusa de radiación reflejada por la superficie terrestre y que incide en el sensor o superficie colectora.

Se puede decir que los albedos para la radiación difusa y directa no coinciden. Es decir, si A_g es el área total vista por el plano inclinado, la radiación total reflejada por la parte de tierra en todo el hemisferio se puede escribir por medio de la expresión (112)

$$(I_b\rho_b + I_d\rho_d)A_g \quad (112)$$

donde: ρ_b es el albedo de radiación directa;

ρ_d es el albedo de radiación difusa.

En este trabajo se van a considerar dos casos de estudio: **isótropo**, es el caso en el que toda la radiación reflejada es difusa y **anisótropo** cuando hay parte de reflejada que no es difusa.

a) reflejada isotrópica:

$$H_{r,\beta\gamma} = \frac{1}{2}H\rho(1 - \cos\beta) \quad (113)$$

donde: H es la radiación global diaria sobre superficie horizontal ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{dia}^{-1}$);

ρ es el albedo terrestre.

b) reflejada anisotrópica:

$$H_{r,\beta\gamma} = \sum_{\text{día}} \left[I\rho \left(\frac{1 - \cos\beta}{2} \right) \left(1 + \text{sen}^2 \frac{\theta_z}{2} \right) (|\cos\Delta|) \right] \quad (114)$$

donde: I es la radiación global horaria sobre superficie horizontal ($\text{kJ}\cdot\text{m}^{-2}\cdot\text{h}^{-1}$);

θ_z es el ángulo cenital;

Δ es el acimut de la superficie inclinada respecto al acimut del sol (grados).

Hay que tener en cuenta que las condiciones de cielo influyen de manera considerable sobre el resultado de la medida de radiación reflejada. El cielo totalmente claro y cubierto no son escenarios similares por lo que es fundamental tener en cuenta la diferencia de radiación reflejada incidente sobre el mismo plano en ambas situaciones.

1.3.4.3. Radiación difusa

En este apartado se va a calcular la componente que incide sobre la superficie inclinada y orientada después de sufrir diferentes cambios en su

camino del Sol a la superficie terrestre. Puede ser desviada de su trayectoria por los elementos que haya en la atmosfera, por un cuerpo traslucido, etc...

En este caso, como en el anterior, se van a considerar tres condiciones de cielo totalmente diferentes: ***circunsolar, isótopo y anisótopo***. Se entiende que el cielo es ***circunsolar*** cuando está totalmente claro y sin nubes. En este caso la radiación solar puede ser tratada como la directa.

a) modelo circunsolar:

$$H_{d,\beta\gamma} = H_d R_b \quad (115)$$

donde: H_d es la radiación difusa diaria sobre superficie horizontal
(MJ·m⁻²·dia⁻¹)

b) modelo isotrópico:

$$H_{d,\beta\gamma} = \frac{1}{2} H_d (1 + \cos\beta) \quad (116)$$

c) modelos anisótopos:

En este caso vamos a emplear dos modelos clásicos y muy conocidos, el modelo Klucher y el Hay

c1) modelo Klucher (Klucher, (1979)):

$$H_{d,\beta\gamma} = \sum_{\text{día}} \left\{ I_d \left(\frac{1 + \cos\beta}{2} \right) \left[1 + F \operatorname{sen}^3 \left(\frac{\beta}{2} \right) \right] (1 + F \cos^2 \theta \operatorname{sen}^3 \theta_z) \right\} \quad (117)$$

donde: I_d es la radiación horaria difusa sobre superficie horizontal
(kJ·m⁻²·h⁻¹);

F es una función de modulación definida como

$$F = 1 - \left(\frac{I_d}{I}\right)^2 \quad (118)$$

siendo I la radiación horaria global sobre superficie horizontal
($\text{kJ}\cdot\text{m}^{-2}\cdot\text{h}^{-1}$);

c2) modelo Hay (Hay y Davies, (1980):

$$H_{d,\beta\gamma} = H_d \left\{ \left[\frac{(H - H_d)}{H_0} \right] R_b + \frac{1}{2} (1 + \cos\beta) \left[1 - \frac{(H - H_d)}{H_0} \right] \right\} \quad (119)$$

donde: H es la radiación global diaria sobre superficie horizontal

($\text{MJ}\cdot\text{m}^{-2}\cdot\text{dia}^{-1}$);

H_d es la radiación difusa diaria sobre superficie horizontal

($\text{MJ}\cdot\text{m}^{-2}\cdot\text{dia}^{-1}$);

H_0 es la radiación extraterrestre diaria sobre superficie horizontal

($\text{MJ}\cdot\text{m}^{-2}\cdot\text{dia}^{-1}$);

CAPÍTULO 2.

DATOS

2.1 INTRODUCCIÓN

La península ibérica se encuentra en el extremo suroeste de Europa. Se extiende desde 36.004° ($36^{\circ}0'15''$ de latitud norte) de Tarifa y 43.790° ($43^{\circ}47'25''$) de Estaca de Bares, 3.32° ($3^{\circ}19'19''$ de longitud este) de Cabo Creus y $-9,49$ ($9^{\circ}29'54''$ de longitud oeste) de Cabo de la Roca. La Figura 2.1 representa el mapa de la península Ibérica



Figura 2.1. Península Ibérica

http://www.ign.es/ign/img/actividades/cartografia/mapa_1M.jpg

La mayor parte de su superficie está configurada como una meseta con una altura media de 600 metros sobre el nivel del mar, con ligera pendiente hacia poniente. Su litoral es rocoso y con acantilados al norte, nordeste, noroeste y sureste, siendo más suave la mayor parte del litoral este y sur.

La situación geográfica de la península le otorga condiciones especialmente favorables para el aprovechamiento de la energía solar, como queda de manifiesto en la Figura 2.2, en la que se representa el mapa de potencial solar de Europa

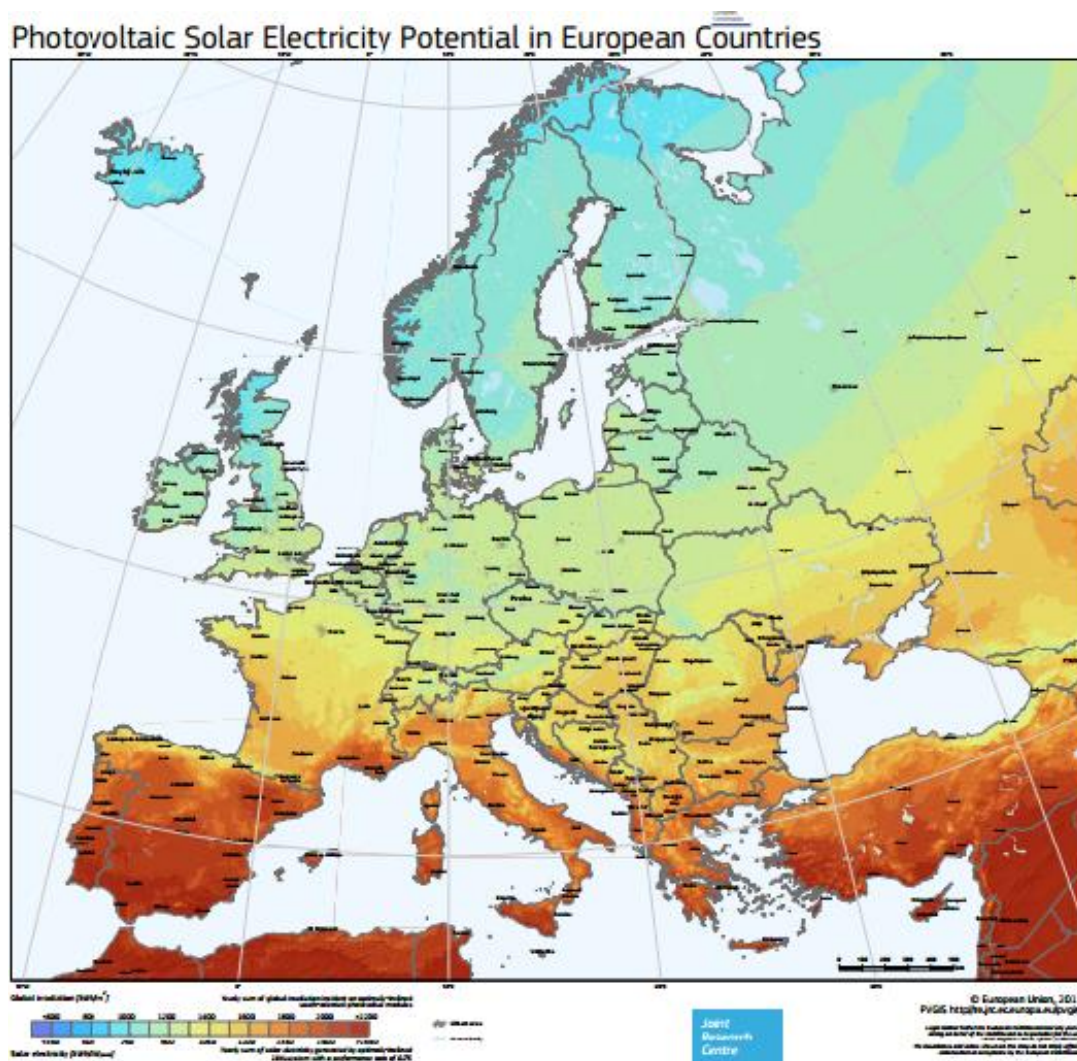


Figura 2.2. Potencial solar fotovoltaico de Europa <http://re.jrc.ec.europa.eu/pvgis/>

Dentro de la península ibérica, Castilla y León es una comunidad que se encuentra situada mayoritariamente en la submeseta Norte, en la cuenca del Duero, con pequeñas porciones de terreno situadas en las cuencas del Ebro (noreste de Burgos), del Tajo (sur de Ávila y Salamanca) y del Sil (noroeste de León). Casi se podría identificar Castilla y León con la parte española de la cuenca del Duero, ocupando prácticamente algo más del 98 % de este territorio. Ocupa una superficie de 94224 km² lo que supone un 18.7 % de la superficie nacional, configurándose como la región más extensa de España

y la tercera de Europa, y tienen una densidad de población de 27 habitantes por km².

La Meseta Norte tiene una altitud media de 750 m y está rodeada de cadenas montañosas al norte, al sur y al este. La Figura 2.3. recoge la ubicación de la Castilla y León en el mapa de la Península Ibérica

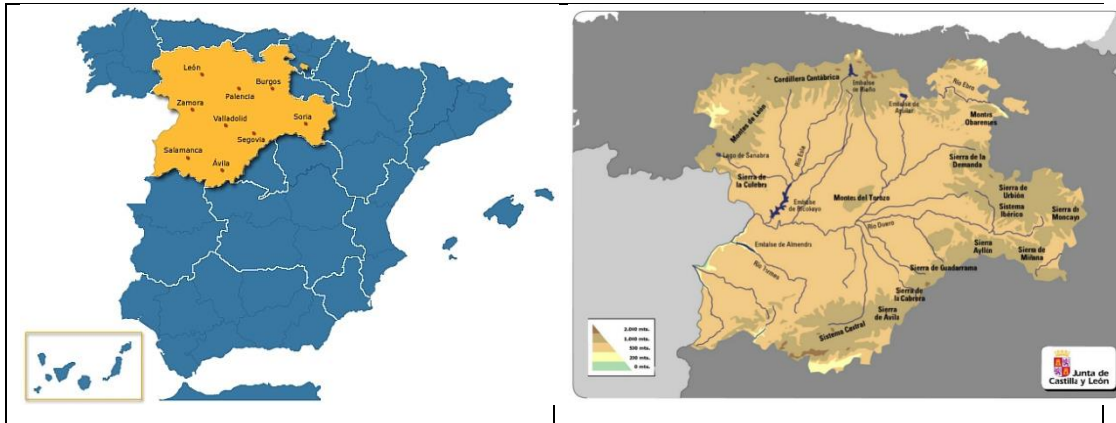


Figura 2.3. Mapa de Castilla y León dentro de la península Ibérica.

<http://navalmanzano.com/> y <http://mapasinteractivos.didactalia.net/>

El clima en la cuenca del Duero es de tipo mediterráneo, aunque marcadamente continental, debido a su aislamiento orográfico. A pesar de su cercanía al mar Cantábrico (menos de 30 km en el norte de León), la cordillera Cantábrica que la separa de él, impide el efecto moderador del mismo. Hay que tener en cuenta que esta cordillera alcanza cotas de 2560 m y es el límite norte que abarca toda la región. Solamente en la parte más occidental de Castilla y León, en las proximidades a la frontera portuguesa, el clima es ligeramente más suave, ya que la ausencia de cadenas montañosas permite la influencia del Océano Atlántico. En cuanto a la pluviometría, la zona perimetral (cordillera Cantábrica, Ibérica y Central) es la que recibe mayor precipitación, disminuyendo a medida que nos alejamos de las montañas. Las precipitaciones se suelen producir entre otoño y primavera y son raras en los meses de julio y agosto. De todas formas, es un fenómeno muy irregular en el tiempo, habiendo mucha diferencia de unos años a otros.

Castilla y León tiene un clima mediterráneo continentalizado, de inviernos largos y fríos con temperaturas medias entre 4 °C y 7 °C en enero. Los veranos no son extremados, siendo pocos los lugares en los que la temperatura media de las máximas del mes más cálido, julio, sobrepasa los 31 °C. Se puede afirmar que los veranos son cortos y calurosos (temperaturas medias entre 19 °C y 22 °C). Los meses de verano se caracterizan por su aridez, como corresponde a un clima mediterráneo. La pluviosidad es escasa acentuándose en la zona baja. Debido a la barrera montañosa, los vientos se frenan, deteniendo las precipitaciones y haciendo que las lluvias caigan de manera muy desigual.

Se pueden distinguir tres zonas climáticas diferentes. Al norte, en la parte más elevada de la cordillera cantábrica, hay un clima atlántico, de suaves inviernos y veranos templados. Cuando el terreno es menos alto se aprecian zonas típicamente atlánticas de montaña con inviernos muy fríos. La parte central de la meseta está dominada por el clima mediterráneo continental, con veranos caluroso e inviernos especialmente severos, excepto la parte este de Zamora, dominada por un clima mucho más seco. En las zonas montañosas del nordeste, el este y el sur, el clima es típicamente mediterráneo de montaña, con lluvias poco abundantes, veranos calurosos e inviernos fríos.

Si tenemos en cuenta que la Meseta limita al este con la cordillera Ibérica, que la separa del valle del Ebro, y que el macizo Central divide la meseta en las dos submesetas, con la particularidad de que la altitud media de la submeseta Norte es mayor que la de la submeseta Sur (750 m frente a 600-700 m) es posible que la altura sobre el nivel del mar influya tanto o más que el resto de las características geográficas y climáticas en la radiación solar recibida en Castilla y León.

Su localización geográfica, entre los 40.083° y 43.233° de latitud norte, conlleva una gran diferencia entre la duración del día en verano e invierno, pasando de poco más de nueve horas en diciembre a algo más de 15 en junio, siendo este uno de los motivos por lo que la irradiación media es mucho menor en invierno que en verano.

La red AEMET está compuesta de 59 estaciones distribuidas por toda la geografía nacional, de las que 24 son estaciones radiométricas principales (que miden radiación global, directa y difusa) y 22 estaciones sólo miden radiación global. Estas estaciones están dotadas de piranómetros termoeléctricos Kipp-Zonen modelos CM-11 y CM-21, con un rango espectral de 305-2800 nm y con una incertidumbre de ± 2 % para la medida de radiación global y difusa, pirheliómetros para la medida de radiación directa, pirgeómetros para la radiación infrarroja y piranómetros de banda ancha para medida de radiación ultravioleta. Estos instrumentos son calibrados cada dos años por el *Centro Radiométrico Nacional* (RRN) de la Agencia Estatal de Meteorología. La calibración de los piranómetros y pirheliómetros de la RRN se lleva a cabo por comparación al sol frente a patrones secundarios calibrados en el PMOD/WRC (*Physikalisch - Meteorologisches Observatorium Davos, World Radiation Center*) que trasladan la referencia radiométrica WRR (*World Radiometric Reference*) del grupo de mundial de normalización WSG (*World Standard Group*).

Para realizar el mapa de radiación global de Castilla y León es necesario disponer de un número suficiente de datos de estaciones de superficie. El número de estaciones de la AEMET se reduce a seis en la región en estudio y se dispone de datos de radiación global únicamente en las estaciones de León, Salamanca, Soria y Valladolid. Por esta razón, resulta imprescindible para realizar el trabajo propuesto encontrar otra fuente de datos más amplia para realizar el mapa.

2.3 DATOS DE LA RED SIAR

2.3.1 DATOS DE LA REGIÓN

La Comisión Europea aprobó en 1998 el Programa Operativo INTERREG II-C: Lucha contra la sequía, con el objetivo de apoyar acciones de cooperación transnacionales en el campo de la ordenación del territorio entre regiones de Francia, España y Portugal. En ese marco, el antiguo Ministerio

de Agricultura, Pesca y Alimentación creó la red nacional SIAR (*Sistema de Información Agroclimática para el Regadío*), que tenía como objetivo principal la mejora de la eficiencia en el uso del agua en los regadíos a través de un mejor conocimiento de las necesidades hídricas de los cultivos. A partir de entonces se mantiene y actualiza la red permitiendo obtener información de datos agroclimáticos necesarios para la gestión del agua en zonas de regadío.

La red SIAR, a nivel nacional, se estructuró mediante la creación de un Centro Nacional, 12 Centros Zonales (uno en cada comunidad autónoma adherida a la red) y Estaciones Agroclimáticas distribuidas por esas comunidades.

El Centro Nacional, que está situado en Madrid en dependencias del Ministerio, se encarga de la recepción de la información recogida diariamente en los Centros Zonales para su explotación, divulgación y almacenamiento, así como de la coordinación del Sistema. Cada uno de los Centros Zonales, está dotado de los medios informáticos necesarios para la adquisición, almacenamiento y explotación de los datos obtenidos diariamente y de forma automática por las estaciones instaladas en la Comunidad Autónoma. Las Estaciones Agroclimáticas están repartidas por toda la geografía española; hay 468 estaciones que capturan los datos climáticos y transmiten la información al Centro Zonal. De estas estaciones, 361 pertenecen al Ministerio y 107 son propiedad de las Comunidades Autónomas.

Cada estación agroclimática está diseñada para registrar, de modo automático, los valores de temperatura y humedad del aire, velocidad y dirección del viento, radiación solar y precipitación, necesarios para el cálculo de la evapotranspiración del cultivo de referencia (*ET_o*).

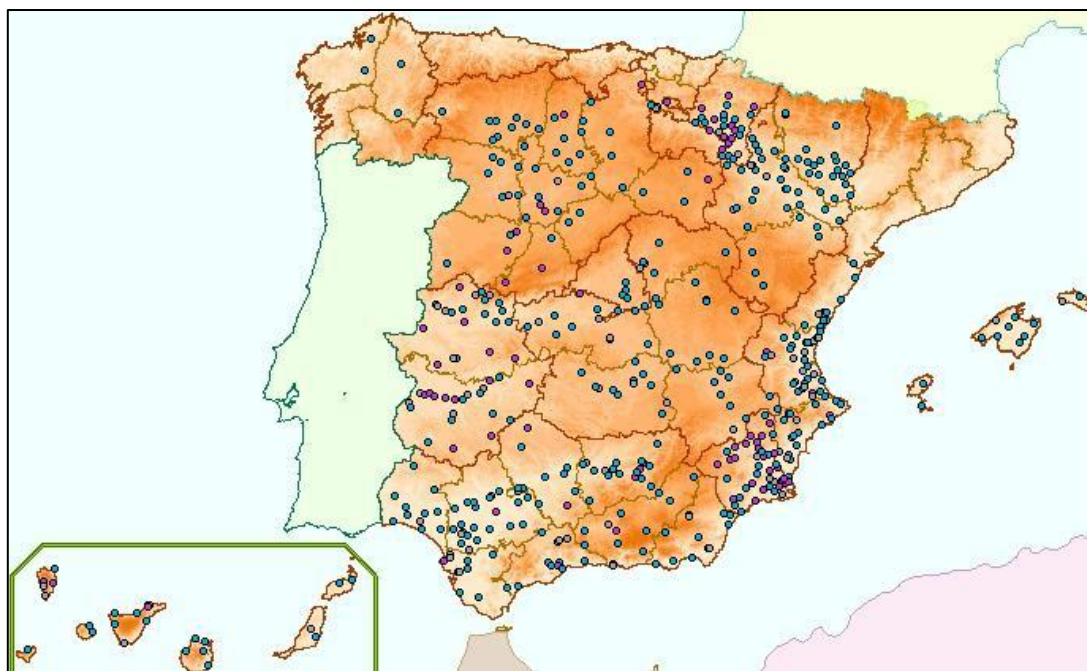


Figura 2.5. Distribución de estaciones de la red SIAR en España.
<http://eportal.magrama.gob.es/websiar>

Cada estación está dotada con una unidad central, que está formada por un datalogger (Campbell CR10X o Campbell CR1000), un módem y un regulador de carga. Esta unidad es la encargada del control de los sensores, registro y transmisión de datos y control de la alimentación eléctrica del sistema. Conectado al regulador de carga, un panel solar recarga la batería de 12 V que sirve de alimentación tanto del datalogger como del módulo de comunicaciones.

En un trípode, que les sirve de soporte, están situados los sensores de temperatura y humedad, sensor de viento (anemo-veleta) y el sensor de radiación. El sensor de radiación es un piranómetro SKYE SP1110, que es una fotocélula de silicio sensible a la radiación entre 350 y 1100 nm montada sobre una base de nivelación y orientada al sur. Su rango de medida es de 0 a $1370 \text{ W}\cdot\text{m}^{-2}$ con una corrección del coseno de 0 para 0° - 70° y menor del 10 % para 85° - 90° . El rango de señal es 1 mV por $100 \text{ W}\cdot\text{m}^{-2}$. La temperatura de trabajo va de $-35 \text{ }^\circ\text{C}$ a $75 \text{ }^\circ\text{C}$ y una precisión de $\pm 5 \%$. El tiempo de muestreo de los piranómetros es de 10 segundos.

Para la ubicación de las estaciones agroclimáticas se han considerado las recomendaciones de la *Organización Meteorológica Mundial* (OMM) la

Agencia Estatal de Meteorología, la Organización para la Agricultura y la Alimentación (FAO), y la *American Society of Agricultural Engineers* (ASAE), teniendo en cuenta que la elección de la ubicación de una estación es de primordial importancia, puesto que de ella depende la representatividad de los datos registrados en la misma. Se han buscado emplazamientos en zonas abiertas, con libre circulación del viento, representativas de la elevación media y de las características agrícolas, así como la ausencia de obstáculos en los alrededores (árboles, arbustos edificios, etc.) y, de existir, deberán estar situados a suficiente distancia para que no proyecten sombras durante la mayor parte del día, aunque sea posible que pueda haber alguna sombra a amanecer o al anochecer.

Las estaciones son sometidas a dos tipos de mantenimiento, uno preventivo (semestral), que en el caso de los piranómetros se realiza mediante la comparación con un sensor patrón, y una calibración anual, que tiene como finalidad la verificación de la precisión de las medidas registradas por los sensores. La calibración de los piranómetros se realiza siguiendo las recomendaciones de la norma "*International Standard ISO 9847 Solar energy – Calibration of field pyranometers by comparison to a pyranometer, primera edición, 1992, International Standardization Organization*", mediante comparación de los valores registrados por el sensor con los registrados por un sensor de referencia de termopila, disponiéndose de otro sensor de termopila para el contraste de datos del sensor de referencia. Se realiza bajo condiciones de luz natural con el fin de simular las mismas condiciones ambientales en las que trabajarán en campo, tomando datos de radiación siempre y cuando el ángulo solar sea superior a 20° y la radiación global superior a $300 \text{ W}\cdot\text{m}^{-2}$. Además, si se detectan anomalías, se realiza un mantenimiento correctivo, de forma que a lo sumo, en cinco días, se solventa cualquier eventualidad.

Cada estación genera tres tipos de registros (un registro semi-horario y dos registros diarios) que son recogidos en un único archivo ASCII. Los registros semi-horarios que se generan cada 30 minutos contienen medias y totales semi-horarios en todos los parámetros recogidos por SIAR. El tiempo de

muestreo de radiación, precipitación, velocidad y dirección del viento son 10 segundos y los de temperatura y humedad 10 minutos.

Los registros diarios tipo 1 presentan la recopilación de la información recogida durante el día (máximas, mínimas, totales, etc.). Los registros diario tipo 2 recogen el tanto por uno que ha estado el viento en calma durante el día, así como el tanto por uno que la temperatura ha estado en distintos tramos de temperatura.

Los valores semi-horarios de irradiancia se presentan en $W \cdot m^{-2}$ y corresponden a la irradiancia solar global media en el período, calculada a partir de medidas de irradiancia realizadas cada 10 segundos; Los valores diarios se presentan en $J \cdot m^{-2}$ y corresponden a la irradiación solar global acumulada a lo largo del día.

Una vez recibidos los datos se realiza una validación de los mismos basado en la *Norma UNE 500540:2004 "Redes de estaciones meteorológicas automáticas: Directrices para la validación de registros meteorológicos procedentes de redes de estaciones automáticas."* de forma que los valores obtenidos tiene que cumplir una serie de condiciones, como por ejemplo que la radiación horaria tiene que tomar valores entre 0 y $1350 W \cdot m^{-2}$, y la diaria tiene que tomar valores entre 0 y $35 MJ \cdot m^{-2}$, no dándose por bueno ningún dato que no esté en esos rangos.

Para obtener el mapa de radiación de Castilla y León se han localizado 56 estaciones de la red SIAR situadas en la región. De ellas 46 pertenecen al Ministerio y 10 a la Comunidad Autónoma. Los datos de estas estaciones son de acceso libre en Internet a través de la página web *Inforiego* de la Junta de Castilla y León, pudiéndose descargar los datos diarios por estación y período de tiempo en formato *.csv, y los datos semi-horarios se pueden pasar a formato Excel después de descargarlos y copiarlos.

Se han obtenido los datos diarios y semi-horarios de siete años, desde 2007 a 2013, y se ha realizado un primer cribado, en el que se han desechado las estaciones de las que no se dispone de al menos cinco de los siete años válidos, entendiendo como año válido aquel que dispone de más de 335

valores diarios. Después, se ha realizado un segundo cribado, comparando, con ayuda de una subrutina programada en MATLAB, los valores de radiación diarios obtenidos directamente a partir del fichero *.csv de datos diarios, con los valores diarios obtenidos como suma de las radiaciones globales obtenidas para intervalos de 30 minutos mediante integración (método de los trapecios (*Conte y De Boor, (1972)*)) de los valores semi-horarios de radiación global media.

Teniendo en cuenta que los valores de radiación global diaria deben ser iguales a la suma de los valores semi-horarios, se han desechado aquellos días en los que la diferencia de los valores de radiación global obtenidos directamente y los obtenidos mediante integración a partir de los datos 30 minutos, no están comprendidos en el intervalo $[-0.02, 0.02]$ MJ·m⁻².

Con el segundo cribado de datos, la red disponible ha quedado reducida a 46 estaciones. Se ha prescindido también de otras dos estaciones, ya que no cumplían la norma de que los valores de un mismo parámetro medidos al mismo tiempo en estaciones cercanas no pueden diferir demasiado unos de otros (validación espacial *UNE 500540:2004*), considerando estaciones situadas espacialmente a distancia inferior a 50 km de la desechada.

Tras aplicar todos los controles anteriores, se han seleccionado 44 estaciones de la red SIAR de las que se disponen de datos diarios de radiación global de al menos cinco años con, al menos, 335 datos al año. De las 44 estaciones elegidas, en dos de ellas se disponen de cinco años de datos, de diez estaciones se disponen de seis años de datos y del resto de estaciones (32) se disponen de siete años de datos. Es decir se ha considerado aptas para ser utilizadas en el estudio, más de 110000 datos diarios de radiación global en la región.

Se ha calculado la radiación global media diaria de cada estación en cada uno de los años de 2007 a 2013 y se ha obtenido la media de esos siete valores, asignando finalmente a cada estación un único valor como representación de su radiación global diaria anual.

2.3.2. DATOS DE LAS ZONAS CIRCUNDANTES

Una vez escogidas las estaciones de la región consideradas idóneas para la elaboración de su mapa de radiación, se ha advertido de la necesidad de disponer de valores de radiación en las zonas limítrofes exteriores, pues su ausencia supone que, al emplear cualquier método de interpolación espacial, a las zonas de las que no se tienen datos se les asigna valores nulos de radiación, imponiendo una fuerte distorsión al estudio. Se han buscado datos de las zonas limítrofes y se han empleado los siguientes conjuntos de datos:

Galicia: Consultada la página web de *Meteogalicia* de la *Conselleria de Medio Ambiente, Territorio e Infraestructuras de la Xunta de Galicia*, se ha accedido a los datos meteorológicos de la región, buscando estaciones en las provincias limítrofes a Castilla y León, próximas a ella y que tengan datos de los años 2007 a 2013 en las mismas condiciones que se ha exigido a las estaciones de la red SIAR. En cada estación se puede elegir el tipo de consulta: datos diezminutales, datos diarios y datos mensuales. Los datos se presentan con el correspondiente código de validación: 0: Datos sin validar, 1: Dato válido original, 2: Dato sospechoso, 3: Dato erróneo, 4: Dato acumulado, 5: Dato válido interpolado, y 9: Dato no registrado. Se ha elegido utilizar los datos diarios, en los que la radiación global diaria se presenta en $10 \text{ kJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$. Se han escogido dos estaciones de Lugo y tres de Orense para incorporarlas a la base de datos necesaria para los cálculos.

Asturias: En el *Boletín Solar del Principado de Asturias*, de la página web de la *Federación Asturiana de la Energía* están recogidas las medias mensuales y anuales de radiación de los años 2007 a 2010 de 11 estaciones repartidas por el Principado de las que se ha elegido tres como más cercanas en características geográficas a la zona norte de Castilla y León.

Cantabria: en la página Web www.airecantabria.com de la Consejería de Medio Ambiente, Ordenación del Territorio y Urbanismo del Gobierno de Cantabria, se tiene acceso a los datos históricos de 11 estaciones fijas y una móvil con datos quince-minutales, horarios y diarios, de distintos parámetros ambientales, con la opción de descargar datos temporales o datos validados.

Se han elegido dos estaciones con datos diarios de radiación global, validados y expresados en $W \cdot m^{-2}$.

Euskadi: En la página www.euskalmet.euskadi.net de la *Agencia Vasca de Meteorología* se tiene acceso a los informes meteorológicos desde el año 2001 hasta 2014, en los que figuran distintos parámetros climatológicos para las 79 estaciones de la región. En cada año se presentan los valores medios mensuales y anuales de radiación global expresados en $MJ \cdot m^{-2}$. Se han escogido tres estaciones de las que una es de la Agencia Vasca de Meteorología, pero está enclavada en el norte de la provincia de Burgos (Ordunte).

La Rioja: El Gobierno de la Rioja dispone de 17 estaciones agroclimáticas que ponen sus datos en abierto a través de la página web *SOS Rioja*. Se tiene acceso a los datos diarios desde 2001 hasta la actualidad. La radiación global se presenta en $W \cdot m^{-2}$, y se han escogido tres estaciones para disponer datos al noreste de Castilla y León.

Aragón: De las 21 estaciones que la red SIAR tiene en Zaragoza se ha elegido una para completar la información del extremo este de la región.

Castilla la Mancha: De las estaciones de la red SIAR de Castilla-la Mancha se ha elegido dos estaciones en Guadalajara y una en Toledo, próximas a Castilla y León que han servido para delimitar el sur de la región junto con Madrid y Extremadura.

Madrid: Del servidor de la estación meteorológica del *Instituto de Energía Solar (UPM)* situado en la ciudad universitaria se ha obtenido los datos diarios de radiación solar de los años de referencia para utilizarlos en la obtención del mapa de radiación. Los datos horarios de radiación global horizontal se presentan en formato $W \cdot m^{-2}$ y recogen la media de las medidas correspondientes a los 60 minutos en torno a una hora determinada.

Extremadura: La red *REDAREX* englobada en la red SIAR dispone en Extremadura de 37 estaciones información al regante que proporciona datos agroclimáticos. Se han elegido tres estaciones que tienen datos en los años

2007 a 2013, y que son próximas y con características similares al sur de la comunidad.

Portugal: a partir de la página web del *Laboratorio Nacional de Energia e Geologia de Portugal (LNEG)* (www.lneg.pt/servicos/328/2263/) tomando los valores de radiación de un año tipo en tres regiones portuguesas: Beira Interior Norte, Alto Tras os Montes y Douro, se han asignado datos de radiación en tres puntos de Portugal que completan las regiones fronterizas con Castilla y León

Como resumen, para obtener el mapa de radiación global en plano horizontal de Castilla y León se han utilizado los datos de 71 puntos de datos de radiación global, 44 de los cuales pertenecen a la región y 27 son externos a ella.

2.4 DATOS DE LA ESCUELA POLITÉCNICA SUPERIOR DE BURGOS

Uno de los objetivos a conseguir en este trabajo consiste en calcular la energía producida por una instalación fotovoltaica integrada en una fachada. Un problema muy generalizado al hacer los estudios correspondientes es que no se dispone de valores de irradiancia global sobre superficie vertical. Existen estaciones radiométricas, como se ha comentado en los dos apartados anteriores, pero ninguna dispone de sensores verticales que puedan medir el valor de la irradiancia para esa inclinación. Pero sí existen modelos matemáticos que permiten calcular los valores de irradiancia para superficies inclinadas con cualquier ángulo de inclinación, por lo tanto también para superficies verticales, a partir de datos de irradiancia global horizontal. Dependiendo del modelo utilizado se requiere también las componentes, difusa y directa de la irradiancia.

Ahora bien, si lo que se quiere es conocer el valor de la irradiancia global vertical en un punto, los modelos solo permiten obtener aproximaciones y estas no se pueden determinar, si no se dispone de datos experimentales en

ese punto. Esta es la razón por la que el grupo de investigación SWIFT (*Solar and Wind Feasibility Technologies*) ha implementado dos estaciones radiométricas, una en la Escuela Politécnica Superior de la Universidad de Burgos y otra en *el Centro para la Investigación y el Desarrollo de la Energía Solar* (CIDESOL) en Astudillo (Palencia). Ambas estaciones meteorológicas constan de varios sensores de radiación, entre ellos piranómetros montados en una torre que miden irradiancia global vertical en las cuatro direcciones (norte, sur, este y oeste). Para este trabajo se va a utilizar los datos recogidos en la estación de la EPS de Burgos.

Como alguno de los modelos matemáticos de radiación emplean valores de irradiancia difusa y directa para poder obtener el valor de irradiancia global en superficie inclinada se necesitan sensores para medir esos parámetros. La estación de Burgos dispone de un seguidor solar que mide la irradiancia solar directa, difusa y global. Para completar la estación se han añadido otros sensores de variables climáticas y el sistema de adquisición de datos, con el objetivo de poder realizar futuros estudios complementarios del que aquí se presenta.

2.4.1 DESCRIPCIÓN DE LA ESTACIÓN

Los datos con los que se trabaja se obtienen de la estación radiométrica ubicada en Burgos, más concretamente en la azotea de la Escuela Politécnica Superior de la Universidad de Burgos Campus del Vena (42.351°; -3.69°; 857 m s.n.m.). Se trata de un edificio de cinco alturas con una cubierta plana y de acceso restringido. Debido a la altura de la azotea y a estar situado en una zona despejada de edificios, la influencia del entorno con la instalación es mínima, puesto que sólo existen edificaciones con una mayor altura al sur y oeste. Estos edificios no generan interferencias importantes en las medidas ya que, debido a la distancia a la que se encuentran, el ángulo del horizonte tiene una elevación menor a 10° respecto de la azotea donde está emplazada la estación de referencia. En la Figura 2.6 se representa una vista satélite de la azotea del edificio donde está ubicada la instalación.



Figura 2.6. Emplazamiento de la instalación de medida. Escuela Politécnica Superior de Burgos. (Google Maps)

La instalación cuenta actualmente con un total de 17 sensores que nos permiten disponer tanto de valores climáticos (temperatura, viento (velocidad y dirección), pluviometría, presión y humedad) y valores de radiación solar en diferentes inclinaciones, componentes y direcciones. La instalación fue puesta en funcionamiento a finales del año 2011, con dos piranómetros, una sonda de temperatura y un datalogger, adquiridos gracias a los fondos conseguidos a través del proyecto de investigación de la Junta de Castilla y León (BU019A08) con una dotación de 11100 €.

En el año 2012, el grupo de investigación SWIFT obtiene fondos del Ministerio de Ciencia e Innovación (MICIIN), en la Convocatoria del Programa Nacional de Proyectos de Investigación Fundamental mediante el Proyecto ENE2011-27511, dotado con 121000 €, que permite completar la estación de medida: se añaden sensores de velocidad y dirección de viento, cuatro piranómetros para obtener datos de irradiancia vertical (N, S, E y O), seguidor con tres piranómetros (global, directa y difusa). A mediados del 2014, como consecuencia de una tesis desarrollada en el grupo, surge la idea de montar un sistema para la medida de radiación difusa en las cuatro direcciones. El desarrollo de la idea ha conducido a la solicitud de una

patente de la estructura soporte de dicho sistema de medida de irradiancia difusa.

Desde el primer momento del diseño de la instalación, se han tenido en cuenta las recomendaciones y normativas de la OMM. Dicha documentación ha sufrido actualizaciones con el paso de los años pero, incluso después de las numerosas modificaciones, la instalación sigue cumpliendo las indicaciones especificadas en los documentos de OMM, 2013. En la Figura 2.7 se muestran algunos de los sensores de la instalación en la actualidad.

Como se ha mencionado anteriormente, la instalación cuenta con diferentes sensores climatológicos y radiométricos, por ello, para la descripción de la misma se ha dividido la instalación en dos subestaciones: subestación climatológica y subestación radiométrica.

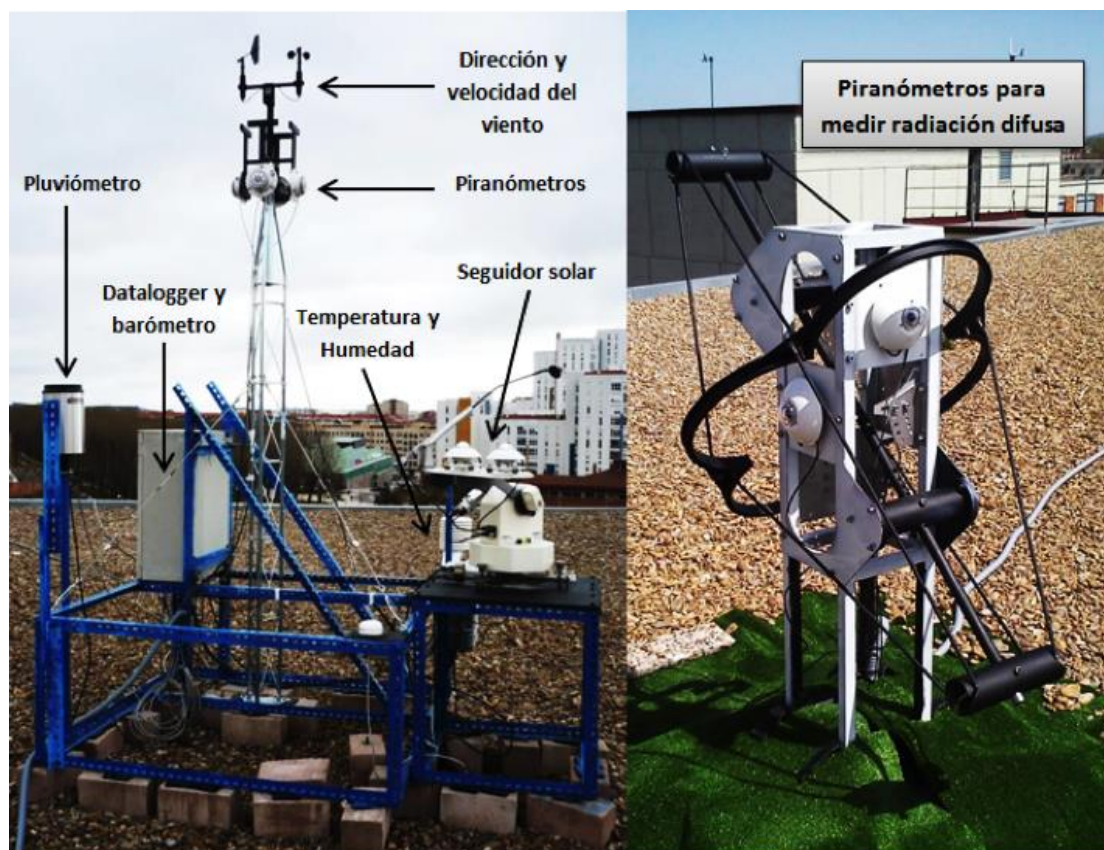


Figura 2.7. Vistas de la estación radiométrica del grupo SWIFT ubicada en la EPS (Burgos)

Subestación climatológica:

Esta parte de la instalación es la que se ocupa del registro de los diferentes parámetros climatológicos del entorno de Burgos. Cuenta con un total de seis sensores que generan registros de las seis variables climáticas: dirección y velocidad de viento, temperatura ambiente, humedad relativa, pluviometría y presión atmosférica. En la Tabla 2.1 se presenta una descripción de las características fundamentales de los diferentes sensores.

Tabla 2.1. Relación de sensores climáticos.

| Tipo de sensor | Modelo | Parámetro registrado | Unidad |
|---------------------------------------|---------------------------|---|----------------------------|
| Pluviómetro | RM YOUNG/52203 | Volumen de precipitación. | (mm.m ⁻²) |
| Anemómetro | RM YOUNG/03101 | Dirección y velocidad del viento. | (°) - (m.s ⁻¹) |
| Barómetro | Campbell Scientific/CS106 | Presión atmosférica. | (mmHg) |
| Sonda de temperatura y humedad | VAISALA/HMD50U/Y | Temperatura ambiente y humedad relativa del aire. | (°C) - (%) |

Con el pluviómetro (RM YOUNG/52203) se registra el volumen de la precipitación acumulada, tanto horaria como diaria. El sensor utilizado para ello es de tipo balancín (contador de pulsos), que proporciona un pulso en el circuito eléctrico cada 0.1 mm de precipitación y cuenta con un área de captación de 200 cm². Este sensor dispone de una pequeña resistencia eléctrica que permite realizar registros de la variable incluso para precipitaciones de nieve, ya que se ocupa de fundirla y poder así cuantificarla. La precisión del sensor es de un 2 % hasta 25 mm·h⁻¹ y de un 3 % hasta una intensidad de precipitación de 50 mm·h⁻¹.

El anemómetro (RM YOUNG/03101) permite determinar la dirección y velocidad del viento. Para la medida de la dirección, el sensor integra una veleta que nos proporciona la dirección del viento con una precisión de ±5° y el rango de medida es de 352°. Por otro lado también integra un anemómetro de cazoletas semiesféricas que permite el registro de valores con un rango de 0-50 m·s⁻¹ y una precisión de ±0.5 m·s⁻¹. Para evitar interferencias con otros sensores y con la propia estructura de la estación se

ha emplazado en la parte superior de la misma. En la Figura 2.8 se muestran los dos sensores descritos.



Figura 2.8: Veleta, anemómetro (Izquierda) y Pluviómetro (derecha)

Otro parámetro fundamental en la caracterización climática es la temperatura ambiente y la humedad relativa. Para la medida de la temperatura ambiente se ha dispuesto una sonda de temperatura de tipo Pt-1000 (VAISALA/HMD50U/Y) Este sensor tiene un rango de medida que abarca desde los $-40\text{ }^{\circ}\text{C}$ hasta los $60\text{ }^{\circ}\text{C}$ con una precisión de $\pm 0.3\text{ }^{\circ}\text{C}$. También se dispone de un sensor de humedad relativa, que puede realizar medidas con una precisión de $\pm 2\%$, y cuenta con un rango de medida comprendido entre 10 y 90 %. Para asegurar su buen funcionamiento, ambos sensores se han colocado dentro de una carcasa que permite una correcta ventilación natural y al mismo tiempo los protege de la radiación solar.

Otro parámetro climático registrado es la presión atmosférica. Para esta medida se emplea un barómetro electrónico (Campbell Scientific/CS106). Se trata de un sensor robusto que ofrece medidas con una precisión de $\pm 1.5\text{ mb}$ para el rango de temperaturas de $-40\text{ }^{\circ}\text{C}$ a $60\text{ }^{\circ}\text{C}$. Cuenta con un rango de medida que abarca desde los 500 mbar hasta los 1100 mbar, suficiente para emplazamiento en el que se encuentra colocado. Debido a que el sensor no cuenta con la certificación IP adecuada para poderse colocar a la intemperie, se ha ubicado en el interior del armario eléctrico, donde están también el

datalogger y otros equipos eléctricos. Ambos sensores se muestran en la Figura 2.9

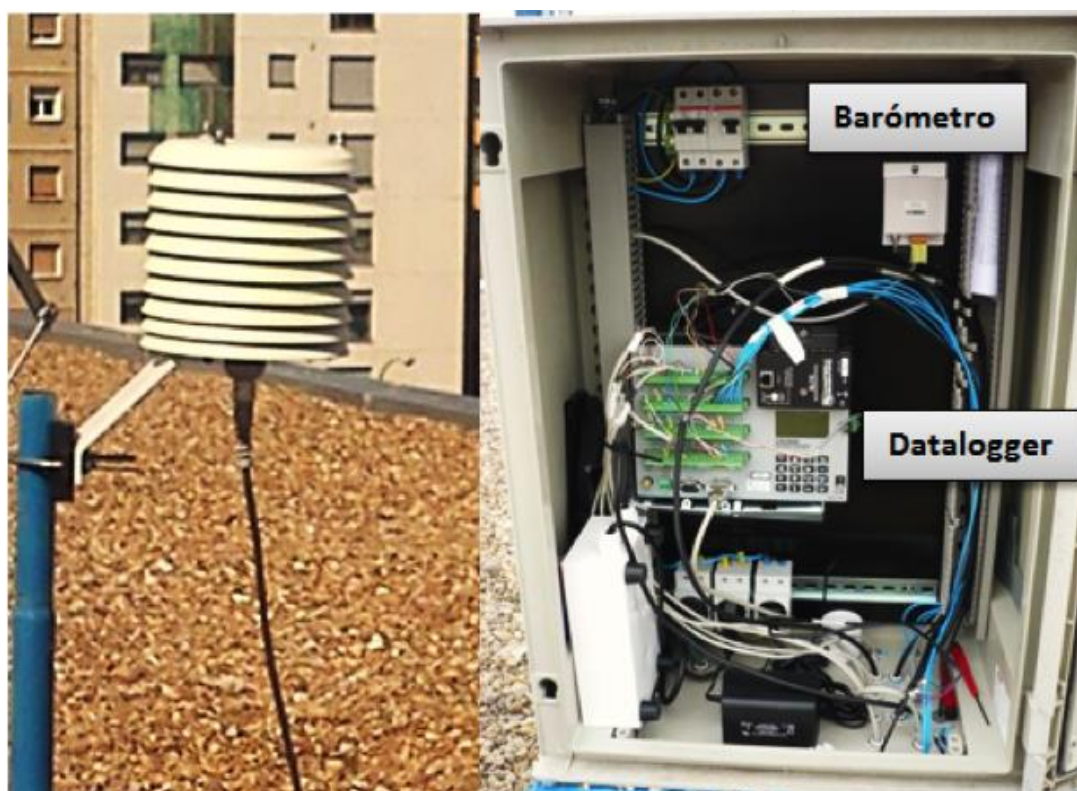


Figura 2.9: Izquierda: Sensor de temperatura y humedad relativa. Derecha: Barómetro.

Subestación radiométrica:

En este campo la estación se ocupa del registro de diferentes parámetros de radiación solar. Cuenta con 10 piranómetros y un pirheliómetro, todos ellos de clase 1, que registran valores de radiación en diferentes inclinaciones (vertical, horizontal o inclinada) y orientaciones (norte, sur, este y oeste). Se dispone también de un equipo con seguimiento solar que permite la medida de las diferentes componentes de la radiación solar (global, difusa, reflejada y directa). En la Figura 2.10 se representan los diferentes sensores.

En la imagen izquierda de la Figura 2.10 está el seguidor solar SunTracker-3000, que permite el registro de la irradiancia solar sobre el plano horizontal de sus tres componentes: global, difusa y directa. Para la medida de la irradiancia global, dispone de un piranómetros Hukseflux/SR11 al igual que para la medida de la irradiancia difusa, salvo que en este caso el seguidor desplaza un disco que proyecta una sombra sobre el sensor. El registro de la

irradiancia directa se lleva a cabo mediante un pirheliómetro Hukseflux/DR01 el cual se orienta de forma que sigue continuamente al sol, lo realiza automáticamente gracias el seguidor solar del dispositivo.



Figura 2.10: Subestación radiométrica. Izquierda: seguidor solar. Centro: difusa vertical MK6. Derecha: Global vertical.

Un elemento distintivo de esta instalación con respecto a las de la Red Radiométrica Nacional es la posibilidad de medir la irradiancia vertical. Concretamente, la instalación cuenta con una estructura diseñada por el grupo de investigación SWIFT que permite la medida de la radiación solar en el plano vertical en diferentes orientaciones (norte, sur, este y oeste). La imagen derecha de la Figura 2.10 muestra la estructura en forma de cruz sobre la que se colocan los diferentes piranómetros PH.SCHENK 8101 en vertical orientados en las cuatro direcciones cardinales.

Por último cabe destacar el diseño particular realizado por el grupo (tiene presentada una solicitud de patente), que mediante un anillo lobular permite sombrear los cuatro piranómetros simultáneamente. En la imagen central de la Figura 2.10 podemos ver el primer prototipo, actualmente en funcionamiento. Este dispositivo no sólo permite registrar valores de irradiancia difusa en vertical para cada una de las diferentes orientaciones, sino que cuenta con un sistema de giro en los soportes de los sensores

pudiendo hacer registros en inclinaciones de hasta 60°. En este caso los sensores son también Hukseflux/SR11.

Registro, almacenamiento y tratamiento de datos:

Tan importante como los sensores en la instalación es el dispositivo de registro de los valores medidos. Para ello se dispone de un datalogger CR3000 del fabricante Campbell Scientific. Se trata de un datalogger robusto y fiable que permite tener una buena resolución y precisión en las medidas generadas.



Figura 2.11: Datalogger Campbell Scientific CR-3000.

Este dispositivo cuenta con un total de 28 entradas analógicas simples o 14 diferenciales con una entrada de tensión configurable desde ± 20 mV hasta ± 5000 mV. La resolución de la medida, depende del rango de entrada que se haya configurado y oscila entre $167 \mu\text{V}$ y $0.67 \mu\text{V}$. También cuenta con ocho entradas digitales y dos contadores de pulsos de hasta ± 20 V.

El almacenamiento de los registros se realiza en tablas fácilmente configurables en la programación del dispositivo. Previamente hay que configurar cada equipo en función de las características del mismo y de la variable a medir. En este caso se realizan registros, de cada una de las variables descritas en párrafos anteriores, en tablas clasificadas por meses con diferentes intervalos de integración, unas almacenan valores cada 10 minutos y en otras se almacenan valores horarios.

Pese a disponer de una memoria suficiente como para registrar valores durante varios meses, se realizan descargas de datos cada uno o dos meses a lo sumo. Una vez descargados se hace un tratamiento de los

mismos, pasándoles los filtros necesarios para cumplir las condiciones de fiabilidad de la OMM y que permite descartar los valores de baja calidad o erróneos para su posterior uso. También se puede obtener una previsualización que permite detectar problemas puntuales, periódicos o desajustes en la instalación.

La imagen de la derecha de la Figura 2.9 se recoge el armario eléctrico en el que se ha colocado y conectado el datalogger junto con otros equipos eléctricos necesarios para alguno de los sensores, como los adaptadores de variables eléctricas y baterías.

El programa para la recogida de los datos se ha editado siguiendo las instrucciones del fabricante y las de cada uno de los elementos que están conectados a él. Los datos almacenados en el datalogger se recogen en tres tipos de archivos: instantáneos, horarios y diarios. En la Figura 2.12 se han representado las cuatro curvas de irradiancia vertical (norte, sur, este y oeste), a partir de los datos directamente registrados en el datalogger.

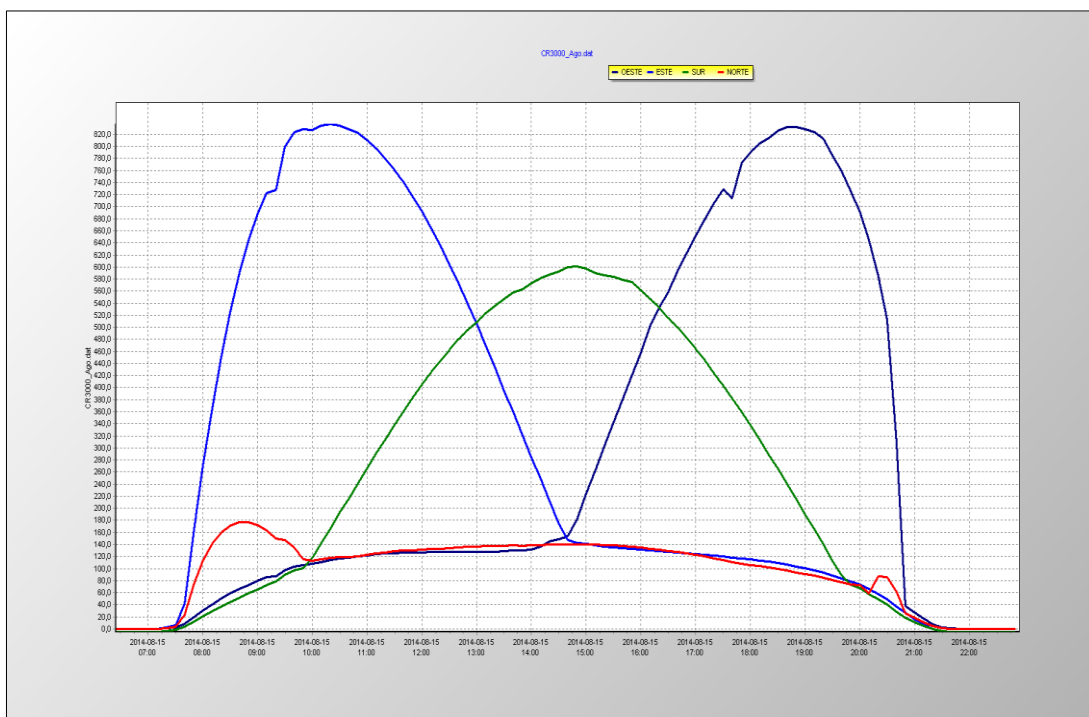


Figura 2.12. Irradiancia global vertical de las cuatro orientaciones: norte, sur, este, y oeste, para el día 15 de agosto de 2014

En la Figura 2.13 se representa las curvas de irradiancia horizontal global, difusa y directa para un día, recogidas también en el datalogger.

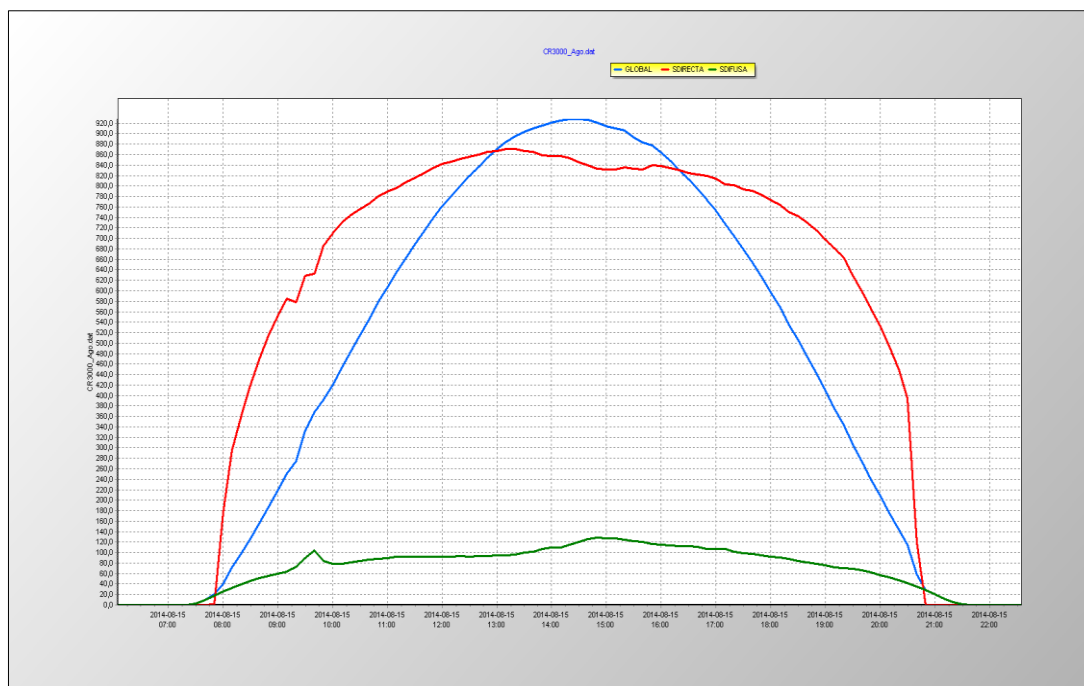


Figura 2.13. Irradiancia horizontal global, directa y difusa para el día 15 de agosto de 2014

2.4.2 CALIBRACIÓN Y MANTENIMIENTO

Las tareas de mantenimiento son un aspecto fundamental para poder garantizar el buen funcionamiento y la calidad de las medidas. Cada uno de los diferentes sensores tiene unas exigencias de mantenimiento que se especifican en los diferentes manuales. Las principales tareas que se realizan con periodicidad mensual son inspecciones visuales de los diferentes sensores y cables de conexión así como de la limpieza de los mismos. Más específico de otros sensores es la necesidad del cambio de las sales deshumidificadoras de los piranómetros cada 12 meses como máximo.

Un piranómetro ideal debería tener siempre una ratio constante, que sería el cociente entre la tensión de salida y el nivel de radiación. Esta ratio se denomina *sensibilidad*. Los valores de radiación se pueden computar simplemente dividiendo la señal de salida del piranómetro por su

sensibilidad, o multiplicando el valor de la señal por el recíproco de la sensibilidad (a la sensibilidad se le llama *factor de calibrado*).

Cada sensor tiene su curva de sensibilidad y ésta se determina en el laboratorio de fabricación por comparación con un piranómetro patrón. La sensibilidad del piranómetro cambia con el tiempo y con la exposición al sol. Una de las razones de este cambio puede ser el deterioro de la pintura negra, por lo que es necesario hacer calibraciones periódicas. La normativa establece que se debe realizar una calibración de los diferentes piranómetros cada 24 meses de uso en exterior. Se suelen realizar campañas de calibración cada 12 meses. El Grupo de Investigación dispone para la calibración, de un piranómetro patrón, el cuál se usa como referencia.

2.4.3. TRATAMIENTO DE LOS DATOS

Los valores registrados por el datalogger se leen con un programa implementado en MATLAB, que realiza tres funciones básicas: la primera consiste en construir una matriz con los datos repartidos en columnas; la segunda aplica las correcciones geométricas y totales a la irradiancia difusa y la tercera aplica el control de calidad a todos los valores medidos.

La primera labor consiste en ordenar los datos matricialmente. Se efectúa colocándolos en una matriz de 16 columnas, para el fichero de valores de radiación verticales, y el número de filas dependerá del tipo de fichero con el que se vaya a trabajar. En este caso se maneja un fichero con valores diezminutales (se necesitan para el cálculo de la irradiancia global vertical en alguno de los modelos elegidos), es decir 52560 filas en el fichero de los valores de radiación vertical de datos de un año.

Es necesario puntualizar que en este caso se trabaja con los datos en horario determinado como GMT+0, es decir, se toma como referencia el meridiano de Greenwich y se suma el tiempo correspondiente a la modificación horaria española (verano e invierno implementado por el propio datalogger), pero no se trabaja con HSV (*Horario Solar Verdadero*) como se

suele hacer cuando se trabaja con valores de irradiancia ($W \cdot m^{-2}$). Se ha decidido por esta opción ya que hay otras variables en el conjunto de datos registrados que no precisan esta conversión.

Los valores de irradiancia recogidos por el sensor de difusa no necesitan corrección de ningún tipo puesto que el elemento de sombreado es un disco y no una banda. El tamaño del disco es lo suficientemente pequeño como para conseguir tapar el sensor y lo suficientemente grande como para que la sombra generada por él dure el tiempo necesario hasta el próximo desplazamiento del sensor. Según las características del propio sensor dadas por el fabricante, este sensor no necesita corrección.

La siguiente labor consiste en seleccionar los valores válidos de todos los recogidos. Este proceso se hace siguiendo los criterios establecidos en el “*Daylight Research Group*” (Kambezidis et al., (1994)) de la Comisión Europea. Los criterios que se han aplicado en este caso son los siguientes:

- se han rechazado los valores de radiación difusa mayores que 1.15 veces al correspondiente valor de radiación global
- se han rechazado los valores de radiación global mayores que 1.2 veces la constante solar (I_{sc})
- se han rechazado los valores de radiación difusa mayores que 0.8 veces la constante solar (I_{sc})
- se han rechazado los valores de radiación global iguales o menores que $0.0007 MJ \cdot m^{-2} \cdot h^{-1}$
- se han rechazado los valores de radiación correspondientes a horas cuyo ángulo de elevación solar es menor de 5°
- se han rechazado los valores de radiación cuya componente directa, es decir aquellos en los que la $[(\text{global-difusa})/\text{senh}]$ sea mayor que el valor de la radiación extraterrestre siendo h el ángulo de elevación

Estas tareas permiten obtener una matriz de valores registrados con calidad suficiente para proceder a trabajar con sus valores en los apartados siguientes. En la Tabla 2.2 se ha recogido parte de la matriz del fichero de

datos recogidos por el seguidor SunTracker-3000. Los recogidos en el otro fichero tienen el mismo aspecto salvo el número de columnas.

Tabla 2.2. Fichero de datos con los que se trabaja

| Año | Mes | Dia | Hora | Minuto | Directa | Difusa | Global |
|------|-----|-----|------|--------|----------|---------|---------|
| 2014 | 5 | 7 | 8 | 20 | 8,74055 | 316,715 | 319,727 |
| 2014 | 5 | 7 | 8 | 30 | 24,4147 | 325,52 | 338,771 |
| 2014 | 5 | 7 | 8 | 40 | 16,7868 | 343,495 | 352,469 |
| 2014 | 5 | 7 | 8 | 50 | 6,4382 | 258,292 | 260,839 |
| 2014 | 5 | 7 | 9 | 20 | 4,79848 | 368,753 | 369,977 |
| 2014 | 5 | 7 | 9 | 50 | 5,21174 | 449,466 | 451,124 |
| 2014 | 5 | 7 | 10 | 10 | 3,38811 | 310,321 | 311,509 |
| 2014 | 5 | 7 | 10 | 20 | 5,30316 | 410,136 | 412,282 |
| 2014 | 5 | 7 | 10 | 40 | 31,4471 | 503,14 | 530,182 |
| 2014 | 5 | 7 | 10 | 50 | 11,6958 | 474,052 | 483,252 |
| 2014 | 5 | 7 | 11 | 0 | 9,85611 | 444,863 | 452,25 |
| 2014 | 5 | 7 | 11 | 10 | 2,30527 | 364,474 | 365,353 |
| 2014 | 5 | 7 | 11 | 20 | 2,48996 | 322,051 | 323,188 |
| 2014 | 5 | 7 | 11 | 30 | 0,883498 | 237,872 | 238,152 |
| 2014 | 5 | 7 | 11 | 40 | 3,15545 | 332,651 | 334,139 |
| 2014 | 5 | 7 | 11 | 50 | 2,5917 | 348,739 | 350,09 |
| 2014 | 5 | 7 | 12 | 0 | 71,2411 | 407,882 | 475,806 |
| 2014 | 5 | 7 | 12 | 10 | 9,34945 | 351,697 | 359,665 |
| 2014 | 5 | 7 | 12 | 20 | 8,55747 | 351,496 | 358,633 |
| 2014 | 5 | 7 | 12 | 30 | 9,57889 | 357,617 | 365,172 |
| 2014 | 5 | 7 | 12 | 50 | 5,88884 | 428,857 | 431,527 |
| 2014 | 5 | 7 | 13 | 0 | 50,0971 | 528,014 | 573,052 |
| 2014 | 5 | 7 | 13 | 10 | 14,2153 | 469,514 | 479,744 |
| 2014 | 5 | 7 | 14 | 10 | 3,15455 | 305,347 | 306,316 |
| 2014 | 5 | 7 | 14 | 40 | 5,38815 | 338,987 | 339,748 |
| 2014 | 5 | 8 | 8 | 30 | 14,065 | 324,835 | 331,472 |
| 2014 | 5 | 8 | 8 | 50 | 3,82325 | 327,66 | 328,158 |
| 2014 | 5 | 8 | 9 | 10 | 2,72408 | 320,576 | 320,63 |
| 2014 | 5 | 8 | 9 | 20 | 4,36681 | 320,711 | 322,011 |
| 2014 | 5 | 8 | 10 | 0 | 3,16424 | 436,099 | 436,603 |
| 2014 | 5 | 8 | 10 | 10 | 3,23128 | 455,53 | 456,04 |
| 2014 | 5 | 8 | 10 | 40 | 2,14149 | 441,434 | 441,581 |

CAPÍTULO 3.

METODOLOGÍA

3.1 INTRODUCCIÓN

Recordamos que el objetivo fundamental de la tesis doctoral reside en la obtención de valores de irradiación global media anual diaria, sobre superficie vertical, en las cuatro direcciones (norte, sur, este y oeste). Para conseguir ese objetivo, se ha trabajado con diferentes datos, procedimientos y aplicaciones. En este capítulo se describe el tratamiento realizado con las diferentes fuentes de datos, SIAR, AEMET y EPS, con el objeto de obtener un conjunto de valores de alta calidad. Los resultados obtenidos, cuando se trabaja con valores registrados, deben ser muy fiables y para ello es necesario asegurarse que cumplen los criterios exigidos por la normativa vigente en este campo. Es también necesario seleccionar de entre los valores registrados, los considerados válidos bajo diferentes criterios de calidad.

El tratamiento de los datos de irradiación solar de la red SIAR pone de manifiesto cierto sesgo en los valores. Por ese motivo se ha realizado un estudio estadístico que determine el comportamiento de los datos siguiendo tres procedimientos: a) un estudio de los valores de irradiación media anual diaria para ver la influencia de los años en el valor medio de dicha variable; b) se ha analizado la relación entre la irradiación y las coordenadas geográficas de las estaciones, como la latitud, longitud y altitud, y c) se ha estudiado la normalidad de los datos.

Para obtener los mapas de irradiación media anual diaria, aplicando métodos matemáticos de interpolación, se ha trabajado con los datos de la red SIAR junto con los de otras estaciones meteorológicas de diferentes localidades a la frontera de la zona de estudio. Para testar la validez de los

diferentes métodos empleados, y por lo tanto, la de los datos obtenidos por ellos, para cualquier punto de la región de Castilla y León, se han empleado cinco controles con estaciones de la red SIAR, y otros con las estaciones de la AEMET. Toda esta parte del trabajo se ha desarrollado con el software ArcGis.

Por último se han analizado los datos experimentales recogidos en la EPS con modelos matemáticos de radiación solar para superficies inclinadas. Es decir modelos que permiten conocer los valores de la irradiación global para una determinada inclinación a partir de datos de irradiación global, directa y difusa horizontal (modelos de trasposición o composición). En este trabajo se han empleado cuatro modelos clásicos, seleccionados en la literatura, que contemplan todas las posibilidades de la atmósfera: circunsolar, isotrópica y anisotrópica, buscando el modelo que mejor se ajusta a la zona de estudio.

3.2 TRATAMIENTO DE LOS DATOS DE LA RED SIAR

3.2.1 DATOS DE LA RED SIAR

Los datos de la red SIAR (*Sistema de Información Agroalimentaria para el Regadío*), son públicos y se pueden descargar por parte del usuario en diferentes formatos. Unas estaciones de esta red pertenecen al sistema nacional, es decir al Ministerio de Agricultura, y otras a cada una de las Comunidades Autónomas. De todas ellas, en este trabajo se han empleado las que están ubicadas en la Comunidad de Castilla y León.

La primera labor ha consistido en descargar los datos de todas las estaciones de la zona en estudio (Castilla y León de la red SIAR). Como se ha comentado en el capítulo anterior, se han recogido los valores correspondientes a siete años, en el periodo 2007 a 2013, de todas las variables climáticas medidas en cada una de las estaciones. Se ha trabajado con dos tipos de datos, y por lo tanto, dos ficheros distintos: valores diarios,

“*datdia*”, y los valores semihorarios, “*datsemhor*”. Se ha hecho así, para poder utilizar estos últimos como sistema de validación (como se verá más adelante) para los valores diarios con los que se va a trabajar.

Cuando se utilizan datos registrados por sensores, es muy importante que éstos sean fiables. Una tarea fundamental es establecer diferentes controles con el fin de asegurar la calidad de los mismos. El primer control realizado, ha consistido en ver cuantas estaciones tenían suficientes datos registrados, en el intervalo de tiempo seleccionado para el estudio (2007 a 2013). Se ha comprobado que sólo 47 estaciones de las 53 recogidas, cumplen la condición, la Figura 3.1 representa la distribución de las 53 estaciones de la red SIAR en Castilla y León. El número total de ficheros con el que se ha trabajado es de 658, estos contienen datos de temperatura, pluviometría, humedad relativa, velocidad y dirección del viento, irradiación global sobre superficie horizontal.

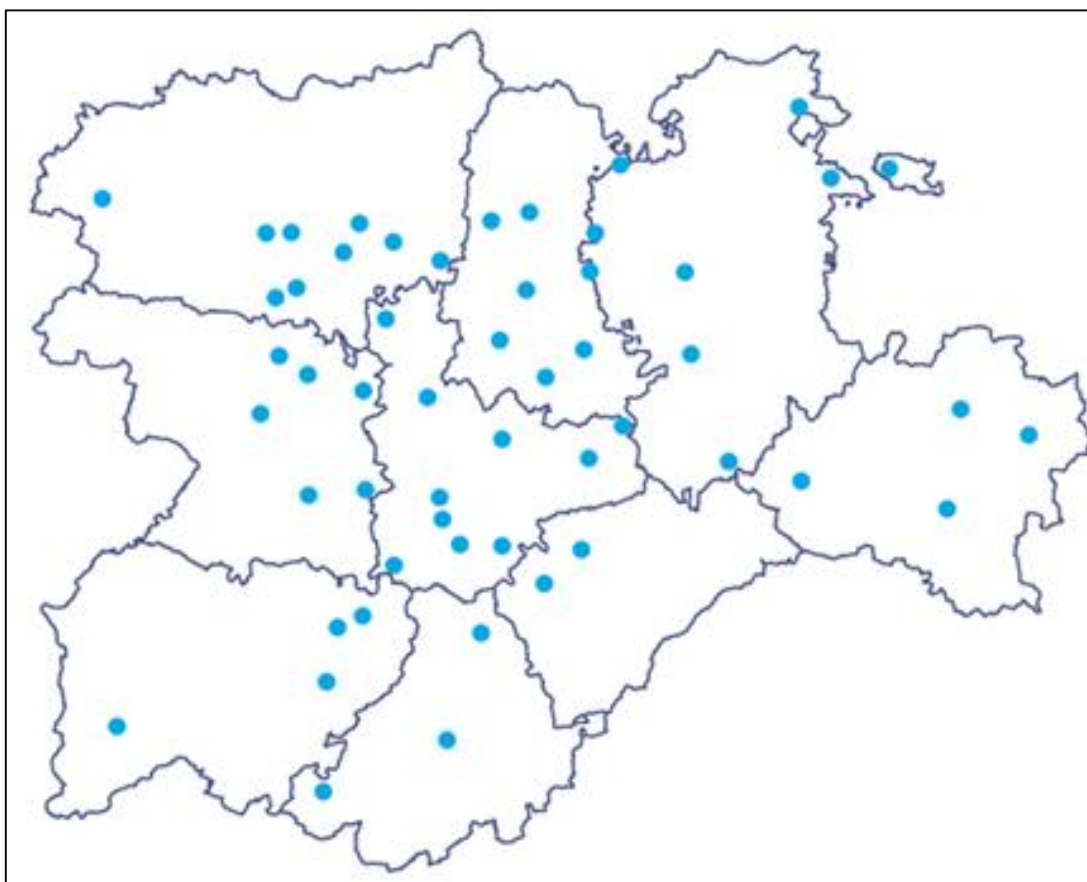


Figura. 3.1. Distribución de las 53 estaciones de la red SIAR en Castilla y León.
(Fuente: <http://www.inforiego.org/opencms/opencms>)

Como lo que se quiere es garantizar la fiabilidad de los datos diarios de irradiación global (que son con los que se va a trabajar), se calcula el valor diario de dicha variable con los datos del fichero “*datsemhor*”. Este cálculo se ha realizado por integración con el “*método de los trapecios*” (Conte y De Boor, (1972)). Para poder comparar los valores así obtenidos con los del fichero “*datdia*”, es necesario que en ambos ficheros exista el mismo número de valores diarios. Se implementa un programa en MATLAB que realiza esta función, entre otras.

El programa lee los dos ficheros y calcula, con el fichero “*datsemhor*”, el valor medio diario de irradiación global, por el método de integración de la regla de los trapecios. Este valor lo guarda con el código del día y año al que pertenece. Compara el valor calculado con el correspondiente del fichero “*datdia*”. Si alguno de los dos no existe, se elimina de los dos ficheros. Para validar el dato comparado en el fichero “*datdia*”, se ha establecido que la diferencia entre los dos no debe ser superior a $0.02 \text{ MJ}\cdot\text{m}^{-2}$. Cuando hay que rechazar el valor por no cumplir la condición, se entiende que no se dispone de elementos suficientes para garantizar la fiabilidad del dato. De este proceso se obtiene un número de estaciones válidas, cada una de ellas con siete años de medidas y cada año con un número determinado de valores diarios de irradiación global sobre superficie horizontal.

Es necesario saber, del conjunto de estaciones obtenidas, cuántas son válidas para desarrollar el estudio. Para ello, se establece como premisa fundamental que, al menos tengan cinco años disponibles. También se impone la condición para que un año sea válido que éste disponga, como mínimo, de 335 valores diarios de irradiación global. Sólo existen en la base de datos construida, 46 estaciones que cumplan estas dos condiciones (se ha eliminado una estación en este cribado). Esto confirma un adecuado tratamiento previo de los datos, por parte del personal de mantenimiento de las instalaciones, suficientemente bueno como para garantizar la calidad y fiabilidad de los datos suministrados por la red SIAR.

De las 46 estaciones con las que se va a trabajar, sólo 33 tienen datos de los siete años seleccionados para el estudio (2007 a 2013), 10 tienen datos

correspondientes a seis años del periodo y tres estaciones poseen datos de cinco años. El filtrado previo de los datos ha eliminado aquellos valores diarios que no coincidan con los calculados a partir de los datos semihorarios. El número total de valores diarios de irradiación global es de 112258, es decir se han eliminado 5364 valores, lo que representa un 4.5 %.

La subrutina generada, también nos permite obtener más información de los datos del fichero “*datdia*”; nos calcula el valor máximo y mínimo de irradiación global horizontal diaria de cada estación y el índice de claridad diario. Esta información nos será válida para tener una visión general de la distribución, tanto de la irradiación en la zona de estudio, como de la de los días claros, parcialmente cubiertos y cubiertos. En la Tabla 3.1 se recogen los valores de irradiación global horizontal máxima y mínima de todo el conjunto de datos y la fecha en la cual se ha registrado el valor.

Tabla 3.1. Valores máximos (RGM) y mínimos (RGm) de irradiación global horizontal diaria ($\text{MJ}\cdot\text{m}^{-2}$). Se indica la fecha y la estación en la que se ha recogido el valor.

| | RGM ($\text{MJ}\cdot\text{m}^{-2}$) | | RGm ($\text{MJ}\cdot\text{m}^{-2}$) | |
|----------|--|---------|--|----------|
| | Máx. | Min. | Máx. | Mín. |
| Valor | 37.73 | 29.17 | 1.14 | 0.059 |
| Estación | ZA040 | VA050 | SO101 | PA080 |
| Día | 15/5/13 | 21/6/09 | 16/1/12 | 18/10/12 |

Analizando los valores de la Tabla 3.1, se puede observar que la estación ZA040, presenta el valor más alto $37.73 \text{ MJ}\cdot\text{m}^{-2}$ (el día 15 de mayo de 2013), y el valor más pequeño (dentro de los máximos) de irradiación diaria que se ha recogido en Valladolid (VA050) con $29.17 \text{ MJ}\cdot\text{m}^{-2}$. Esta diferencia entre ellos (22.6 %) se puede deber a la variedad de las características geográficas y climatológicas de la región.

Respecto a los valores mínimos de irradiación diaria global horizontal, la estación SO101 tiene el más alto y la PA080 el más bajo. En este caso, la diferencia es mucho mayor (94.8 %), pero no se considera relevante, ya que se trata de valores mínimos. Consideramos que estos valores no son

representativos (ni de la zona ni de la irradiación recibida), y por lo tanto no reflejan la distribución de la irradiación solar en la zona, simplemente son valores puntuales y extremos.

Se ha trabajado también con valores medios por provincias para ver la distribución de irradiación media anual diaria en la zona. Se trabaja con los valores medios anuales diarios considerando que, en este tipo de medidas, el valor medio diario anual es el representativo, ya que la irradiación global horizontal en otro intervalo de tiempo menor (hora, diez minutos) es demasiado influenciado, y puede ser afectada por elementos y variables climáticas, algunas puntuales y muy significativas, como nieblas y/o nubes.

Se calculan los valores medios anuales (en cada estación considerando el total de datos válidos existentes), y luego se realiza la media de todas las estaciones de cada provincia. De esta forma se comprobará la distribución por provincias de la irradiación global media anual diaria en la región. La Figura 4.2 recoge la distribución de estos valores. Se observa que Zamora y Ávila son las que más irradiación media anual diaria reciben (hay muy poca diferencia entre ellas: 17.08 y 17.05 MJ·m⁻²) y Burgos la que menos (14.92 MJ·m⁻²). La Figura 3.2 recoge la distribución de estos valores

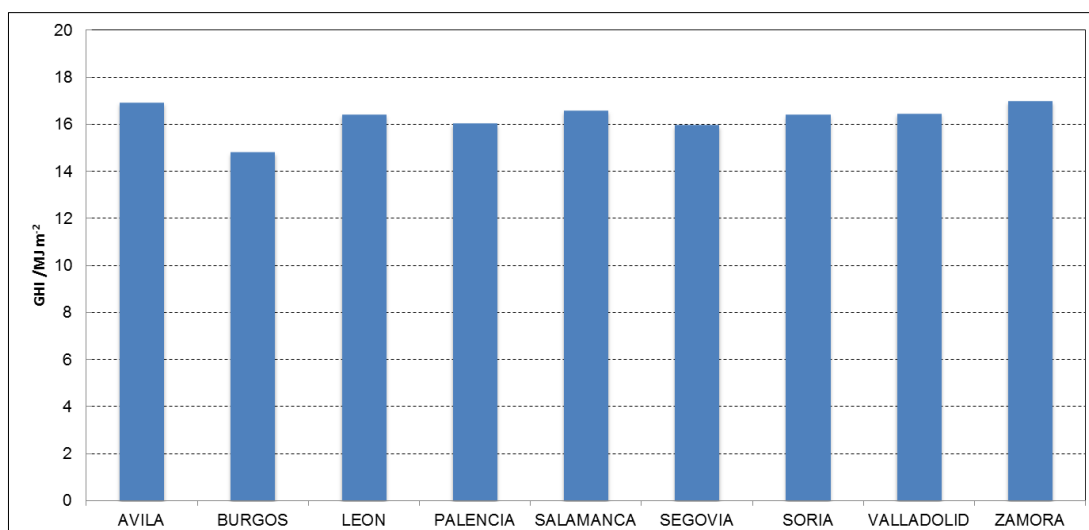


Figura 3.2. Valores medios anuales diarios de irradiación global horizontal de las nueve provincias de Castilla y León

La diferencia entre los valores mayor y menor (Zamora y Burgos) no es muy grande (12.7 %), pero si es significativa cuando se traduce a producción energética. Para una instalación tipo de 100 kW_p, la producción anual en ambas provincias presenta una diferencia aproximadamente de 18700 kWh.

Si en lugar de trabajar, como se ha comentado en el párrafo anterior, con valores medios por provincias, se calcula el valor medio anual diario de cada una de las estaciones, (es el valor medio de los valores anuales, teniendo en cuenta que cada año tiene un valor medio, que es la media de los diarios), se observa que la estación BU101 recoge la menor irradiación media anual diaria (12.41 MJ·m⁻²), y el mayor valor lo presenta la estación de ZA050 (17.47 MJ·m⁻²). La Figura 3.3 recoge los 44 valores obtenidos, uno para cada estación.

Si se compara la media anual diaria obtenida con los siete años, con la media de cada año se observa que, las diferencias entre ellas son muy pequeñas: el valor más grande lo presenta la estación SG020 en el año 2007 (3.14 MJ·m⁻²). Curiosamente esta estación presenta las mayores diferencias en los años 2008, 2009 y 2013. En los restantes años, 2010, 2011 y 2012, la estación que muestra las mayores diferencias con la media es LE090.

Para ver el comportamiento de cada uno de los años, se ha comparado el valor medio de irradiación media diaria de cada año con el valor medio (irradiación media anual diaria) de los siete años con los que se trabaja. Esto permite analizar, por un lado las diferencias de cada uno de los años (2007 a 2013) con la media de los siete, y por otro, analizar la media anual desde el punto de vista energético. De esta forma se puede concluir si un año es bueno o malo, en términos energéticos (si tiene un valor de irradiación media anual diaria por encima o por debajo de la media de los siete años).

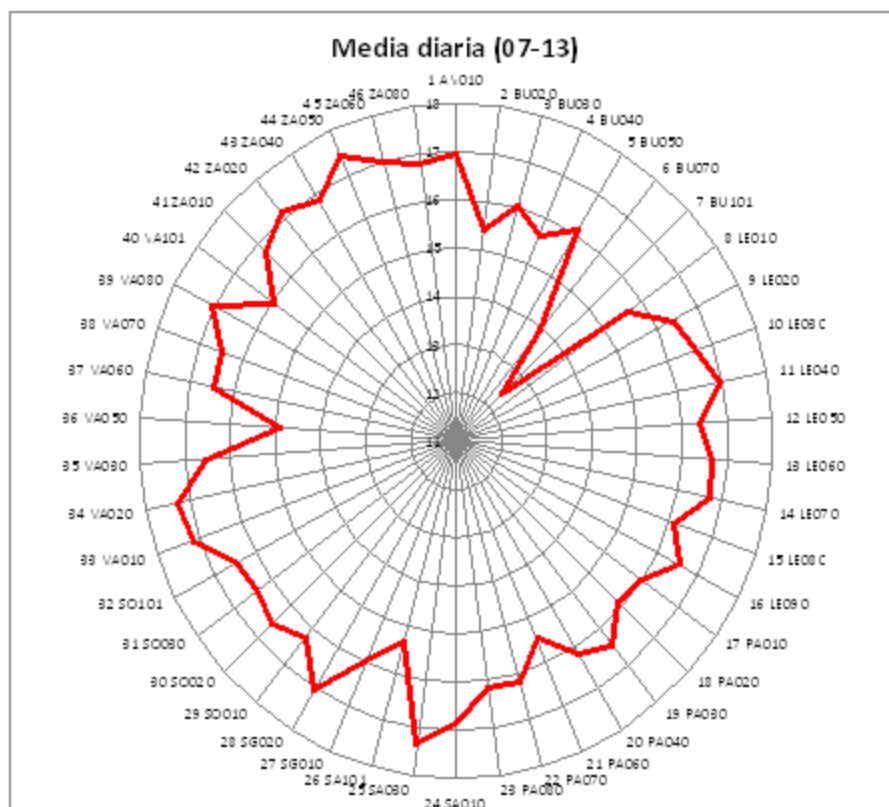


Figura 3.3. Valores medios de irradiación media anual diaria de las 46 estaciones de la red SIAR empleadas para el estudio

La Figura 3.4 representa el comportamiento de la media diaria anual con la media anual diaria de los siete años. En ella se puede ver que cada estación presenta un valor medio diario que puede ser mayor o menor que el correspondiente a la media de los siete años. Pero hay dos años que presentan un comportamiento totalmente diferente: en el año 2009 todas las estaciones tienen un valor medio anual diario mayor que el valor medio anual diario de los siete años. El año 2013 presenta el comportamiento contrario: los valores medios de ese año son todos menores que los valores medios anuales diarios (siete años).

Se ha realizado un estudio particular para esos dos años discordantes (2009 y 2013) a fin de analizar el efecto que producen en el resto de los datos cuando se trabaja con ellos para calcular un valor medio de irradiación anual diaria. Para ello se ha calculado la media de seis años, prescindiendo en primer lugar de 2009 y posteriormente de 2013. Se ha comparado con la media de los siete años. La Figura 3.5 recoge estos dos casos.



Figura 3.4. Valores medios de irradiación de cada año frente a los valores medios de irradiación de los 7 años

Como se puede ver en la Figura 3.5 a) el valor medio anual diario de todas las estaciones en el año 2009, es mayor que el valor medio anual diario de los seis años restantes. Se puede decir que este año es muy bueno desde el punto de vista de la irradiación recibida en la zona de estudio. Por el contrario el año 2013 es malo, todos los valores medios anual diario de ese año son menores que los valores medios anuales diarios de los otros 6. Esta situación puede observarse en la Figura 3.5 b).

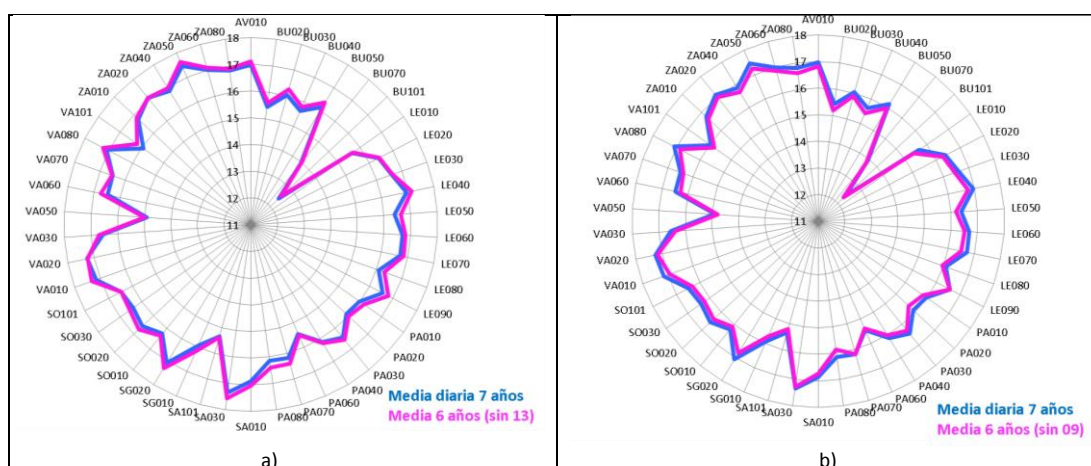


Figura 3.5. Comparación entre valores medios de irradiación anual media ($\text{MJ}\cdot\text{m}^{-2}$) de los 7 años con: a) media de 6 años (eliminando el año 2013) y b) media de 6 años (eliminando el año 2009).

Se puede afirmar que la “calidad” de los valores medios anuales diarios de irradiación global está fuertemente influenciada por los años con los que se realice la media. El número de años de los que se dispone de datos es importante para obtener una media fiable, tanto mayor cuantos más años recoja el estudio, puesto que de este modo la influencia de uno será difuminada por el resto, y la media será más aproximada a la realidad. También se observa que el “tipo” de años juega un papel importante, ya que no es lo mismo trabajar con años aproximadamente iguales, que hacerlo con años buenos (valores medios altos de irradiación), o con años malos (valores medios bajos de irradiación).

Naturalmente, si se elimina del estudio un “año malo” la media de los seis años está por encima de la media obtenida previamente y este valor quedará por debajo si se prescinde de los valores del “año bueno”. Este hecho tiene como resultado un reflejo inmediato en el cálculo de la producción. Los

valores de la serie de trabajo son fundamentales para obtener un valor medio anual diario de irradiación global fiable.

Intentando determinar los valores de irradiación más representativos de la zona, se ha analizado la distribución de la irradiación media anual diaria en intervalos. Se observa que la mayor parte de las estaciones tiene un valor medio anual diario comprendido entre 15 a 18 MJ·m⁻². Esta distribución se repite casi todos los años. Por el contrario, hay pocas estaciones con valores pequeños de irradiación y estos datos aparecen muy equitativamente distribuidos.

Todos los años, el intervalo 16-17 MJ·m⁻² es el más representativo para las estaciones (más estaciones recogen datos dentro de este intervalo). Destaca el año 2011 en el que 29 estaciones presentan este valor como característico. Es interesante recordar el carácter “anómalamente bajo” en términos energéticos, de este año. El mejor año energético, el 2009, presenta 27 estaciones con un valor medio anual diario de irradiación, entre 17 y 18 MJ·m⁻². La Figura 3.6 muestra dicha distribución.

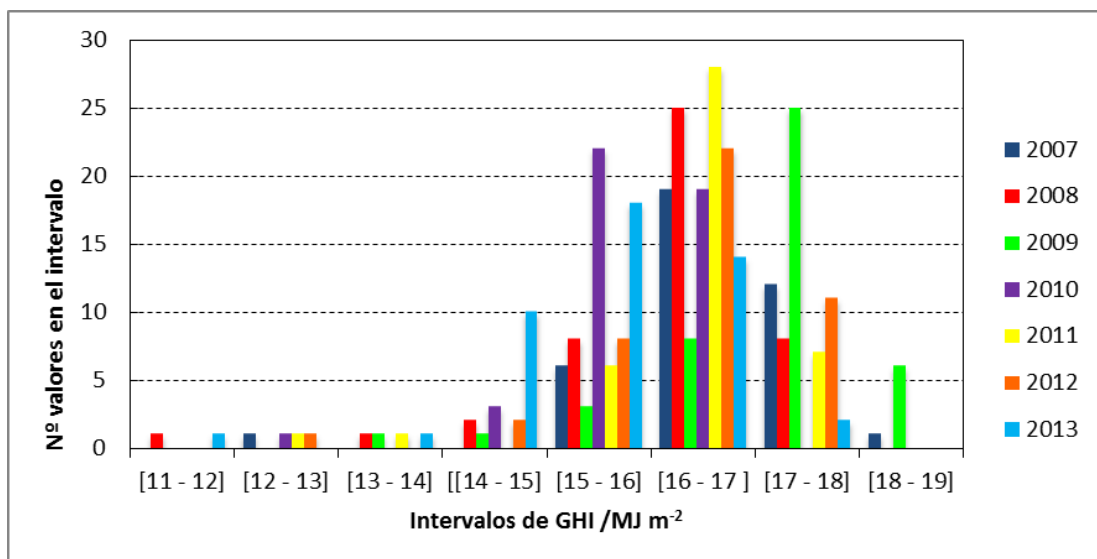


Figura 3.6. Distribución del número de valores de irradiación anual media diaria en intervalos

Para analizar la influencia de la latitud del lugar en el que está situada la estación sobre el valor de irradiación media anual registrado, se han comparado los valores medidos en las estaciones de latitud mayor y menor,

que se corresponden con las estaciones SA010 (40.590°) en Ciudad Rodrigo, Salamanca, y BU101 (42.967°), en Valle de Losa, Burgos. Para estas dos estaciones el valor medio anual, 16.83 y 12.41 MJ·m⁻², respectivamente, sí que presenta una gran diferencia (26.3 %). Si se analizan de forma independiente los diferentes años, se observa que esta diferencia es aproximadamente constante a lo largo del periodo temporal analizado, como se pone de manifiesto en Figura 3.7.

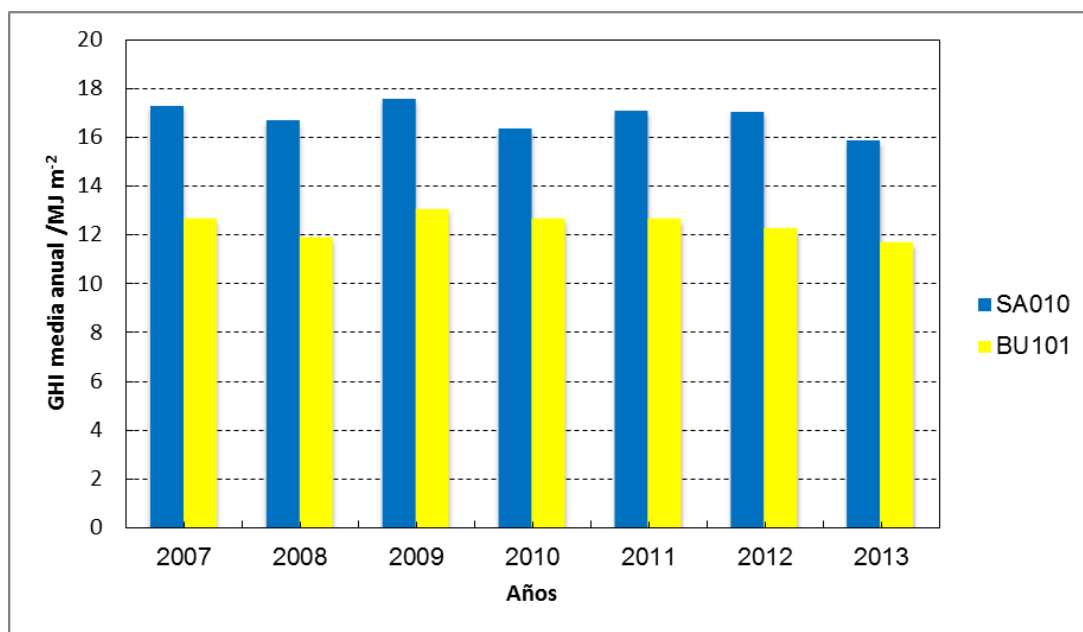


Figura 3.7. Irradiación media anual diaria (GHI / MJ·m⁻²) para las estaciones con latitud extrema en la zona de estudio (Burgos, Valle de Losa 42.967 y Salamanca, Ciudad Rodrigo 40.590)

A partir de la observación anterior, se podría concluir una relación inversamente proporcional entre la latitud y la irradiación media anual recibida, pero esta afirmación decae cuando se analizan todas las estaciones. En este caso se atribuye una componente aleatoria a la relación entre ambas variables puesto que es habitual encontrar valores de irradiación elevados para latitudes pequeñas y viceversa. Por ejemplo: La estación LE010, de latitud 42.565° tiene una irradiación media anual diaria de 15.63 MJ·m⁻² y la estación AV010 con una latitud de 40.969° muestra un valor de 16.98 MJ·m⁻². También se ha observado que las estaciones SO020 y LE090 tienen el mismo valor medio de irradiación anual diaria, 16.56 MJ·m⁻², con latitudes de 41.566° y 42.441° respectivamente. Por lo tanto y como se

mostrará en el apartado 3.3 la latitud no es la única variable determinante en la irradiación.

La relación con la longitud geográfica de la estación de medida ha arrojado un resultado semejante: las estaciones con los valores extremos de irradiación media anual diaria, ZA050 y BU101 tienen longitudes geográficas de -5.90° y -3.24° e igual sucede con las estaciones que presentan el mismo valor de irradiación media anual diaria, SO020 con longitud -3.22° y LE090 y -5.26° .

La tercera variable geográfica, la altura sobre el nivel del mar, tampoco ha demostrado ser determinante para los valores de irradiación media anual diaria recibida. Estos resultados inducen a considerar interesante un análisis estadístico con regresión simple y múltiple, que permita estudiar la relación conjunta de los tres parámetros geográficos latitud, longitud y altitud con la irradiación (ver capítulo 4).

Si se analiza la distribución de los valores del índice de claridad, K_t , se observa una cierta analogía con los de irradiación, de forma que Zamora, tiene el valor más alto y Burgos el más bajo. La Tabla 3.2 recoge los valores medios anuales diarios del índice de claridad por provincias.

Tabla 3.2. Valores medios del índice de claridad, K_t , diario para cada provincia

| | AV | BU | LE | PA | SA | SG | SO | VA | ZA |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| K_t | 0.5792 | 0.5103 | 0.5694 | 0.5722 | 0.5780 | 0.5591 | 0.5644 | 0.5612 | 0.5806 |

La Figura 3.8 representa la distribución de los valores medios del índice de claridad por provincias. En esta figura, se observa que Zamora, tiene el índice de claridad mayor (0.5806), lo que repercute en una mayor probabilidad de encontrar un día claro en esta provincia que en cualquier otra y Burgos presenta el índice de claridad más bajo (0.5103), con lo que será la provincia con menos días claros de la zona en estudio.

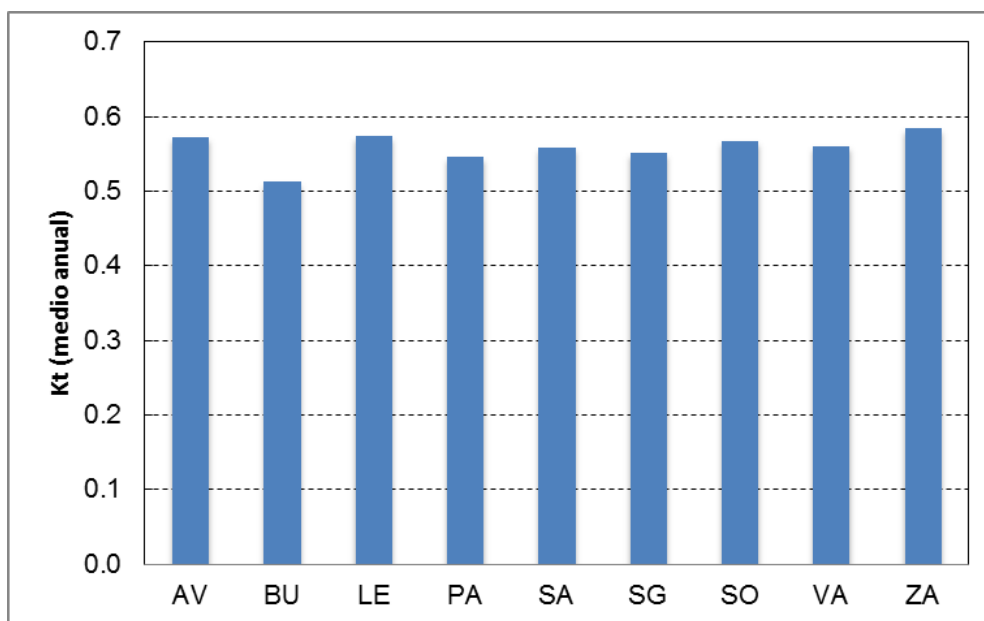


Figura 3.8. Valores medios del índice de claridad, K_t , por provincias

La diferencia entre estos valores no es grande y se ha observado que en todas las provincias el valor medio está en el intervalo central, que corresponde con un día parcialmente cubierto. Es conveniente resaltar que cada provincia tiene un número de estaciones diferente y que en algunos casos, como Ávila, sólo hay una estación por lo que los valores no son muy representativos.

Para completar el estudio se ha calculado el promedio de días que hay en cada uno de los intervalos representativos de cielo cubierto, intermedio o despejado, tal y como se indica en *lqbal*, (1983), para cada provincia buscando si existe alguna tendencia a un tipo determinado de cielo. El índice de claridad K_t de un día cubierto varía entre 0 y 0.3, de un día parcialmente claro entre 0.3 y 0.65 y de un día claro entre 0.65 y 1. La Figura 3.9 pone de manifiesto que, todas las provincias tienen un porcentaje semejante en cada intervalo. Burgos presenta el mayor número de días cubiertos (18.3 %) y Zamora el mayor número días claros (casi del 50 %) Palencia, Salamanca, Segovia, Soria y Valladolid muestran un número muy similar de días parcialmente cubiertos y días claros. Por último, la provincia con menos días cubiertos es Ávila, inferior al 10 %, con lo que en esta provincia la probabilidad de encontrar días claros es superior a la de días parcialmente cubiertos.

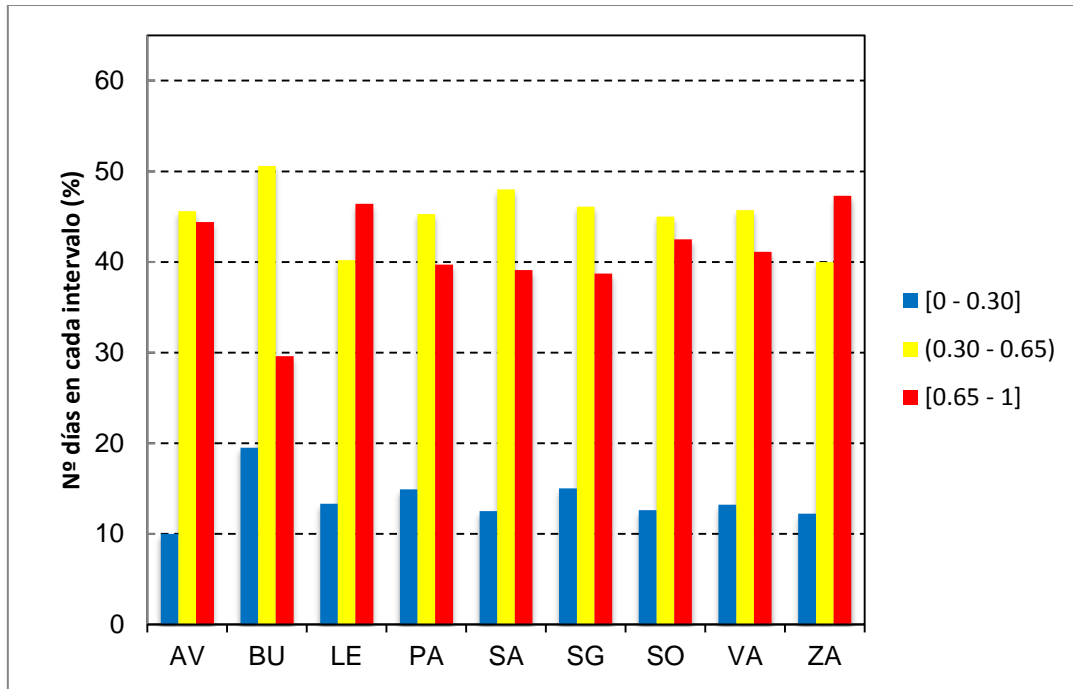


Figura 3.9. Distribución del número de días en (%) en cada intervalo del índice de claridad por provincias

En general, no hay diferencias demasiado acusadas entre las provincias todas ellas se caracterizan por cielos parcialmente cubiertos respecto al índice de claridad.

Dado que el número de estaciones por provincia es diferente, se ha decidido hacer un estudio individualizado de cada una de las provincias.

La provincia de Ávila, en la que sólo hay una estación, Nava de Arévalo, próxima al sur de la provincia de Valladolid da como resultado para todos los años estudiados un número bajo de días cubiertos (media de 35.3 días al año, correspondiente a un 9.9 %). El resto se reparte de forma muy similar entre días parcialmente cubiertos y días claros (166 y 161, correspondiendo al 45.5 % y 44.4 %, respectivamente). Esto daría un buen comportamiento general de la provincia para la producción fotovoltaica con un 90 % de los días parcialmente cubiertos o claros.

En la provincia de Burgos se dispone de seis estaciones, tres de ellas ubicadas en la zona central de la provincia (de norte a sur). Las tres estaciones centrales están situadas en la Cuenca del Duero. Otra estación está situada al oeste, también en la Cuenca del Duero, y las otras dos muy

al este, al borde de los límites de la provincia y pertenecen a la Cuenca del Ebro. Analizando el comportamiento de todas las estaciones de esta provincia, se observa que hay dos grupos con un comportamiento muy distintos. Las estaciones BU020, BU030, BU040, y BU050 tienen una tendencia semejante, con un porcentaje pequeño de días cubiertos y un porcentaje parecido para días parcialmente cubiertos y claros. La estación BU020 presenta pequeñas diferencias en el grupo. Por otro lado, las estaciones BU070 y BU101 muestran un porcentaje muy alto de días cubiertos y de días parcialmente cubiertos, superior al de los días claros. Una explicación de este comportamiento puede ser que mientras el primer grupo pertenece a la Cuenca del Duero y el segundo a la Cuenca del Ebro. La Figura 3.10 recoge la distribución de porcentaje de los tres intervalos de clasificación en todas las estaciones.

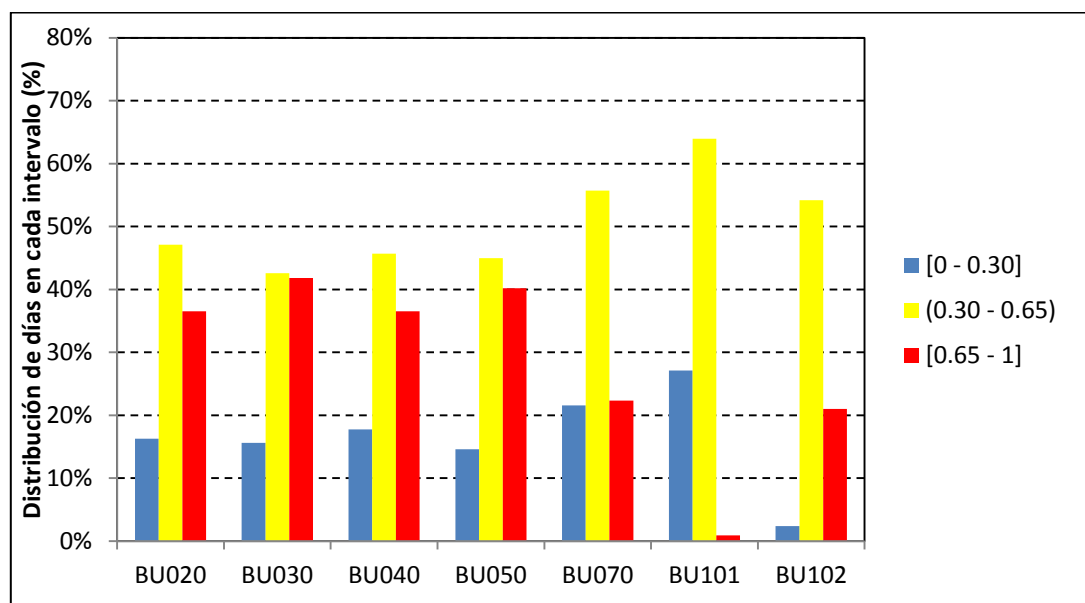


Figura 3.10. Distribución del número de días en (%) en cada intervalo del índice de claridad por estaciones para Burgos

La influencia de los parámetros geográficos sobre este comportamiento se ha analizado detalladamente sin obtener resultados concluyentes. La variación de la latitud entre ellas está entre 42.967° de BU101 y 41.637° de BU050, diferencia que consideramos no influyente. En cuanto a la longitud, ésta varía entre -3.24° y -3.58° , lo que tampoco justifica la discrepancia. Se ha observado que la altitud, puede influir de alguna manera en el

comportamiento. En este caso las BU101 y BU070 están a poca altitud, 635 y 532 m, mientras que las otras cuatro están a una altitud mayor, 870 m para BU050 y 975 m para BU020. La diferencia de altitudes puede justificar el distinto comportamiento entre los dos grupos de estaciones, junto con la influencia de las cuencas hidrológicas.

En la provincia de León se trabaja con nueve estaciones: ocho de ellas están ubicadas en el suroeste de la provincia, próximas entre sí, con longitudes comprendidas entre -5.02° y -5.90° , altitudes que van desde 750 a 885 m y todas ellas en la Cuenca del Duero. La novena estación es la LE010, está ubicada al noroeste de la provincia y tienen menor altitud, 467 m, longitud diferente, -6.72° , y no pertenece a la Cuenca del Duero. En este caso se observa también una similitud en el comportamiento de las estaciones pertenecientes a la Cuenca del Duero, presentando todas unos porcentajes similares para los tres intervalos del índice de claridad. Por el contrario la estación LE010 tiene un comportamiento diferente. En la Tabla 3.3, se han reflejado los porcentajes medios de cada una de las estaciones para los tres intervalos.

Tabla 3.3. Distribución de los porcentajes en los tres intervalos del índice de claridad para cada una de las nueve estaciones de León.

| | LE010 | LE020 | LE030 | LE040 | LE050 | LE060 | LE070 | LE080 | LE090 |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| [0 – 0.30] | 18.06 | 13.54 | 12.79 | 11.96 | 12.77 | 12.22 | 12.16 | 14.92 | 13.47 |
| [0.30 – 0.65) | 41.76 | 37.64 | 39.18 | 36.19 | 40.77 | 37.05 | 37.58 | 41.34 | 36.67 |
| [0.65 - 1) | 40.18 | 48.82 | 48.03 | 51.85 | 46.46 | 50.73 | 50.26 | 43.74 | 49.86 |

Se observa que LE010 presenta el porcentaje mayor de días cubiertos (18 %) que el resto de estaciones. La otra diferencia manifestada es que hay más probabilidad de encontrar un día parcialmente cubierto que un día claro, cosa que no ocurre en ninguna de las otras estaciones.

En Palencia hay siete estaciones incluidas en el estudio, tres de ellas ubicadas al sur de la provincia, dos al oeste y otras dos en el centro, todas pertenecen a la Cuenca del Duero. Las longitudes no son muy diferentes entre sí, oscilando entre -4.25° y -4.78° . En este caso las altitudes sobre las

que están ubicadas tampoco varían mucho. La más alta es PA070 con 927 m y la más baja, 744 m, PA020.

No hay ninguna estación que destaque en el comportamiento. El reparto de días es semejante en las siete estaciones con un promedio es de 14.6 % para días cubiertos, 43.8 % para días parcialmente cubiertos y 41.7 % para días claros. Cabe destacar que las que más al norte están, como las estaciones PA070 y PA060, mayor porcentaje de días cubiertos presentan, pero con poca diferencia del resto.

En Salamanca se han incluido tres estaciones en el estudio, dos de ellas muy próximas geográficamente y al noreste de la provincia y la otra al sureste, muy alejada y casi en el límite. Todas están en la Cuenca del Duero y respecto a las coordenadas geográficas, presentan variación tanto en la longitud como en la altitud.

Dos estaciones, SA010 y SA030, tienen un comportamiento muy similar. En ambas es más probable encontrar días claros que parcialmente cubiertos, 49.2 % respecto a 51.4 %. Esto se puede traducir en un buen comportamiento de las instalaciones fotovoltaicas en estas zonas ya que la probabilidad de encontrar días claros es alta. La tercera estación, SA101, destaca por una baja probabilidad de encontrar un día claro, 26.32 %, y mayor que el resto en la provincia para días cubierto (un 13.1 %).

En Segovia se localizan dos estaciones, próximas entre sí y con valores muy similares de longitud, y de altitud. Ambas estaciones pertenecen a la Cuenca del Duero. El estudio de los índices de claridad no presenta diferencias destacables entre las dos estaciones. La estación SG010 ubicada al oeste tienen un mayor porcentaje de días cubiertos (15 %) y un 40.7 % de días claros, mientras que la estación SG030 muestra 11.4 % de días cubiertos y un 46.9 % de días claros.

En Soria hay cuatro estaciones disponibles, repartidas por la provincia de forma que cubren todas las direcciones, pero todas ellas están en la Cuenca del Duero. Las diferencias entre las longitudes son mínimas y la altitud tampoco presenta grandes diferencias. El comportamiento de las cuatro

estaciones es muy similar entre ellas en el reparto de días claros, cubiertos o parcialmente cubiertos, de forma que podemos encontrarnos aproximadamente un 12.5 % de días cubiertos, un 43.3 % de días parcialmente cubiertos y un 44.1 % de días claros.

En la provincia de Valladolid se ha trabajado con ocho estaciones de medida, distribuidas aleatoriamente por toda la provincia y todas en la Cuenca del Duero con variables geográficas de longitud y altitud parecidas. La estación VA050 presenta un comportamiento totalmente diferente al resto con el porcentaje más elevado de días cubiertos (15.1 %), también el mayor porcentaje de días parcialmente cubiertos (61.7 %), mientras que la posibilidad de encontrar días claros es sólo del 23.2 %. También VA101 tiene un comportamiento parecido aunque la diferencia entre los días parcialmente cubiertos y claros es menor (45.4 y 39.18 % respectivamente), y la probabilidad de encontrar un día cubierto es 15.42 %. El resto de estaciones se puede dividir en dos grupos: aquellas que tiene un porcentaje de días parcialmente cubiertos mayor que el de los días claros, VA010, VA020 y VA080, y aquellas cuya la probabilidad de encontrar un día parcialmente cubierto y claro es prácticamente la misma, VA030, VA060 y VA070.

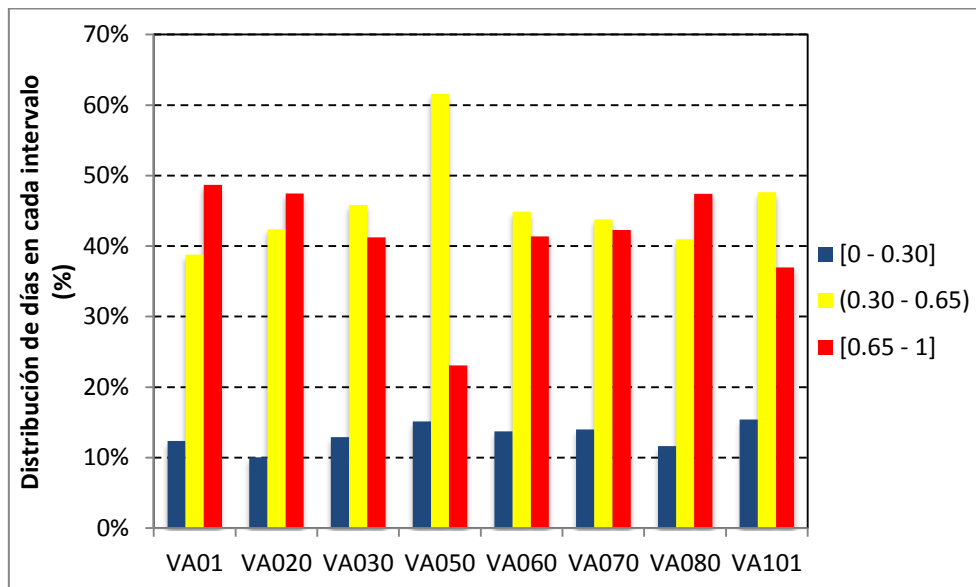


Figura 3.11. Distribución del número de días en (%) en cada intervalo del índice de claridad por estaciones para la provincia de Valladolid

La Figura 3.11 muestra la distribución del número de días (%) de los tres intervalos de índice de claridad, para todas las estaciones.

Si se analizan los parámetros geográficos fundamentales de las estaciones, se observa que las dos primeras (VA050 y VA101), tienen longitudes menores que el resto y la altitud de VA050 es la más alta de la provincia 816 m. Aquellas que presentan un comportamiento similar, con mayor porcentaje de días claros tienen una longitud parecida, alrededor de -5° y una altitud también semejante, unos 770 m. Aquellas que tienen muy similar los porcentajes de días parcialmente cubiertos y claros presentan una longitud entre -4.3° y -4.7° y la altitud está entre 750 y 756 m.

En Zamora se cuenta con seis estaciones, todas ellas en la Cuenca del Duero. Presentan gran similitud en los valores de los parámetros geográficos, longitud y altitud, entre -5.37° y -5.90° y 650 m de ZA080 y 738m de ZA060, lo que contribuye a que las diferencias en el valor de la irradiación sean mínimas. La Tabla 3.4 presenta la distribución de los porcentajes en los tres intervalos del índice de claridad para todas las estaciones de esta provincia.

Se puede observar que el comportamiento de todas las estaciones es parecido. La mayor probabilidad mayor aparece en el intervalo correspondiente a días claros. Esto confirma los resultados obtenidos en el estudio de la irradiación (Zamora era la provincia que mayor radiación global).

Tabla 3.4. Distribución de los porcentajes en los tres intervalos del índice de claridad para cada estación de Zamora.

| | ZA010 | ZA020 | ZA040 | ZA050 | ZA060 | ZA080 |
|---------------|-------|-------|-------|-------|-------|-------|
| [0 – 0.30] | 11.96 | 11.73 | 12.50 | 11.09 | 12.68 | 12.17 |
| [0.30 – 0.65) | 38.86 | 39.31 | 39.02 | 35.12 | 37.57 | 41.86 |
| [0.65 - 1) | 49.19 | 48.97 | 48.48 | 53.79 | 49.74 | 45.97 |

Por otra parte el porcentaje de los días cubiertos es relativamente pequeño en todas las estaciones y muy similar, lo que permite decir que en Zamora

en general es donde más días claros nos vamos a encontrar y por lo tanto donde mayor rendimiento pueden tener las instalaciones fotovoltaicas.

3.2.2 DATOS DE LA AEMET

Con los datos obtenidos de las estaciones de la red SIAR se quiere obtener mapas de irradiación media anual diaria (global horizontal), para Castilla y León, para lo que se va a trabajar con métodos matemáticos de interpolación. Los métodos permiten obtener valores de irradiación en cualquier punto de la zona de estudio, a partir de un conjunto de valores experimentales de la misma variable repartidos en la zona. Lo importante es comprobar que los valores así generados (por los métodos de interpolación) son fiables. Para ello se van a establecer unos procedimientos, que hemos denominado “controles”, que permiten comparar los valores generados por interpolación con los experimentales, recogidos en alguna de las estaciones. Uno de estos controles se va a realizar con el conjunto de estaciones de la AEMET. Para ello se han solicitado a la Agencia Estatal los datos de sus estaciones en esta zona. La Agencia sólo dispone de cuatro estaciones en la región, ubicadas en Soria, Salamanca, Valladolid y León.

Los datos han pasado los controles necesarios para cumplir las condiciones establecidas por la OMM, por lo que únicamente se trabaja con ellos para unificar el formato con los datos de la red SIAR. La labor consiste en transformar los datos horarios, que es el formato en el que los suministra la AEMET, y pasarlos a datos diarios.

3.3 MÉTODOS DE INTERPOLACIÓN CON ARCGIS

El objetivo es representar mapas de irradiación global horizontal (valores medios diarios), de la región de Castilla y León, de forma que se pueda conocer el valor de la variable, en cualquier punto de la zona. Se ha elegido el software ArcGis, que dispone de diferentes aplicaciones tanto, para el

diseño como para el análisis y tratamiento de sistemas de la información geográfica.

Una de las aplicaciones del ArcGis permite obtener el valor de una determinada propiedad, en cualquier punto de una zona en estudio, a partir de un conjunto de valores, aplicando métodos matemáticos de interpolación. Estos valores se pueden representar en mapas logrando una visión continua de la variación de la característica representada. En este caso, la propiedad con la que se trabaja es la irradiación global horizontal, que está caracterizada por las coordenadas geográficas (latitud, longitud y altitud). La idea es introducir los valores de irradiación de las 46 estaciones, aplicar los métodos de interpolación y obtener los mapas de la distribución de irradiación en todos los puntos de la región de Castilla y León.

3.3.1 ANÁLISIS PREVIOS

Como lo que se quiere es visualizar las estaciones en el mapa de Castilla y León se emplea el ArcMap, que es la aplicación de ArcGis que permite crear tanto mapas sencillos como cartografías complejas, pudiéndose representar la información geográfica como un conjunto de capas y elementos de un mapa. Para poder visualizar los datos en un mapa es necesario utilizar un sistema de coordenadas, es decir, hay que determinar la proyección cartográfica del mapa en el conjunto de datos. Para ello es necesario tener en cuenta que las unidades de mapa son las que se visualizan en las capas (en el marco de datos) y por lo tanto, las que se utilizan. Estas quedan determinadas por el sistema de coordenadas. En nuestro caso se quiere trabajar con el mapa de las nueve provincias de Castilla y León en el sistema de referencia “*Geographics Coodinate Systems, World, WGS1984*”, que permite obtener los planos con el elipsoide terrestre.

Para poder introducir los datos (en este caso los valores de irradiación de las 46 estaciones), se crea una base de datos a partir de la tabla obtenida del tratamiento de los mismos. En ella figuran el nombre de la estación, la identificación de la estación, latitud, longitud, altitud y ocho columnas más de

las que las siete primeras representan el valor medio diario anual de cada uno de los siete años de estudio y la última el valor medio de las siete anteriores.

Es importante reseñar que en este punto sólo se va a trabajar con el valor de la última columna, valor medio anual diario de irradiación (como se ha indicado anteriormente en el punto 3.2.1), es decir trabajamos con 46 valores de irradiación, uno por estación.

Se añade a la Tabla de Contenidos, indicando las variables con las que se quiere trabajar, latitud y longitud y como variable auxiliar la altitud. La Tabla 3.5 recoge los valores de las 46 estaciones de la red SIAR con las que se trabaja.

Los métodos de interpolación se encuentran en la caja de herramientas (de análisis espacial). Se selecciona el método de interpolación que se quiere aplicar y se elige la columna de valores con los que se va a realizar la interpolación. En cada método de interpolación se requiere unos determinados parámetros, número de puntos con el que se va a trabajar en cada celda, peso, tipo de variograma, o los parámetros que corresponda en cada caso. Seleccionando el modo de interpolación se van eligiendo los parámetros con los que se quiere trabajar. Con esto se obtiene el mapa de interpolación correspondiente, del que se puede extraer el valor de la irradiación en un punto cualquiera de la zona de trabajo.

Tabla 3.5. Tabla con datos de las estaciones

| | Estación | latitud | longitud | altitud | GHI /MJ·m ⁻² |
|----|----------|---------|----------|---------|-------------------------|
| 1 | AV010 | 40.969 | -4.77 | 921 | 16.9801 |
| 2 | BU020 | 42.746 | -4.13 | 975 | 15.4433 |
| 3 | BU030 | 42.040 | -3.77 | 840 | 16.0322 |
| 4 | BU040 | 42.346 | -3.80 | 770 | 15.6402 |
| 5 | BU050 | 41.637 | -3.58 | 870 | 16.1454 |
| 6 | BU070 | 42.702 | -3.08 | 532 | 13.9664 |
| 7 | BU101 | 42.967 | -3.24 | 635 | 12.4121 |
| 8 | LE010 | 42.565 | -6.72 | 467 | 15.6328 |
| 9 | LE020 | 42.506 | -5.43 | 791 | 16.3909 |
| 10 | LE030 | 42.396 | -5.51 | 777 | 16.6033 |
| 11 | LE040 | 42.258 | -5.74 | 779 | 16.9565 |
| 12 | LE050 | 42.218 | -5.84 | 750 | 16.4009 |
| 13 | LE060 | 42.460 | -5.90 | 835 | 16.6910 |
| 14 | LE070 | 42.464 | -5.77 | 874 | 16.7260 |
| 15 | LE080 | 42.374 | -5.02 | 856 | 16.0886 |
| 16 | LE090 | 42.441 | -5.26 | 885 | 16.5645 |
| 17 | PA010 | 42.051 | -4.30 | 868 | 15.9863 |
| 18 | PA020 | 41.946 | -4.49 | 750 | 15.8980 |
| 19 | PA030 | 42.081 | -4.72 | 744 | 16.4519 |
| 20 | PA040 | 42.271 | -4.59 | 817 | 16.1456 |
| 21 | PA060 | 42.489 | -4.25 | 821 | 15.4552 |
| 22 | PA070 | 42.528 | -4.78 | 927 | 16.1854 |
| 23 | PA080 | 42.344 | -4.28 | 798 | 16.1486 |
| 24 | SA010 | 40.590 | -6.54 | 635 | 16.8607 |
| 25 | SA030 | 40.990 | -5.48 | 815 | 17.3493 |
| 26 | SA101 | 41.036 | -5.36 | 847 | 15.3326 |
| 27 | SG010 | 41.301 | -4.30 | 870 | 15.9209 |
| 28 | SG020 | 41.172 | -4.48 | 822 | 17.0629 |
| 29 | SO010 | 41.461 | -2.50 | 943 | 16.2537 |
| 30 | SO020 | 41.566 | -3.22 | 855 | 16.5615 |
| 31 | SO030 | 41.834 | -2.43 | 1063 | 16.3908 |
| 32 | SO101 | 41.734 | -2.10 | 1043 | 16.4801 |
| 33 | VA010 | 42.151 | -5.29 | 748 | 17.1594 |
| 34 | VA020 | 41.228 | -5.21 | 793 | 17.2612 |
| 35 | VA030 | 41.310 | -4.69 | 750 | 16.5539 |
| 36 | VA050 | 41.767 | -4.10 | 816 | 14.9012 |
| 37 | VA060 | 41,487 | -5.0 | 658 | 16,4831 |
| 38 | VA070 | 41.643 | -4.27 | 756 | 16.4894 |
| 39 | VA080 | 41.861 | -5.07 | 739 | 17.0883 |
| 40 | VA101 | 41.711 | -4.70 | 714 | 15.9403 |
| 41 | ZA010 | 42.000 | -5.81 | 709 | 16.7621 |
| 42 | ZA020 | 41.482 | -5.65 | 659 | 17.1188 |
| 43 | ZA040 | 41.880 | -5.39 | 701 | 16.8603 |
| 44 | ZA050 | 41.782 | -5.90 | 714 | 17.4741 |
| 45 | ZA060 | 41.934 | -5.67 | 738 | 17.0064 |
| 46 | ZA080 | 41.508 | -5.37 | 650 | 16.8128 |

Observando los mapas obtenidos mediante diferentes métodos de interpolación con las 46 estaciones de la red SIAR, se ha llegado a tres conclusiones importantes:

- hay dos estaciones que proporcionaban datos anómalos de irradiación, por lo que se ha determinado eliminarlas del conjunto de estaciones base (estaciones de la red SIAR).
- de todos los métodos de interpolación disponibles en el software, hay cuatro que presentaban valores interesantes para analizar, el resto no son capaces de ajustar los resultados a los valores de la zona, la tendencia no es ajustada a las características de la zona o de los datos.
- se observa que, al no disponer de datos en las zonas límites, los resultados en los puntos próximos a la frontera de la región no son coherentes con los esperados y obtenidos en otros estudios (mapa de radiación de la AEMET para Castilla y León).

Como resultado de las conclusiones anteriores se ha decidido:

- trabajar con 44 estaciones (se han eliminado una de la provincia de Salamanca y otra de la provincia de Valladolid)
- emplear cuatro métodos de interpolación, IDW, Vecino Natural, Kriging y Splines
- buscar valores de irradiación, en zonas próximas a todo el perímetro de Castilla y León (lo más importante es encontrar valores de irradiación anual (medias diarias), en el mismo intervalo temporal de las provincias limítrofes).

Las estaciones elegidas con las que se ha trabajado (para añadir al conjunto de la red SIAR), que permitirán suavizar todas las superficies de interpolación son, del norte de Portugal, sureste de Galicia, Asturias, Cantabria, País Vasco, La Rioja, Aragón, norte de Castilla la Mancha, norte de Madrid y norte de Extremadura. La Figura 3.12 muestra la distribución de todas las estaciones de la red de trabajo.

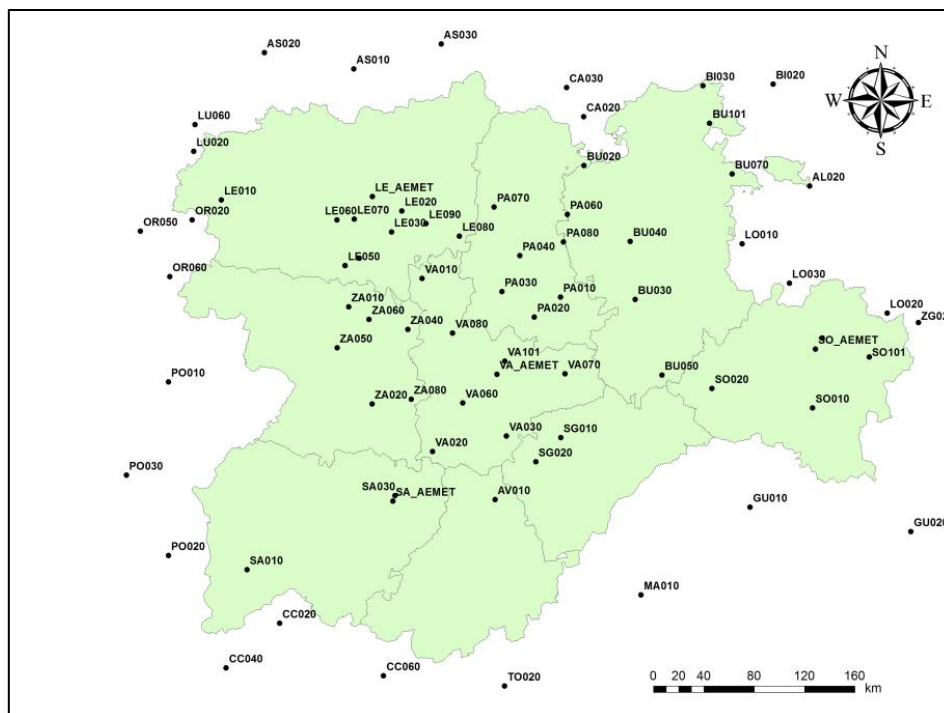


Figura 3.12. Estaciones empleadas para hacer el estudio. Interiores (SIAR), exteriores (varias fuentes) círculos las estaciones de la AEMET

Localizadas las estaciones en las provincias próximas y recogidos los datos necesarios, se vuelven a crear las nuevas bases de datos con las estaciones complementarias. El número de estaciones de la red ahora, es de 71. De ellas 44 son de la red SIAR, tres de Portugal, cinco de Galicia, tres de Asturias, dos de Cantabria, tres del País Vasco, tres de la Rioja, una de Zaragoza, una de Madrid, tres de Castilla La Mancha y tres de Extremadura.

Cuando se ejecuta uno de los métodos de interpolación, se obtiene el correspondiente mapa de irradiación. Para comprobar la validez del método, se extraen los valores numéricos de irradiación de diferentes puntos de la zona y se comparan con los medidos en esos puntos. Para ello el procedimiento de control consiste en eliminar de la base general de datos, cinco estaciones (estaciones de validación). Se obtienen los valores de irradiación para cada modelo de interpolación con las restantes estaciones y se comparan los valores obtenidos con los medidos en las estaciones de validación.

Los cuatro métodos de interpolación seleccionados se aplican a las 39 estaciones de la red SIAR (cumpliendo la teoría que se verá en el punto 4.3),

más 27 complementarias (66 valores en total) como base de interpolación, y cinco estaciones de validación. Mediante un proceso de validación cruzada en el que se van modificando en todos los métodos de interpolación las estaciones seleccionadas para la interpolación y para el control, se calcula el error cometido con cada modelo de interpolación para cada una de las estaciones de validación.

Para realizar la validación de los métodos y de los modelos de irradiación propuestos se han utilizado los parámetros estadísticos habituales en estudios de energía solar: RMSE (*Root Mean Squared Error*), MBE (*Mean Bias Error*), MAE (*Mean Absolute Error*), MAPE (*Mean Absolute Percentage Error*).

Estos parámetros se han definido como sigue:

$$RMSE = \sqrt{\frac{\sum(v_r - v_i)^2}{N}} \quad RMSE \% = \frac{RMSE \times 100}{Promedio(v_r)}$$

$$MBE = \frac{1}{N} \sum (v_r - v_i) \quad MBE \% = \frac{MBE \times 100}{Promedio(v_r)}$$

$$MAE = \frac{1}{N} \sum |v_r - v_i| \quad MAE \% = \frac{MAE \times 100}{Promedio(v_r)}$$

$$MAPE = \frac{1}{N} \sum \left| \frac{v_r - v_i}{v_r} \right| \times 100$$

donde: v_r es el valor real u observado;

v_i es el valor interpolado o calculado

El RMSE da la medida de las diferencias, en promedio entre los valores reales y los interpolados o calculados. Dado que los errores se elevan al cuadrado antes de promediarlos, el RMSE da un peso relativamente alto a errores altos.

MBE proporciona información sobre la tendencia o sesgo del modelo a sobreestimar o subestimar una variable.

MAE es el promedio de los errores, lo que significa que todas las diferencias individuales se equilibran en el promedio.

MAPE da en porcentaje la magnitud del error en comparación con la magnitud de los datos

En cada control se realizan 29 interpolaciones diferentes, se calculan los valores en los cinco puntos geográficos de las estaciones de validación y se comparan con los valores experimentales (medidos en las propias estaciones). En cada caso se calculan los cuatro estadísticos: MBE, RMSE, MAE y MAPE. El valor de estos parámetros permiten establecer el método de interpolación que mejor se ajusta a la zona de estudio. En total se han realizado 203 mapas.

En el Control 1 se han seleccionado dos estaciones de León, una de Valladolid, una de Palencia y una del País Vasco como grupo de control. Esta última se ha elegido para ver la influencia de la zona límite, ya que la estación pertenece al Servicio Vasco de Meteorología, pero está en Castilla y León. En la Figura 3.13 se recogen las estaciones del control y las estaciones base para realizar la interpolación.

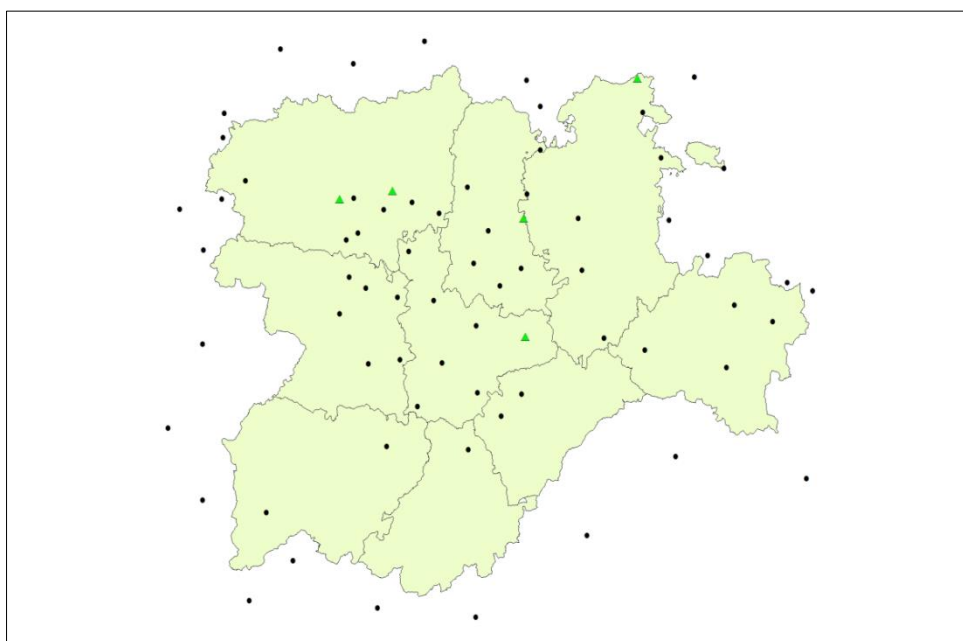


Figura 3.13. Estaciones base para interpolar y las cinco del Control1: León, León, Palencia, Valladolid y Vizcaya

La figura 3.14 representa los planos de interpolación obtenidos para un mismo método. Únicamente varía el número de puntos seleccionados para cada celda del proceso de interpolación. En este caso el método seleccionado es el *SPLIN REGULARIZADO* con un peso de 0.01 para 12 y 35 puntos.

Para el Control 2 se sigue el mismo proceso, pero esta vez las estaciones de validación se han seleccionado de forma que estuvieran distanciadas entre si y que cubrieran la mayor parte de la superficie, teniendo cuidado de no dejar una provincia sin valores en la base de interpolación, es decir que siempre hubiera por lo menos una estación en la base. Las estaciones son: LE020, PA080, SO020, VA060 y ZA060. La Figura 3.15, muestra los mapas de interpolación de Kriging Ordinario Circular y 35 puntos y el Kriging Ordinario Lineal y 35 puntos.

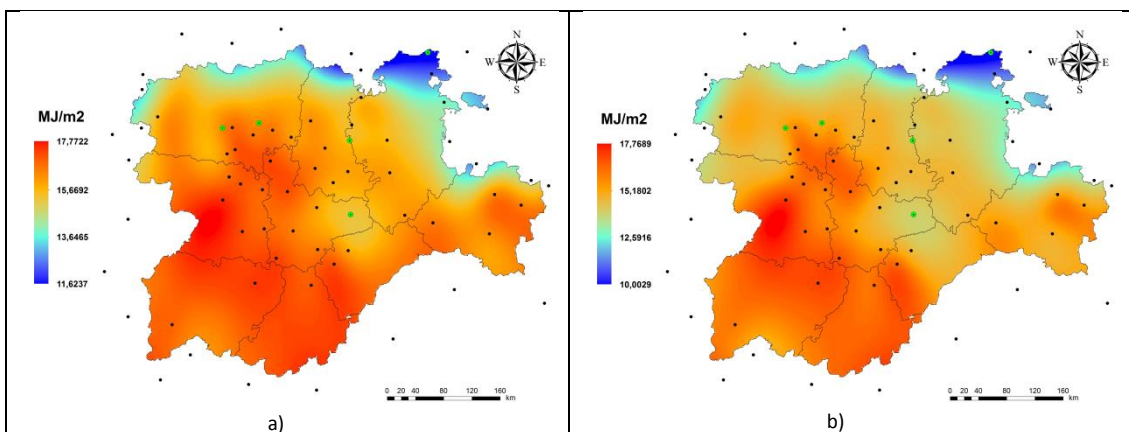


Figura 3.14. Mapas de interpolación para *SPLIN REGULARIZADO* con peso 0.01: a) 12 puntos y b) 35 puntos

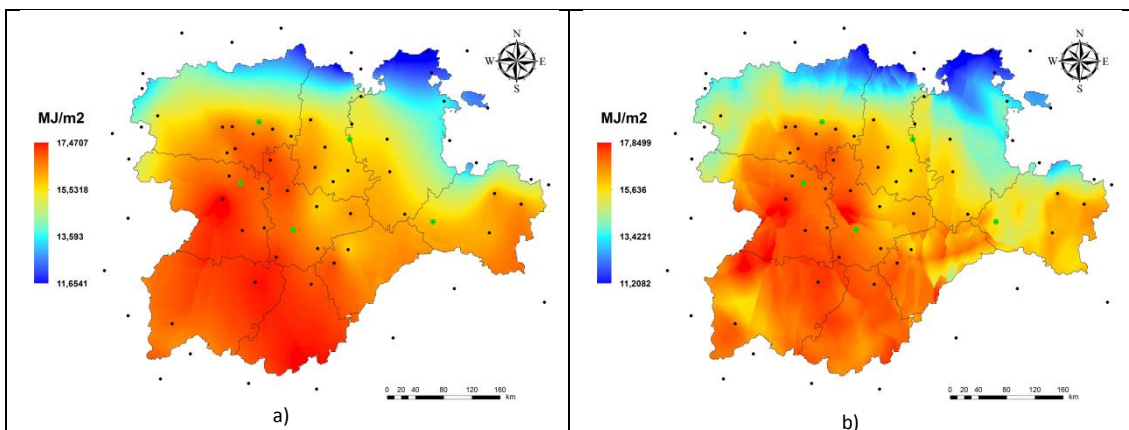


Figura 3.15. Mapas de interpolación Control2, con menor RMSE: a) Kriging Ordinario Circular 35 puntos y menor MAPE, b) Kriging Ordinario Lineal 35 puntos

Para el Control 3 se han seleccionado las estaciones LE020, BU050, SA030, VA080 y ZA080. En este caso se ha seleccionado la provincia de Salamanca, que es una zona muy poco poblada de estaciones, para ver el comportamiento de grandes zonas sin valores en la base de interpolación. El resto de estaciones ha seguido la misma idea que en el control anterior, estaciones distribuidas que cubran la mayor zona posible. La Figura 3.16 representa los mapas de interpolación del método IDW para 12 y 35 puntos.

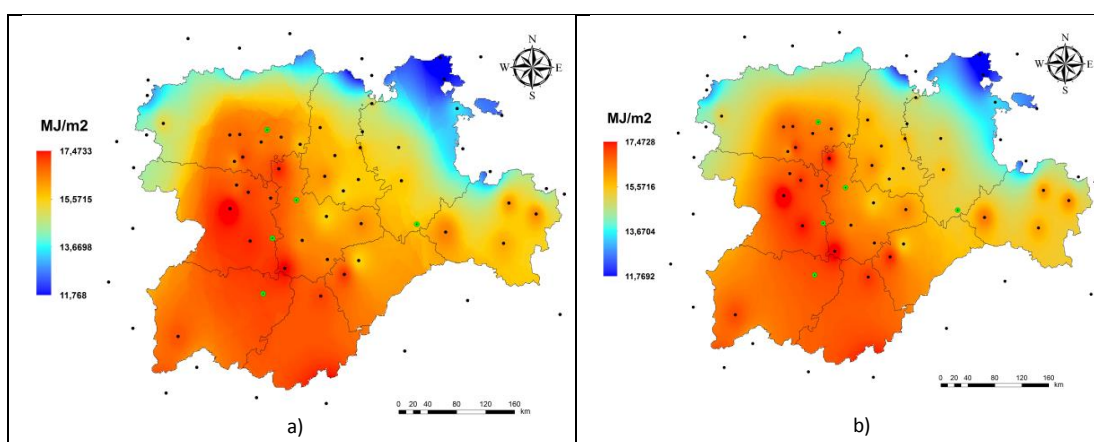


Figura 3.16. Mapas de interpolación con IDW: a) 12 puntos, b) 35 puntos

En el Control 4 se trabaja con las estaciones de validación más centradas en la zona de estudio. Las estaciones seleccionadas son: LE020, PA030, SG010, VA020 y ZA040. En la Figura 3.17 se han representado los mapas de interpolación correspondientes al Kriging Universal Lineal de 12 puntos y Kriging Ordinario Gaussiano de 35 puntos

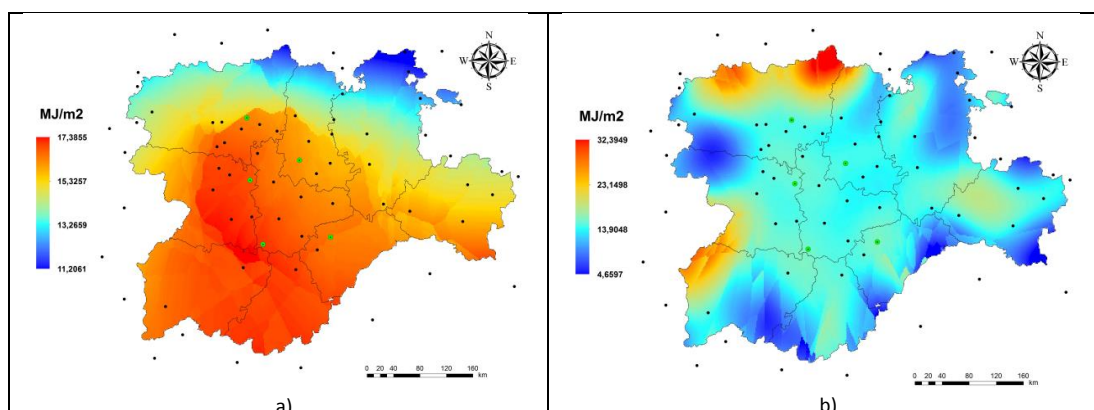


Figura 3.17. Mapas de interpolación con Kriging: a) Universal Lineal de 12 puntos menor RMSE, b) Ordinario Gaussiano de 35 puntos mayor RMSE

Como último caso de estudio con estaciones de la red SIAR como control, se ha trabajado buscando la influencia de la latitud. Para ello se han seleccionado LE020, SA030, VA020, ZA060 y ZA080, las cinco estaciones con un valor de longitud muy próximo. La Figura 3.18 representa el mapa de interpolación con el método Vecino Natural.

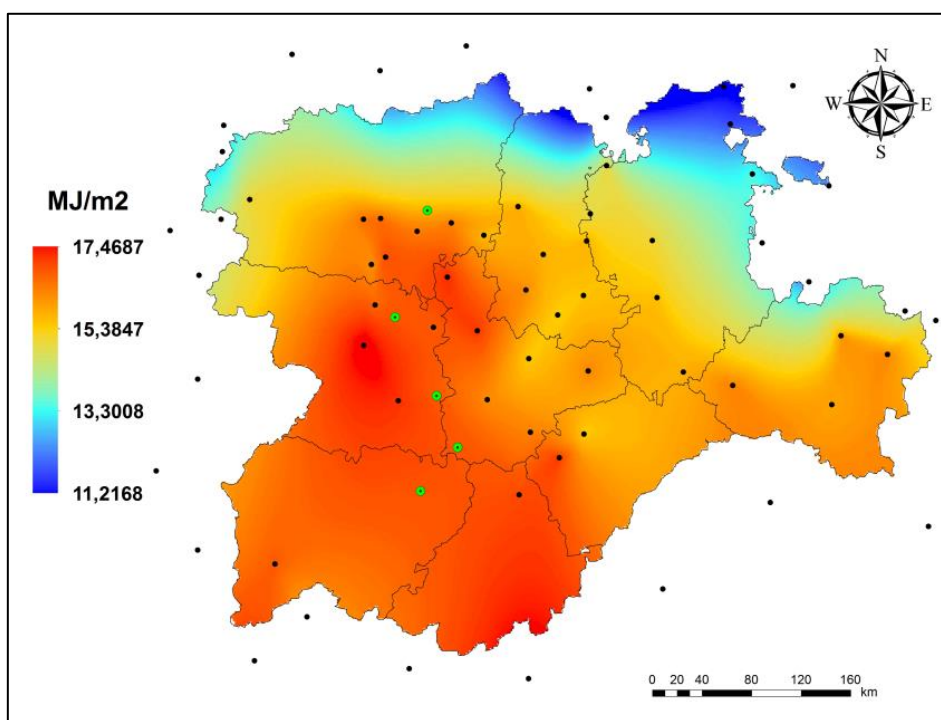


Figura 3.18. Mapa de interpolación con Vecino Natural

Para obtener los resultados finales se ha realizado un control con las estaciones de la AEMET ubicadas en Salamanca, Soria, Valladolid y León. En este caso las estaciones son totalmente diferentes a las empleadas como base de interpolación (red SIAR). Los resultados de este control son los que se van a tomar como definitivos. El proceso de trabajo es el mismo que en los controles anteriores, con la diferencia de tener 71 estaciones como base de interpolación y cuatro de control, en lugar de los anteriores en los que se ha trabajado con 66 estaciones de base y cinco de control. La Figura 3.19 recoge la situación de las estaciones base de interpolación y las cuatro de control.

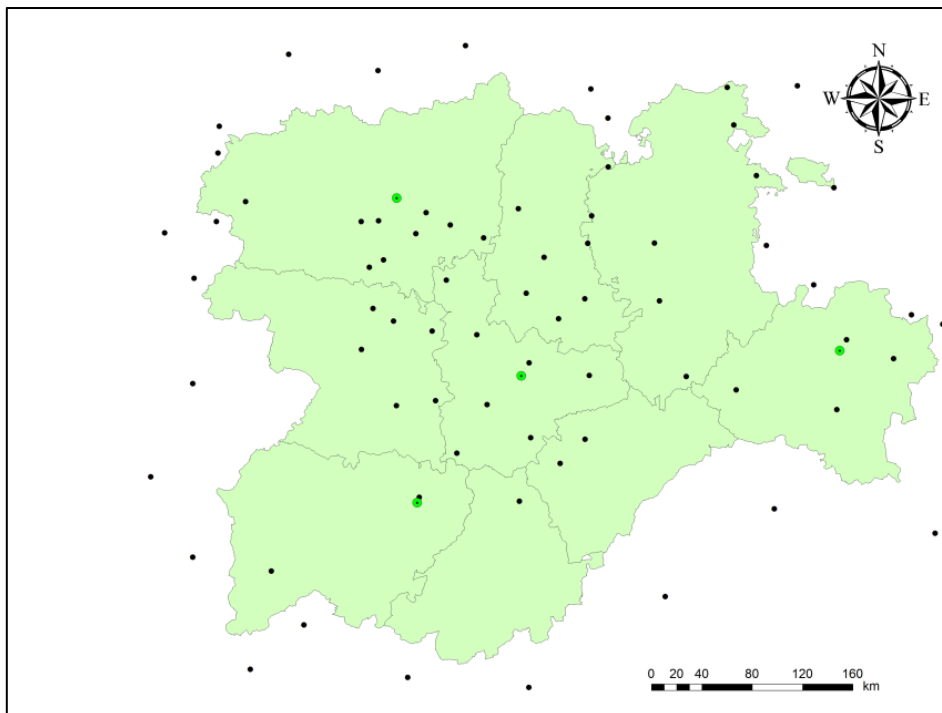


Figura 3.19. Representación de las 71 estaciones base de interpolación (SIAR + diferentes medios) y cuatro de control (AEMET)

La Figura 3.20 representa el mapa de interpolación del Kriging Universal Cuadrático con 35 puntos, para el control de las estaciones de la AEMET.

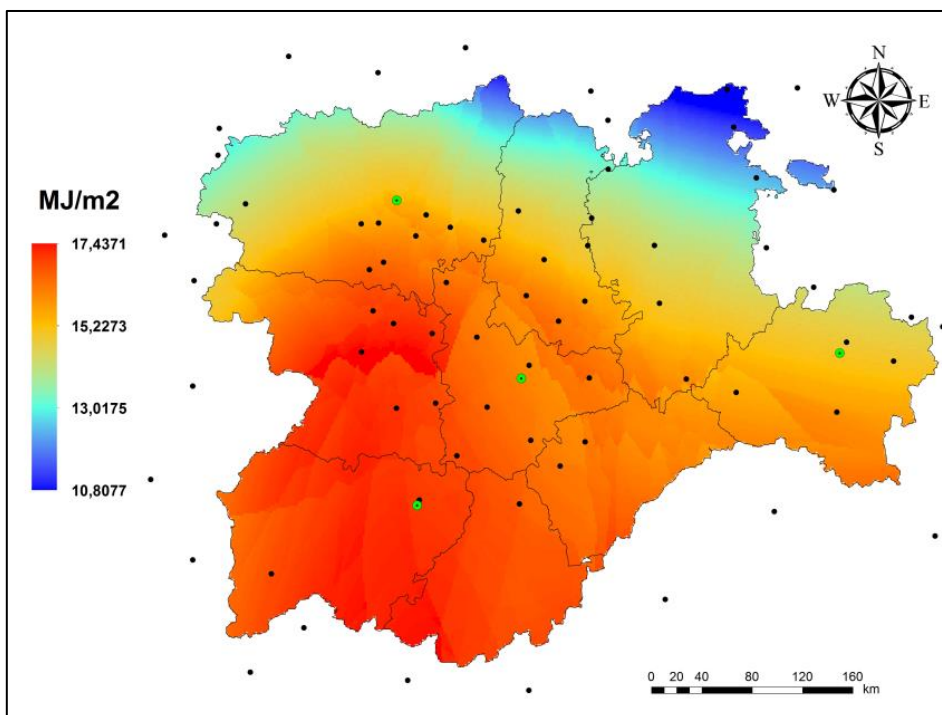


Figura 3.20. Mapa de interpolación Kriging Universal Cuadrático 35 puntos menor RMSE usando 71 estaciones como base y cuatro estaciones de la AEMET como control

Para comprobar que el resultados del control anterior es suficientemente bueno, se ha realizado un control adicional con 67 estaciones base (40 de la red SIAR y 27 de diferentes fuentes para la zona limítrofe) y ocho estaciones de validación: cuatro de la AEMET y cuatro de la red SIAR. Los resultados se han comparado con los resultados del control anterior y con los valores obtenidos para las estaciones de la red SIAR sin las de la AEMET. La Tabla 3.6 recoge los valores para uno de los métodos de interpolación, Splin con todas sus variantes.

De todo este proceso, realizado con los diferentes controles y métodos de interpolación, se ha seleccionado el que menos error presenta en la estimación de valores de irradiación en cualquier punto de la zona en estudio. Este apartado nos permite también generar los mapas de irradiación (objetivo parcial propuesto). Como muestra, la Figura 3.21 representa uno de ellos, el obtenido para el método de interpolación SPLIN de tensión con peso 5 y 12 puntos, que permite obtener valores de irradiación media anual diaria, en cualquier punto de la región, con un error menor del 3 %

Tabla 3.6. Errores de Splin con 12 puntos, regularizado y de tensión con estaciones de AEMET, con 44 y 40 estaciones del SIAR

| ESTACIONES AEMET (44 SIAR) | | | | RMSE | MBE | MAE | MAPE |
|----------------------------|--------------|------|-------|---------|----------|---------|---------|
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.30630 | -1.40188 | 2.97850 | 2.98550 |
| | | | 0.01 | 3.19131 | -1.21845 | 2.94316 | 2.94700 |
| | | | 0.001 | 3.07994 | -1.09498 | 2.86831 | 2.86939 |
| | TENSIÓN | PESO | 1 | 2.99927 | -1.01497 | 2.80445 | 2.80388 |
| | | | 5 | 3.00711 | -0.98352 | 2.82405 | 2.82351 |
| | | | 10 | 3.00578 | -0.95922 | 2.82990 | 2.82924 |
| ESTACIONES AEMET (40 SIAR) | | | | RMSE | MBE | MAE | MAPE |
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.26819 | -1.44281 | 2.92112 | 2.92856 |
| | | | 0.01 | 3.15661 | -1.25998 | 2.88924 | 2.89342 |
| | | | 0.001 | 3.05441 | -1.12397 | 2.83025 | 2.83156 |
| | TENSIÓN | PESO | 1 | 2.98135 | -1.04023 | 1.38804 | 2.77567 |
| | | | 5 | 2.99073 | -1.00867 | 2.79738 | 2.79699 |
| | | | 10 | 2.99040 | -0.98364 | 2.80463 | 2.80410 |

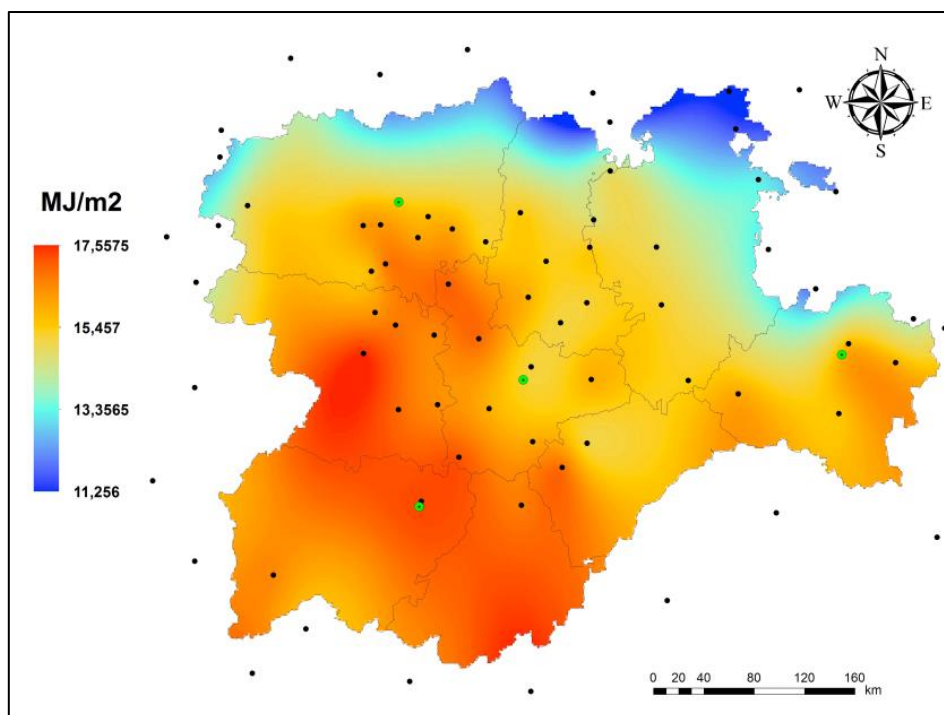


Figura 3.21. Mapa de irradiación de Castilla y León

3.4 MODELOS DE RADIACIÓN SOLAR

La última parte del trabajo se ha realizado con modelos matemáticos, para calcular la irradiación global sobre superficie inclinada (en este caso el ángulo de inclinación es de 90°), a partir de valores de irradiación global, directa y difusa horizontales. Como se ha comentado en otras ocasiones, el objetivo es poder calcular la producción eléctrica de una instalación FV integrada en fachadas. Para ello es necesario conocer el valor de la irradiación solar vertical en las cuatro orientaciones cardinales (norte, sur, este y oeste). Es conveniente recordar que normalmente se trabaja con planos orientados hacia el sur, ya que esa es la orientación considerada óptima para recoger la máxima cantidad de radiación, pero es interesante conocer la irradiación en otras orientaciones, para no descartar de antemano ninguna posibilidad. En este caso, los datos experimentales disponibles serán los de la EPS de Burgos.

Se van a seleccionar modelos matemáticos de radiación con el objeto de obtener los valores de la irradiación solar vertical en las cuatro orientaciones.

Los valores experimentales disponibles para realizar el trabajo son los datos del año 2014 completo, salvo los días que se han eliminado por ausencia de cualquiera de los valores recogidos. Una vez calculados (por medio de los modelos), los valores de irradiación global vertical (para las cuatro orientaciones), se comparan con los medidos o experimentales. Para calcular el error entre los valores teóricos (obtenidos por los modelos) y los experimentales, se emplean los cuatro parámetros estadísticos más frecuentes en este tipo de trabajos. El análisis de los resultados permitirá establecer qué modelo se ajusta más a la zona de estudio. Reseñar que se ha trabajado únicamente con cuatro modelos, los más clásicos, porque el fin de este trabajo no es el de comparar el comportamiento de los modelos que hay en la bibliografía, sino la posibilidad de disponer paneles fotovoltaicos en cualquier orientación con un rendimiento suficiente.

3.4.1. PROCESO DE CÁLCULO

Después de someter a los datos a un proceso de calidad, se obtienen dos ficheros con 52560 filas (valores diezminutales). El primero tiene los valores de todas las variables climáticas incluyendo los cuatro sensores de radiación vertical y el segundo los valores registrados por el seguidor, irradiación directa, difusa y global horizontal recogidos a lo largo del año 2014 en la instalación de la EPS de Burgos.

Los pasos seguidos para realizar el proceso son los siguientes:

1.- Se leen los datos de los dos ficheros establecidos. Se generan dos nuevos ficheros denominados "*Year_2014.mat*" con los valores de irradiación global, directa y difusa horizontales y "*global_90.mat*" con los valores de irradiación global vertical de las cuatro orientaciones.

2.- Por medio de un fichero llamado "*Prueba_indice*", se toman los valores de global vertical de las cuatro orientaciones y los de global, difusa y directa y los ajusta para que coincidan en número de datos con las fechas de la matriz "Caja". Por defecto se crea una matriz de datos.

3.- La matriz resultante con todos los datos se guarda en un fichero, “*Datos_2014.mat*”, que se va a usar para ejecutar el programa de cálculo, “*Rehaciendo.m*”.

4.- Cuando se ejecuta el programa, se obtiene una matriz resultante (fichero de datos) con 120 datos por día (de 3:00 a 23:00 h), un dato cada 10 minutos. La matriz tiene 13 columnas y 43800 filas

5.- El programa, “*Rehaciendo.m*”, ejecuta un lazo idéntico, para cada uno de los valores diezminutales, calculando los parámetros necesarios para cada modelo:

- a) declinación solar δ
- b) factor de corrección de excentricidad de la órbita E_0
- c) número de horas de sol n
- d) ángulos horarios correspondientes al amanecer y al anochecer ω_{sr} y ω_{ss}
- e) ángulos horarios ω_s
- f) ángulo cenital del sol θ_z
- g) ángulo de incidencia de los rayos solares con la perpendicular a la vertical θ
- h) irradiación extraterrestre horaria horizontal I_0
- i) irradiación extraterrestre horaria vertical en cuatro orientaciones $I_{0\beta\gamma}$

6.- Una vez calculados todos los parámetros, se calcula la irradiación global vertical teórica con cada modelo y en cada orientación. En la Tabla 3.26 se han recogido los valores teóricos de irradiación global vertical sur para los cuatro modelos y los valores experimentales para el mes de enero.

7.- El programa permite obtener los valores de irradiación vertical para un modelo y para una cualquiera de las orientaciones, de esta forma se puede

hacer el estudio con uno, dos, tres o los cuatro modelos, en una o varias orientaciones.

8.- A continuación se comparan los valores experimentales con los teóricos (recogidos por los sensores verticales). Este proceso se hace para cada modelo y con los cuatro evaluadores estadísticos más empleados en la bibliografía, RMSE, MBE, MAE y MAPE.

Tabla 3.7. Valores de radiación global vertical orientación sur para los cuatro modelos (circunsolar, isotrópico, Klucher y Hay) y experimentales ($\text{MJ}\cdot\text{día}^{-1}$)

| RESULTADOS - SUR | | | | | |
|------------------|-----------------------|---------------------|--------------------|-----------------|-------------|
| DÍA | VERTICAL EXPERIMENTAL | TEÓRICO CIRCUNSOLAR | TEÓRICO ISOTRÓPICO | TEÓRICO KLUCHER | TEÓRICO HAY |
| 1 | 0,50 | 3,64 | 0,76 | 0,34 | 0,80 |
| 2 | 1,17 | 6,99 | 1,49 | 0,68 | 1,56 |
| 3 | 1,25 | 7,27 | 1,58 | 0,73 | 1,65 |
| 4 | 2,34 | 9,45 | 6,07 | 5,63 | 6,25 |
| 5 | 0,44 | 3,21 | 0,69 | 0,31 | 0,72 |
| 6 | 3,95 | 13,85 | 6,87 | 5,90 | 7,37 |
| 7 | 3,82 | 13,74 | 5,66 | 4,48 | 6,12 |
| 8 | 16,07 | 45,80 | 40,43 | 40,40 | 43,29 |
| 9 | 9,93 | 27,86 | 19,64 | 18,70 | 21,41 |
| 10 | 10,37 | 29,00 | 20,23 | 19,19 | 22,05 |
| 11 | 10,74 | 29,90 | 21,95 | 21,05 | 23,74 |
| 12 | 4,66 | 15,89 | 8,44 | 7,32 | 8,90 |
| 13 | 3,30 | 11,00 | 5,47 | 4,66 | 5,81 |
| 14 | 1,59 | 7,94 | 1,85 | 0,85 | 1,93 |
| 15 | 1,55 | 7,38 | 1,90 | 0,99 | 2,00 |
| 16 | 2,00 | 7,99 | 2,92 | 2,09 | 3,06 |
| 17 | 11,33 | 30,98 | 22,98 | 22,08 | 25,28 |
| 18 | 1,63 | 7,62 | 1,97 | 0,99 | 2,06 |
| 19 | 4,14 | 13,22 | 5,72 | 4,45 | 6,12 |
| 20 | 5,80 | 18,17 | 9,43 | 7,99 | 10,17 |
| 21 | 0,78 | 4,50 | 1,10 | 0,48 | 1,15 |
| 22 | 5,45 | 17,60 | 11,61 | 10,64 | 12,11 |
| 23 | 4,93 | 15,64 | 7,78 | 6,39 | 8,24 |
| 24 | 6,86 | 18,06 | 10,50 | 9,25 | 11,44 |
| 25 | 6,58 | 19,00 | 11,47 | 10,19 | 12,28 |
| 26 | 1,40 | 6,69 | 1,75 | 0,77 | 1,82 |
| 27 | 18,06 | 47,85 | 42,90 | 42,84 | 45,73 |
| 28 | 2,41 | 8,88 | 5,20 | 4,51 | 5,37 |
| 29 | 6,63 | 17,77 | 8,76 | 7,01 | 9,50 |
| 30 | 10,52 | 26,36 | 19,11 | 17,99 | 20,94 |
| 31 | 1,77 | 7,45 | 2,08 | 0,92 | 2,17 |

CAPÍTULO 4.

ANÁLISIS ESTADÍSTICO DE LOS DATOS

4.1 INTRODUCCIÓN

Una vez realizado el control de los datos con los que se va a trabajar, garantizando que la fiabilidad de los mismos es adecuada, se ha realizado un doble análisis estadístico y de cálculo de errores. Por una parte se ha trabajado con los datos para ver la tendencia y la eficiencia de los mismos y por otra se le ha aplicado un proceso de regresión múltiple para ver la relación existente entre la radiación global y el punto geográfico en el que se toma la variable, es decir expresar la irradiación en función de la latitud, longitud y altitud de la estación de medida.

4.2 ESTUDIO DE LOS DATOS DE IRRADIACIÓN

Una vez obtenidos los valores medios de irradiación media anual en las estaciones de Castilla y León se ha realizado un estudio estadístico de los datos, mediante un análisis de la varianza de un factor (*ANOVA Analysis Of Variance (Fisher, (1925))*). Este análisis permite contrastar la hipótesis de que una de las medias difiera del resto, en lo que a su valor esperado se refiere. Para ello se hace un ANOVA Simple con el programa Statgraphics (<http://www.statgraphics.net>). Este proceso hace una descomposición de la variación de los datos respecto a la media, construyendo pruebas y gráficas para analizar el comportamiento de los valores medios de irradiación.

En nuestro caso, los datos registrados son los valores de irradiación recibidos en distintas estaciones (44) a lo largo de siete años, desde el año 2007 hasta el 2013. En este estudio se desea comprobar si el nivel medio de irradiación recibido en estos años es igual frente a la hipótesis alternativa, de que al

menos dos de las medias con las que se trabaja (media de los años 2009 y 2013), son diferentes. El resultado obtenido se recoge en la Tabla 4.1, donde se estudia, por una parte con los valores entre grupos (un grupo por año), que miden la variabilidad entre los valores de irradiación de los distintos años y por otra analiza la variabilidad de los valores dentro de los grupos y la variabilidad total. Se establece el contraste de hipótesis de irradiaciones medias iguales a partir de una distribución F-Snedecor (*Snedecor y Cochran, (1991)*) para las características del problema planteado con los grados de libertad determinados. El estadístico “*Razón-F*”, obtenido como cociente entre los cuadrados medios, es el estadístico de contraste que nos permite aceptar o rechazar la hipótesis de igualdad de irradiación media en los distintos años. El “*P-Valor*” determina el nivel de significancia de la prueba. Un P-Valor menor que 0.05 nos permite rechazar la hipótesis establecida con el 95 % de confianza.

Tabla 4.1. ANOVA para Irradiación Media Anual ($\text{MJ}\cdot\text{m}^{-2}$) por año

| <i>Fuente</i> | <i>Suma de Cuadrados</i> | <i>Gl</i> | <i>Cuadrado Medio</i> | <i>Razón-F</i> | <i>P-Valor</i> |
|---------------|--------------------------|-----------|-----------------------|----------------|----------------|
| Entre grupos | 80.007 | 6 | 13.3344 | 15.08 | 0.0000 |
| Intra grupos | 253.721 | 287 | 0.8841 | | |
| Total (Corr.) | 333.727 | 293 | | | |

Viendo los resultados del ANOVA Simple, especialmente los valores de “*Razón-F*” y “*P-Valor*” se puede afirmar que existen diferencias significativas, en cuanto al nivel medio de irradiación recibida en los distintos años, es decir, no coincide en, al menos, dos de los años.

Dado que existen diferencias significativas, se ha obtenido la tabla de irradiación media para cada año. En ella se muestran los valores medios de irradiación junto con los errores estimados estándar (empleando la desviación estándar agrupada), y los correspondientes intervalos de confianza para dichas medias. Estos resultados se recogen en la Tabla 4.2 donde se puede observar también el número de datos disponibles para cada año. En las últimas columnas se representan los valores límite en los intervalos de confianza del 95 %.

Tabla 4.2. Medias para Irradiación Media Anual ($\text{MJ}\cdot\text{m}^{-2}$) por año con intervalos de confianza del 95 %

| Año | Casos | Error Est. | | | |
|-------|-------|-------------------------------------|--------------|---|---|
| | | Media $\text{MJ}\cdot\text{m}^{-2}$ | (s agrupada) | Límite Inferior $\text{MJ}\cdot\text{m}^{-2}$ | Límite Superior $\text{MJ}\cdot\text{m}^{-2}$ |
| 2007 | 38 | 16.6202 | 0.152527 | 16.4080 | 16.8325 |
| 2008 | 44 | 16.3137 | 0.141746 | 16.1164 | 16.5110 |
| 2009 | 42 | 17.2928 | 0.145082 | 17.0909 | 17.4947 |
| 2010 | 43 | 15.8334 | 0.143385 | 15.6338 | 16.0330 |
| 2011 | 41 | 16.4361 | 0.146840 | 16.2317 | 16.6405 |
| 2012 | 42 | 16.3891 | 0.145082 | 16.1872 | 16.5910 |
| 2013 | 44 | 15.5516 | 0.141746 | 15.3543 | 15.7489 |
| Total | 294 | 16.3367 | | | |

Si los valores de la Tabla 4.2 se representan gráficamente, se obtienen las medias graficadas junto con los intervalos de incertidumbre, recogidos en la Figura 4.1. En ella se puede ver que los valores de los años 2007, 2008, 2011 y 2012 se traslapan. Esto implica que no hay diferencias estadísticamente significativas entre ellos. Los años 2010 y 2013 también se traslapan entre ellos pero no con el grupo anterior y el año 2009 es totalmente independiente de los otros dos grupos. Esto nos permite determinar que entre los valores del año 2009 y el resto hay una diferencia estadísticamente significativa en el nivel de confianza del 95 %.

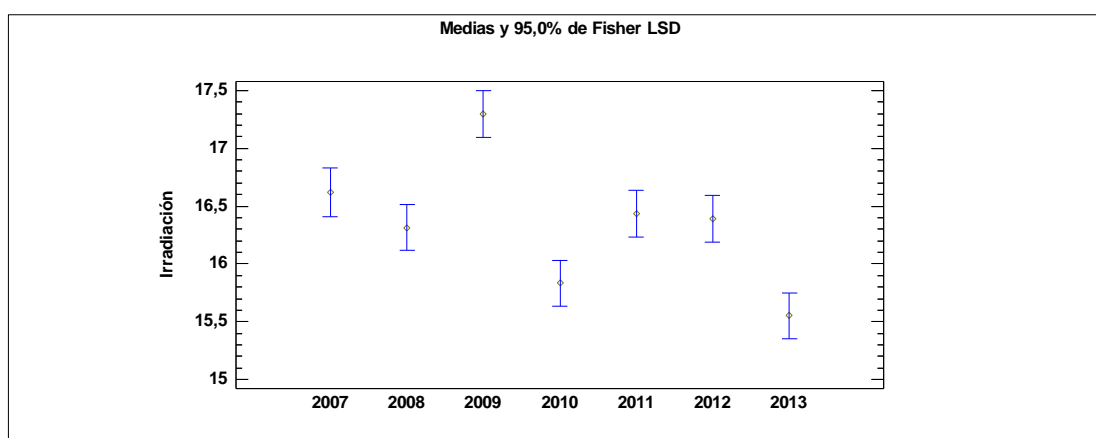


Figura 4.1. Gráfico de valores medios de Irradiación Global Anual ($\text{MJ}\cdot\text{m}^{-2}$) con intervalo de confianza del 95 %

En la Figura 4.1 se puede observar que todos los tamaños muestrales representados son aproximadamente iguales, y se puede determinar qué valor medio de irradiación es significativamente diferente a otro. Para ello se determinan los intervalos LSD (*Least Significant Difference*), obtenidos en torno a la media anual de irradiación y en los que la amplitud está influida por una constante definida a través de la distribución *t de Student*. Si los intervalos asociados a dos años se traslapan, entre dichos años no hay diferencias estadísticamente significativas.

Para conocer qué valores medios de irradiación son significativamente diferentes unos de otros se establece el criterio de “*Prueba de Rangos Múltiples*” con el procedimiento LSD (*Fisher, (1925)*). Para ello se calcula la diferencia entre los valores medios de la muestra y se realiza una estimación de la diferencia en la que interviene una constante que depende del procedimiento de comparación elegido, y que en el caso del método LSD está definido, como anteriormente, a través de la distribución *t de Student*. Los resultados se recogen en la Tabla 4.3. Este método se ha diseñado para comparar cualquier par de medidas con el nivel de confianza establecido.

Tabla 4.3. Pruebas de Rangos Múltiples para Irradiación por año

| <i>Año</i> | <i>Casos</i> | <i>Media / MJ·m⁻²</i> | <i>Grupos Homogéneos</i> |
|------------|--------------|----------------------------------|--------------------------|
| 2013 | 44 | 15.5516 | X |
| 2010 | 43 | 15.8334 | X |
| 2008 | 44 | 16.3137 | X |
| 2012 | 42 | 16.3891 | X |
| 2011 | 41 | 16.4361 | X |
| 2007 | 38 | 16.6202 | X |
| 2009 | 42 | 17.2928 | X |

Como se puede observar, en la tabla quedan clasificados los años por grupos homogéneos. La tabla 4.3 informa los qué años tienen un comportamiento similar en cuanto al nivel de irradiación recibido y pone de manifiesto dónde se encuentran las diferencias puestas anteriormente de manifiesto en el gráfico de medias. Si se analiza la última columna “*Grupos homogéneos*”, se puede ver que hay tres grupos homogéneos, (en cuanto a la irradiación media recibida en

el año), de forma que estadísticamente los años 2010 y 2013 son equivalentes, los años 2008, 2012, 2011 y 2007 también lo son y por último el año 2009 no es homogéneo con ninguno de los anteriores. De esta información se puede concluir que el nivel medio de irradiación del año 2009 es significativamente mayor que en el resto de los años, y es la forma numérica del resultado obtenido en la Figura 4.1.

Para dar más peso a esta conclusión, se ha hecho un estudio de los intervalos de confianza para la diferencia de medias al 95 %, es decir se hace un estudio comparando los años de dos en dos, para ver entre que pares existen las diferencias más significativas. En la Tabla 4.4 se recogen las diferencias de cada año con los años que no pertenecen a su grupo. La primera columna de esta tabla muestra los años comparados; la segunda columna, la diferencia de medias de los dos años; la tercera, los límites establecidos según Fisher; la cuarta, la diferencia entre el límite inferior del primer año y el superior del segundo y la quinta, la diferencia entre el límite superior del primero y el inferior del segundo.

Tabla 4.4. Diferencia de valores medios de Irradiación Global sobre plano horizontal

| <i>Contraste</i> | <i>Diferencia MJ·m⁻²</i> | <i>+/- Límites</i> | <i>Inferior</i> | <i>Superior</i> |
|------------------|---|--------------------|-----------------|-----------------|
| 2007 - 2009 | -0.6726 | 0.4143 | -1.0869 | -0.2583 |
| 2007 - 2010 | 0.7868 | 0.4120 | 0.3747 | 1.1989 |
| 2007 - 2013 | 1.0686 | 0.4098 | 0.6588 | 1.4785 |
| 2008 - 2009 | -0.9791 | 0.3992 | -1.3784 | -0.5799 |
| 2008 - 2010 | 0.4803 | 0.3968 | 0.0834 | 0.8771 |
| 2008 - 2013 | 0.7621 | 0.3946 | 0.3675 | 1.1566 |
| 2009 - 2010 | 1.4594 | 0.4015 | 1.0579 | 1.8609 |
| 2009 - 2011 | 0.8567 | 0.4063 | 0.4504 | 1.2630 |
| 2009 - 2012 | 0.9037 | 0.4038 | 0.4999 | 1.3080 |
| 2009 - 2013 | 1.7412 | 0.3993 | 1.3420 | 2.1405 |
| 2010 - 2011 | -0.6027 | 0.4040 | -1.0067 | -0.1987 |
| 2010 - 2012 | -0.5557 | 0.4015 | -0.9572 | -0.1542 |
| 2011 - 2013 | 0.8845 | 0.4017 | 0.4828 | 1.2862 |
| 2012 - 2013 | 0.8375 | 0.3993 | 0.4383 | 1.2367 |

Para validar el estudio, se ha contrastado la homogeneidad de las varianzas, mediante la prueba de verificación de varianzas. Para ello se establece como hipótesis la nula igualdad de las varianzas. Realizado el estadístico correspondiente, el resultado de la prueba se ha recogido en la Tabla 4.5

Tabla 4.5. Verificación de Varianza

| | Prueba | P-Valor |
|----------|--------|---------|
| Levene's | 0.3961 | 0.8813 |

Al obtener un “*P-valor*” superior a 0.05, no se rechaza la hipótesis nula, lo que nos permite concluir que las varianzas son iguales.

A través de un gráfico de cajas y bigotes, podemos representar los valores que están entre el primer y el tercer cuartil (por medio de cajas), así como los puntos atípicos (puntos cuya distancia a la caja es mayor que 1.5 veces el ancho de la caja). El resultado se puede ver en la Figura 4.2. En el gráfico queda reflejado que la media está situada todos los años a la izquierda del segundo cuartil (mediana) y se observó que todos los años hay puntos atípicos. También se puede comprobar la existencia de puntos atípicos extremos, que son los que están alejados más de tres veces el ancho de la caja. Como puede verse en la Figura 4.2, los puntos extremos aparecen siempre con los menores valores de irradiación y se identifican con las estaciones del noreste de la provincia de Burgos, que no pertenecen a la cuenca del Duero.

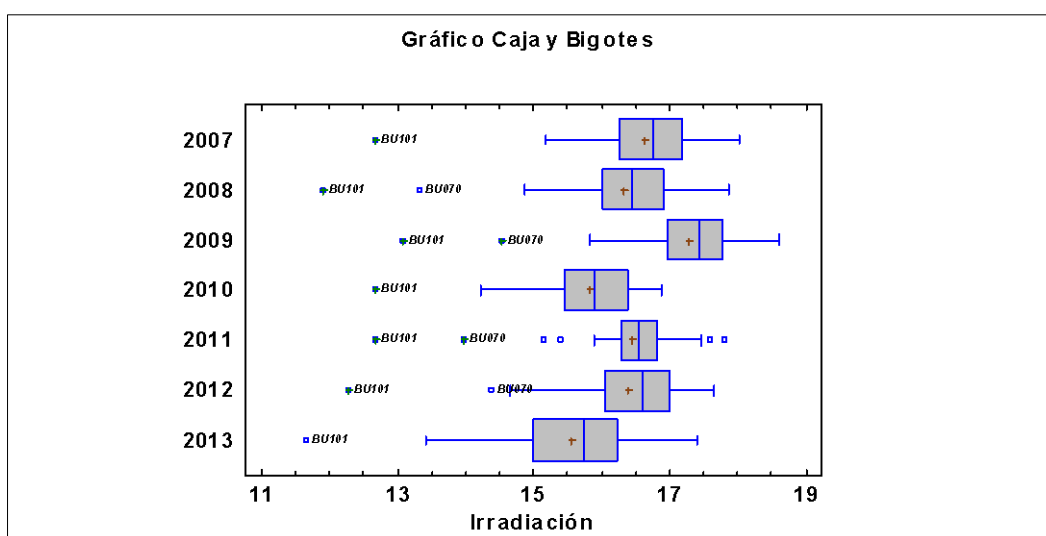


Figura 4.2. Valores entre el primer y tercer cuartil y puntos atípicos (MJ·m⁻²)

Se ha hecho un estudio de la existencia de puntos atípicos para el valor anual de irradiación, y para completar el estudio de los datos se ha realizado una identificación de los valores atípicos de la irradiación media de los siete años. Para ello, se trabaja con los 44 valores medios de irradiación, tomados como variable, que presentan un rango que va desde 12.4121 MJ·m⁻² a 17.4742 MJ·m⁻². En la Tabla 4.6 se representa la estadística que estima el centro de la población con la que trabajamos

Tabla 4.6. Estimado de Localización

| | |
|-------------------|----------------------------|
| Media muestral | 16.3372 MJ·m ⁻² |
| Mediana muestral | 16.4816 MJ·m ⁻² |
| Media recortada | 16.4630 MJ·m ⁻² |
| Media winsorizada | 16.4682 MJ·m ⁻² |

En esta tabla se recoge el valor de la media aritmética, 16.3372 MJ·m⁻²; el centro o valor medio de la muestra, 16.4816 MJ·m⁻²; el valor promedio después de eliminar el 15 % de los valores menores y mayores, 16.463 MJ·m⁻², y por último, el valor promedio después de sustituir el 15 % de los valores menores y mayores con los valores más extremos no incluidos dentro de ese porcentaje, 16.4682 MJ·m⁻².

Se ha procedido a estudiar la dispersión de los datos, realizando una estimación de la desviación estándar, a través de los denominados “Estimados de Escala” cuyos resultados se recogen en la Tabla 4.7

Tabla 4.7. Estimados de Escala

| | |
|---------------------------|---------|
| Desviación estd. muestral | 0.87416 |
| DAM/0.6745 | 0.57234 |
| Sbi | 0.53669 |
| Sigma winsorizada | 0.58867 |

En la Tabla 4.7 se tiene el valor de la desviación estándar muestral, 0.87416 MJ·m⁻²; la estimación basada en la desviación absoluta mediana (la mediana de la diferencia absoluta entre los valores y la mediana muestral): 0.57234; la

estimación basada en una suma ponderada de cuadrados alrededor de la media muestral, donde los pesos disminuyen con la distancia a partir de la mediana: 0.53669 y por último la estimación basada en los desviaciones al cuadrado alrededor de la media “winsorizada”: 0.58870.

La influencia que pueden tener los valores atípicos, se ha estudiado mediante los intervalos de confianza del 95 % para las medias y desviaciones estándar y la estadística “winsorizada”. Los valores se recogen en la Tabla 4.8

Tabla 4.8. Intervalos de confianza del 95.0 % para la media

| | Límite Inferior MJ·m ⁻² | Límite Superior MJ·m ⁻² |
|-------------|---------------------------------------|---------------------------------------|
| Estándar | 16.0714 | 16.6030 |
| winsorizada | 16.2560 | 16.6805 |

Al ser los valores límites próximos, se puede decir que los valores atípicos no son problema para el conjunto de datos.

La gráfica de aberrantes (Figura 4.3) presenta en el eje vertical izquierdo los valores de irradiación en MJ·m⁻², en el eje horizontal el número de fila que representa a cada estación y líneas horizontales. La línea central representa la media de los valores de irradiación y las paralelas están dibujadas a una distancia de la central y entre sí igual al valor de la desviación estándar.

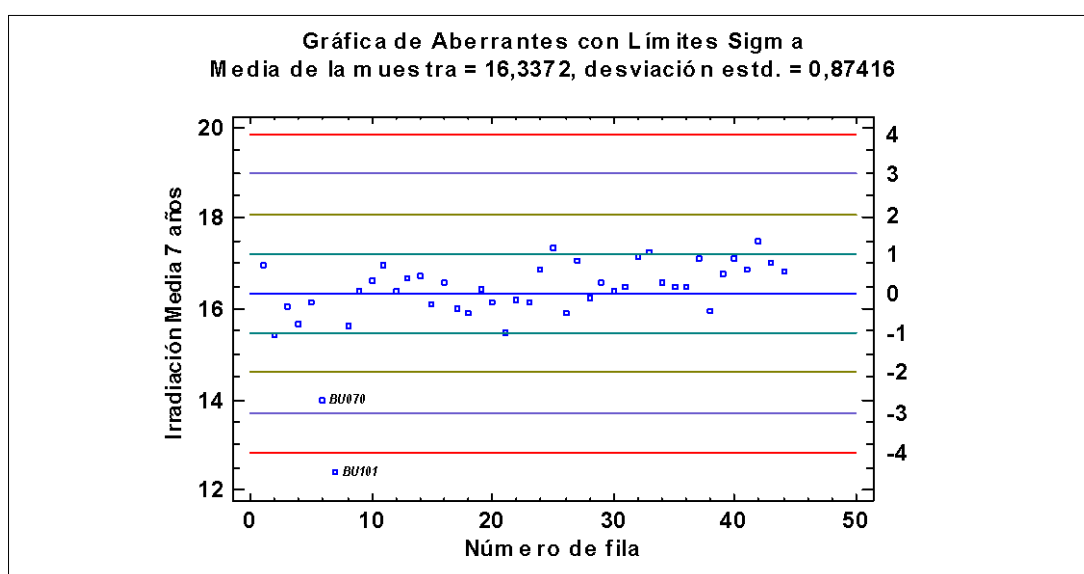


Figura 4.3. Gráfica de aberrantes

A partir de los valores de irradiación media de 7 años, ordenados de menor a mayor, se recogen en la Tabla 4.9 los cinco valores menores y los cinco valores mayores, indicando la fila en la que aparece el valor (determinada por la estación correspondiente), y valores estandarizados que pueden ayudar a identificar los valores atípicos.

La columna central, “*Valores estudentizados sin supresión*”, mide cuantas veces la diferencia entre el valor medio de irradiación y la media muestral es múltiplo de la desviación estándar y corresponde a la escala del eje de la derecha de la gráfica de valores atípicos. La cuarta columna, “*Valores estudentizados con supresión*”, representa la misma situación que el caso anterior, cuando ese valor no se incluye en la muestra. Y la última columna, “*modificados*”, estos valores usan la estimación de sigma basada en la desviación absoluta mediana (DAM).

Tabla 4.9. Valores Ordenados

| Fila | Valor | Valores estudentizados | Valores estudentizados | modificados |
|------|---------|------------------------|------------------------|-------------|
| | | Sin Supresión | Con Supresión | Valor-Z DAM |
| 7 | 12.4121 | -4.4901 | -6.2956 | -7.1103 |
| 6 | 13.9664 | -2.7121 | -3.0197 | -4.3947 |
| 2 | 15.4434 | -1.0225 | -1.0472 | -1.8141 |
| 21 | 15.4552 | -1.0090 | -1.0329 | -1.7934 |
| 8 | 15.6329 | -0.8057 | -0.8212 | -1.4829 |
| ... | ... | ... | ... | ... |
| 40 | 17.1188 | 0.8942 | 0.9129 | 1.1133 |
| 32 | 17.1595 | 0.9406 | 0.9614 | 1.1843 |
| 33 | 17.2612 | 1.0571 | 1.0835 | 1.3621 |
| 25 | 17.3493 | 1.1578 | 1.1900 | 1.5160 |
| 42 | 17.4742 | 1.3006 | 1.3426 | 1.7341 |

Se ha realizado una última prueba formal para los valores atípicos, denominada “*Prueba de Grubbs*” o “*Prueba de la Desviación estudentizada extrema*” (ESD). En ella se obtienen un valor aproximado de “*P-Valor*” calculando la probabilidad, basada en la distribución t de Student, de exceder el estadístico que emplea utilizando el mayor valor estudentizado sin supresión. En este caso

se ha aplicado al valor -4.4901, que a su vez es el que presenta un mayor valor absoluto de “Valor-Z DAM”; al ser el “P-Valor” obtenido de $8.33 \cdot 10^{-6}$, se puede concluir que ese valor extremo es verdaderamente atípico.

Resumiendo se puede decir que, para los 44 valores de irradiación media de 7 años, la media y sigma son 16.3372 y $0.8742 \text{ MJ}\cdot\text{m}^{-2}$, respectivamente. Los correspondientes estimados winsorizados, en los cuales se ha remplazado el 15.0 % de los valores más grandes y más pequeños con valores del interior de la muestra, son 16.4682 y $0.5887 \text{ MJ}\cdot\text{m}^{-2}$.

La tabla de valores ordenados muestra los valores extremos de irradiación global media en 7 años. Los valores estudentizados miden la relación entre la diferencia de cada valor con la media muestral ($16.3372 \text{ MJ}\cdot\text{m}^{-2}$) y la desviación estándar. El valor más extremo se encuentra en la fila siete (correspondiente a la estación BU101), para el cual la relación anterior es 4.49013. Puesto que el P-valor para la prueba de Grubb es menor que 0.05, ese valor es un aberrante significativo con un nivel de significancia del 5 %, asumiendo que todos los demás valores siguen una distribución normal. Al calcular las estadísticas muestrales después de eliminar cada punto, uno a la vez, al igual que cuando la media y la desviación estándar están basadas en la desviación absoluta de la mediana (DAM). Cualquier punto para el que el valor absoluto de “Valor Z DAM” sea mayor que 3.5 se puede considerar observación aberrante. En el caso de estudio, esta circunstancia se presenta en dos casos, las filas seis y siete que corresponden con las estaciones BU070 y BU101.

Si se representa la gráfica de caja y bigotes para la irradiación global media de 7 años, se comprueba que esos posibles valores aberrantes coinciden con los valores atípicos identificados en la gráfica 4.4.

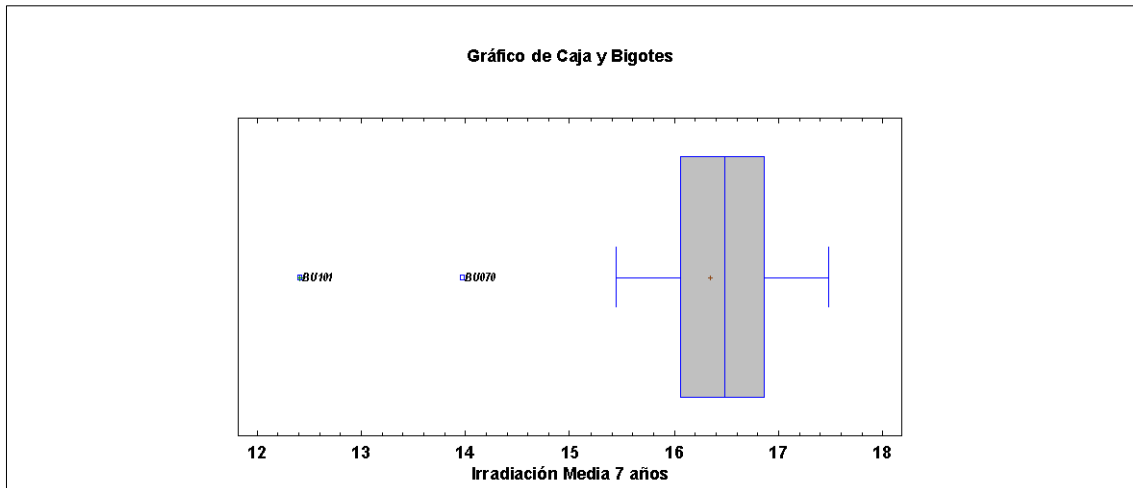


Figura 4.4. Gráfica de Caja y Bigotes para Irradiación Global Media de 7 años ($\text{MJ}\cdot\text{m}^{-2}$)

4.3 IRRADIACIÓN Y COORDENADAS GEOGRÁFICAS DE LAS ESTACIONES

Diversos autores relacionan la irradiación solar con parámetros geográficos (*Kumar et al. (2005)*, *Chen et al. (2006)*). Mediante estudios de regresión simple y múltiple se ha buscado, con ayuda del programa Statgraphics, relacionar los valores medios de irradiación media anual de los siete años con las coordenadas geográficas de las 44 estaciones de la red SIAR.

Se ha realizado mediante el cálculo de regresión simple utilizando como variable dependiente el valor medio de irradiación (es conveniente recordar que este valor se obtiene como media de los valores medios diarios de los siete años con los que se trabaja), y como variable independiente la latitud de la estación, ajustando los valores a un modelo lineal:

$$\text{Irradiación Media 7 años} = \alpha_1 + \alpha_2 * \text{latitud}$$

Los coeficientes característicos se han recogido en la Tabla 4.10.

Tabla 4.10. Coeficientes de la regresión simple de la Irradiación Global Media 7 años con la latitud

| <i>Parámetro</i> | <i>Mínimos Cuadrados Estimado</i> | <i>Error Estándar</i> | <i>Estadístico T</i> | <i>P-Valor</i> |
|------------------|-----------------------------------|-----------------------|----------------------|----------------|
| α_1 | 54.709 | 8.7116 | 6.2800 | 0.0000 |
| α_2 | -0.9149 | 0.2077 | -4.4050 | 0.0001 |

En la Tabla 4.10 el contraste estadístico “estadístico-T”, establecida la hipótesis nula (el parámetro correspondiente es igual a 0) permite, apoyándonos en los correspondientes valores de la t de Student, rechazar la hipótesis nula en los dos casos al obtener un “P-Valor” menor que 0.05 en ambos casos, y el coeficiente “Pendiente”, al ser distinto de cero, prueba la dependencia de la irradiación con la latitud. Es decir se puede afirmar que existe relación estadísticamente significativa entre la irradiación y la latitud con un nivel de confianza del 95 %. La expresión matemática que recoge la dependencia entre ambas variables y que se ajusta al modelo, se puede expresar como:

$$\text{Irradiación Media 7 años} = 54.709 - 0.9149 * \text{latitud}$$

Se ha realizado también un análisis de la varianza, cuyos resultados están recogidos en la Tabla 4.11

Tabla 4.11. Análisis de Varianza en el ajuste lineal de la Irradiación Global Media (7 años) con la latitud

| <i>Fuente</i> | <i>Suma de Cuadrados</i> | <i>Gl</i> | <i>Cuadrado Medio</i> | <i>Razón-F</i> | <i>P-Valor</i> |
|---------------|--------------------------|-----------|-----------------------|----------------|----------------|
| Modelo | 10.3837 | 1 | 10.3837 | 19.40 | 0.0001 |
| Residuo | 22.4750 | 42 | 0.5351 | | |
| Total (Corr.) | 32.8587 | 43 | | | |

Al obtener la suma de cuadrados explicados por el modelo y la suma de cuadrados de los residuos, permite obtener la suma de cuadrados totales como suma de los anteriores y descomponer la varianza total como suma de la varianza explicada por el modelo más la varianza residual.

Los valores más importantes en este estudio son los de la “Razón-F” y su “P-Valor”, que representan la significancia estadística del modelo ajustado. En

este caso como el “P-Valor” es menor del 0.05 indica que existe la relación estadística planteada entre la irradiación y la latitud. Pero no es suficiente información como para determinar si esa relación es lo suficientemente importante o no. Para ello, hay que calcular los valores estadísticos recogidos en la Tabla 4.12.

Tabla 4.12. Estadísticos del ajuste lineal de la Irradiación Global Media 7 años con la latitud

| | |
|---|--|
| <i>Coefficiente de correlación = -0.5622</i> | |
| $R^2 = 31.60 \%$ | $R^2(\text{ajustado para g.l.}) = 29.97 \%$ |
| <i>Error estándar del est. = 0.7315</i> | <i>Error absoluto medio = 0.5040</i> |
| <i>Estadístico Durbin – Wastson =</i> 1.16851 ($P = 0.0013$) | <i>Autocorrelación de residuos con retraso 1 =</i> 0.4143 |

El estadístico R^2 indica que el modelo ajustado explica el 31.60 % de la variabilidad en irradiación global media 7 años. El coeficiente de correlación 0.5622 indica una relación moderadamente fuerte entre las variables. El error estándar del estimador señala que la desviación estándar estimada de los residuos es 0.7315 (el residuo es la diferencia entre el valor real y el predicho).

El error absoluto medio (MAE) de 0.5040 es el valor promedio de los residuos, y el estadístico de Durbin-Watson (DW), que examina los residuos para determinar si hay alguna correlación significativa basada en el orden en el que se presentan en el archivo de datos, al tener un P-Valor de 0.0013 (menor que 0.05), establece que pudiera haber una posible correlación serial con un nivel de confianza del 95 %.

Al obtener un valor de R^2 de 31.6 % parece lógico pensar que la irradiación no depende únicamente de la latitud, por lo que se ha procedido a realizar otros estudios en los que intervengan otros factores, manteniendo la latitud en todos ellos, ya que no pareció lógico prescindir de esta variable al hacer el estudio de la variabilidad de la irradiación con las coordenadas geográficas de las estaciones.

Se ha hecho un estudio análogo al anterior pero haciendo intervenir en la relación de dependencia dos variables independientes. Primero se ha

procedido a realizar una regresión múltiple donde la variable dependiente es la irradiación y las dos variables independientes la latitud y la longitud de las estaciones, según el modelo:

$$\text{Irradiación Media 7 años} = \alpha_1 + \alpha_2 * \text{latitud} + \alpha_3 * \text{longitud}$$

Los resultados obtenidos para los coeficientes se han recogido en la Tabla 4.13.

Tabla 4.13. Regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)

| Parámetro | Estimación | Error Estándar | Estadístico T | P-Valor |
|------------|------------|----------------|---------------|---------|
| α_1 | 51.6374 | 7.5018 | 6.8833 | 0.0000 |
| α_2 | -0.8824 | 0.1781 | -4.9545 | 0.0000 |
| α_3 | -0.3616 | 0.0898 | -4.0286 | 0.0002 |

En la tabla 4.13 se puede observar que el “P-Valor” también es menor de 0.05, por lo que si hay relación estadística entre las variables independientes y la dependiente. El término correspondiente al “P-Valor” más alto es 0.0002, correspondiente a la variable independiente longitud, da a este término un valor estadístico suficientemente significativo. La ecuación del modelo ajustado para este caso se puede expresar:

$$\text{Irradiación Media 7 años} = 51.6374 - 0.8824 * \text{latitud} - 0.3616 * \text{longitud}$$

Como en el caso anterior se ha procedido a realizar un análisis de la varianza para tener mayor información, los resultados están recogidos en la Tabla 4.14.

Tabla 4.14. Análisis de Varianza de la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)

| Fuente | Suma de Cuadrados | Gl | Cuadrado Medio | Razón-F | P-Valor |
|---------------|-------------------|----|----------------|---------|---------|
| Modelo | 16.7572 | 2 | 8.3786 | 21.33 | 0.0000 |
| Residuo | 16.1014 | 41 | 0.3927 | | |
| Total (Corr.) | 32.8587 | 43 | | | |

Los indicadores de variabilidad muestran en la Tabla 4.15

Tabla 4.15. Estadísticos de la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)

| | |
|---|---|
| $R^2 = 51.00 \%$ | $R^2(\text{ajustado para g.l.}) = 48.61 \%$ |
| Error estándar del est. = 0.6267 | Error absoluto medio = 0.4166 |
| Estadístico Durbin – Wastson = 1.1506 (P = 0.0007) | Autocorrelación de residuos con retraso 1 = 0.4225 |

El estadístico R^2 indica que el modelo así ajustado explica el 51 % de la variabilidad en irradiación media 7 años. La R^2 ajustada, que es más apropiada que R^2 para comparar modelos con diferente número de variables independientes, es 48.61 %. El Error estándar del estimador muestra que la desviación estándar estimada de los residuos es 0.6267. El error absoluto medio (MAE) es 0.4166. El estadístico de Durbin-Watson (DW) al ser el “P-Valor” 0.0007 que es menor que 0.05, indica una posible correlación serial con un nivel de confianza del 95 %.

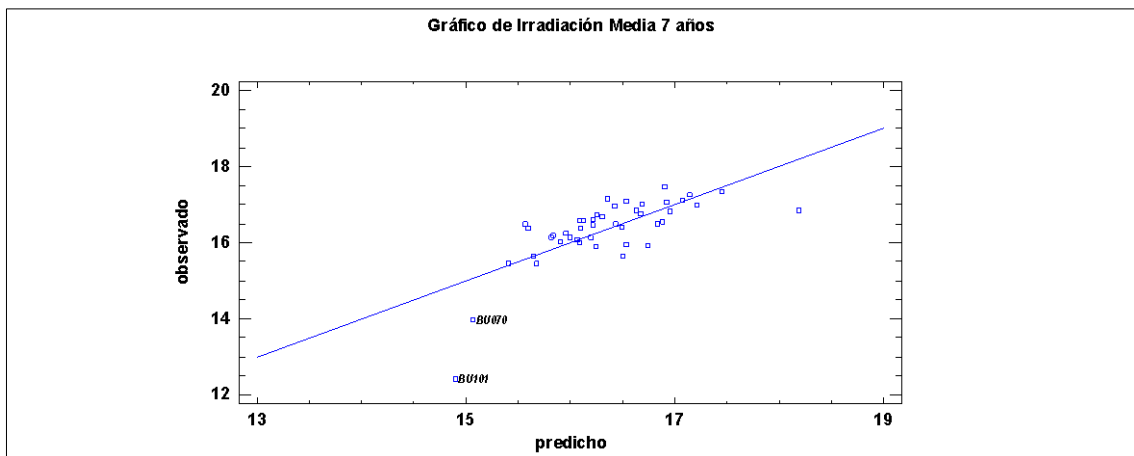


Figura 4.5. Regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)

En la Figura 4.5 se muestran los valores observados frente a los predichos por el modelo. Se puede apreciar que no hay mucha dispersión entre ellos, y la mayoría están en las proximidades de la recta.

Se ha analizado los residuos atípicos, y los resultados obtenidos están reflejados en la Tabla 4.16.

Tabla 4.16. Residuos atípicos de la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y longitud)

| Fila | Y | Y Predicha | Residuo | Resid. estudentizado |
|------|--------|------------|---------|----------------------|
| 7 | 12.412 | 14.895 | -2.4823 | -5.71 |
| 24 | 16.861 | 18.185 | -1.3242 | -2.56 |

Para el estudio de residuos atípicos, el programa Statgraphics, construye una lista con todas las observaciones con residuos estudentizados con valor absoluto mayor a dos. Los residuos estudentizados miden la proporción entre la desviación de cada valor observado de irradiación global media 7 años del modelo ajustado y la desviación estándar, utilizando todos los datos excepto esa observación. Se puede ver en la Tabla 4.16 que hay dos residuos estudentizados mayores que dos, de los que uno es mayor que tres, que corresponde a la estación BU101. Se suele estudiar detenidamente las observaciones con residuos mayores a tres para determinar si son valores aberrantes que debieran ser eliminados del modelo y tratados por separado.

Se ha realizado un tercer estudio para ajustar linealmente la irradiación global media cambiando las variables independientes a latitud y altitud. El modelo propuesto es:

$$\text{Irradiación Media 7 años} = \alpha_1 + \alpha_2 * \text{latitud} + \alpha_3 * \text{altitud}$$

Los resultados obtenidos se recogen en las Tabla 4.17, Tabla 4.18 y Tabla 4.19

Tabla 4.17. Regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud)

| Parámetro | Estimación | Error Estándar | Estadístico T | P-Valor |
|------------|------------|----------------------|---------------|---------|
| α_1 | 53.116 | 8.7342 | 6.0814 | 0.0000 |
| α_2 | -0.8996 | 0.2065 | -4.3570 | 0.0001 |
| α_3 | 0.0012 | 9.4×10^{-4} | 1.2826 | 0.2068 |

A partir de la Tabla 4.17 se obtienen los parámetros para ajustar el modelo propuesto para describir la relación entre la irradiación media 7 años y las dos variables independientes. La ecuación del modelo ajustado es:

$$\text{Irradiación Media 7 años} = 53,116 - 0,8996 * \text{latitud} + 0,0012 * \text{altitud}$$

Al observar que el “P-Valor” más alto de las variables independientes, que corresponde a altitud (0.2068) es mayor que 0.05, el término correspondiente no era estadísticamente significativo con un nivel de confianza del 95 % ó mayor, por lo que en un modelo en el que las variables independientes fueran únicamente la latitud y la altitud, se podría eliminar esta última, pero eso nos llevaría al caso ya estudiado en primer lugar

Tabla 4.18. Análisis de Varianza en la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud)

| <i>Fuente</i> | <i>Suma de Cuadrados</i> | <i>Gl</i> | <i>Cuadrado Medio</i> | <i>Razón-F</i> | <i>P-Valor</i> |
|---------------|--------------------------|-----------|-----------------------|----------------|----------------|
| Modelo | 11.251 | 2 | 5.6254 | 10.67 | 0.0002 |
| Residuo | 21.608 | 41 | 0.5270 | | |
| Total (Corr.) | 32.858 | 43 | | | |

Puesto que el “P-Valor” en la tabla ANOVA es menor que 0.05, existe una relación estadísticamente significativa entre las variables con un nivel de confianza del 95 %.

Tabla 4.19. Estadísticos en la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud)

| | |
|---|--|
| $R^2 = 34.24 \%$ | $R^2(\text{ajustado para g.l.}) = 31.03 \%$ |
| <i>Error estándar del est.</i> = 0.7260 | <i>Error absoluto medio</i> = 0.5249 |
| <i>Estadístico Durbin – Wastson</i> = 1.2427 ($P = 0.0024$) | <i>Autocorrelación de residuos con retaso 1</i> = 0.3736 |

El valor de R^2 indica que el modelo así ajustado explica el 34.24 % de la variabilidad en la irradiación. La R^2 ajustada es 31.03 %. El error estándar del estimador muestra que la desviación estándar de los residuos es 0.7260. El error absoluto medio (MAE) de 0.5249. El estadístico de Durbin-Watson (DW) presenta un el P-Valor menor que 0.05, lo que significa una posible correlación serial con un nivel de confianza del 95 %. El modelo se ha representado en la Figura 4.6.

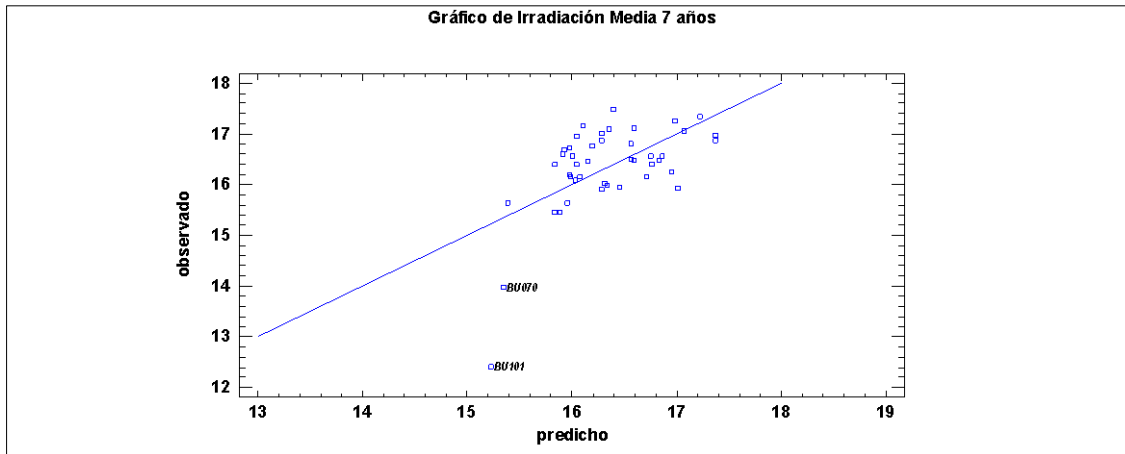


Figura 4.6. Regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud)

El análisis de los residuos atípicos, se muestra en la Tabla 4.20.

Tabla 4.20. Residuos atípicos en la regresión múltiple de la Irradiación Global Media 7 años con dos variables independientes (latitud y altitud)

| Fila | Y | Y Predicha | Residuo | Resid. estudentizado |
|------|--------|------------|---------|----------------------|
| 6 | 13.966 | 15.345 | -1.3785 | -2.18 |
| 7 | 12.412 | 15.231 | -2.8189 | -5.47 |

Este estudio permite establecer la posible existencia de valores aberrantes que debieran ser tratados por separado, y se ha comprobado la existencia de dos residuos estudentizados mayores que dos, de los que uno es mayor que tres. Ese valor se corresponde con la estación BU101, igual que en el caso anterior.

Por último se ha realizado el análisis de regresión múltiple con tres variables independientes, latitud, longitud y altitud, dejando como en los casos anteriores la irradiación media 7 años como variable dependiente.

$$Irradiación\ Media\ 7\ años = \alpha_1 + * latitud + \alpha_3 * longitud + \alpha_4 * altitud$$

Los resultados se muestran en tabla de coeficientes (Tabla 4.21), tabla de análisis de varianza (Tabla 4.22) y tabla de estadísticos (Tabla 4.23).

Tabla 4.21. Coeficientes de la regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud, longitud y altitud)

| Parámetro | Estimación | Error Estándar | Estadístico T | P-Valor |
|------------|-----------------------|-------------------------|---------------|---------|
| α_1 | 44.8062 | 6.0199 | 7.4431 | 0.0000 |
| α_2 | -0.8150 | 0.1401 | -5.8186 | 0.0000 |
| α_3 | -0.5733 | 0.0813 | -7.0532 | 0.0000 |
| α_4 | 3.81×10^{-3} | 7.3488×10^{-4} | 5.1843 | 0.0000 |

La relación entre la irradiación y las tres variables independientes (latitud, longitud y altitud) se puede expresar con la ecuación del modelo propuesto:

$$\text{Irradiación Media 7 años} = 44.8062 - 0.8150 * \text{latitud} - 0.5733 * \text{longitud} + 0.0038 * \text{altitud}$$

Un aspecto curioso en este caso es que todos los “P-Valores” de las variables independientes son “0” por lo que, al ser menores de 0.05, todos los términos son estadísticamente significativos (con un nivel de confianza del 95 %).

Tabla 4.22. Análisis de Varianza de la regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud, longitud y altitud)

| Fuente | Suma de Cuadrados | Gl | Cuadrado Medio | Razón-F | P-Valor |
|---------------|-------------------|----|----------------|---------|---------|
| Modelo | 23.228 | 3 | 7.7427 | 32.16 | 0.0000 |
| Residuo | 9.631 | 40 | 0.2408 | | |
| Total (Corr.) | 32.86 | 43 | | | |

Tabla 4.23. Estadísticos de la regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud, longitud y altitud)

| | |
|--|---|
| $R^2 = 70.69 \%$ | $R^2(\text{ajustado para g.l.}) = 68.49 \%$ |
| Error estándar del est. = 0.49068 | Error absoluto medio = 0.34881 |
| Estadístico Durbin – Wastson = 1.5259 (P = 0.0285) | Autocorrelación de residuos con retaso 1 = 0.2089 |

Puesto que el “P-Valor” en la tabla ANOVA es menor que 0.05, existe una relación estadísticamente significativa entre las variables con un nivel de confianza del 95 %. En este caso el R^2 indica que el modelo así ajustado explica el 70.69 % de la variabilidad en la irradiación. El estadístico R^2 ajustada, es 68.49 %. El error estándar del estimador muestra que la

desviación estándar de los residuos es 0.4907. El error absoluto medio (MAE) de 0.34881. El estadístico de Durbin-Watson (DW) al obtener un “P-Valor” menor que 0.05, supone una posible correlación serial con un nivel de confianza del 95 %.

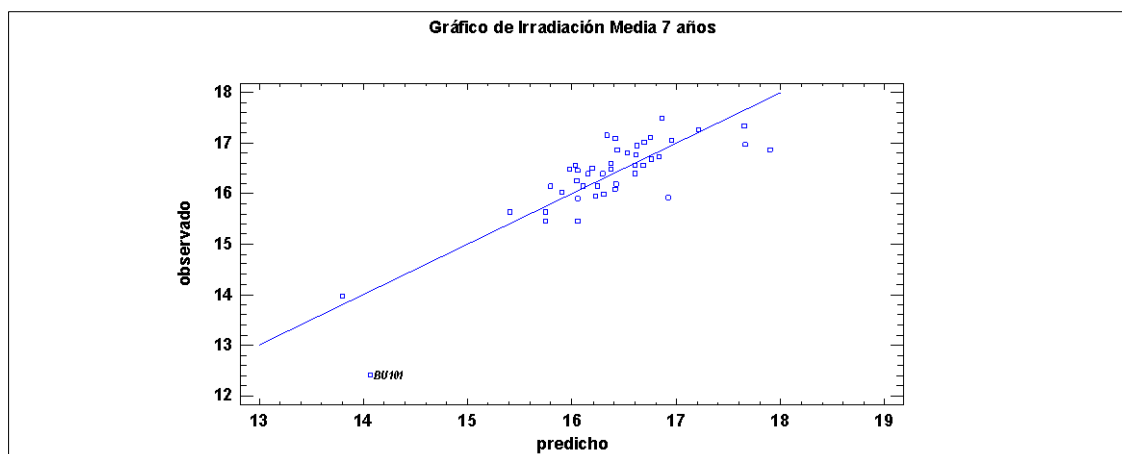


Figura 4.7. Regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud longitud y altitud)

Se han estudiado los residuos atípicos para determinar la existencia de valores aberrantes que debieran ser tratados por separado cuyos resultados se recogen en la Tabla 4.24. Al analizar los valores se ha comprobado que hay tres residuos estudentizados mayores que dos, de los cuales uno es mayor que tres, que corresponde a la estación BU101.

Tabla 4.24. Residuos atípicos de la regresión múltiple de la Irradiación Global Media 7 años con tres variables independientes (latitud, longitud y altitud)

| Fila | Y | Y Predicha | Residuo | Res. estudentizado |
|------|--------|------------|---------|--------------------|
| 7 | 12.412 | 14.067 | -1.6548 | -4.90 |
| 24 | 16.861 | 17.896 | -1.0351 | -2.59 |
| 26 | 15.921 | 16.927 | -1.0061 | -2.22 |

Este mismo estudio de regresión lineal múltiple de la irradiación global media se ha realizado para cada año de los siete del estudio. La ecuación que relaciona los cuatro parámetros es del tipo:

$$Irradiación\ año\ i = \alpha_{i1} + \alpha_{i2} * latitud + \alpha_{i3} * longitud + \alpha_{i4} * altitud$$

Los coeficientes de las siete ecuaciones se recogen en la Tabla 4.25

Tabla 4.25. Coeficientes de la regresión múltiple de la Irradiación Global Media 1 año con tres variables independientes (latitud, longitud y altitud) para los años 2007 a 2013

| | α_{i1} | α_{i2} | α_{i3} | α_{i4} |
|-------|---------------|---------------|---------------|-----------------------|
| 2007 | 47.2945 | - 0.87478 | - 0.57901 | 4.07×10^{-3} |
| 2008 | 46.1112 | - 0.87142 | - 0.69910 | 4.37×10^{-3} |
| 2009 | 42.5873 | - 0.75687 | - 0.64221 | 4.35×10^{-3} |
| 2010 | 32.1826 | - 0.51389 | - 0.58780 | 3.11×10^{-3} |
| 2011 | 49.8536 | - 0.93421 | - 0.51699 | 4.19×10^{-3} |
| 2012 | 44.1955 | - 0.80308 | - 0.48693 | 4.54×10^{-3} |
| 2013 | 41.4236 | - 0.74290 | - 0.56528 | 3.32×10^{-3} |
| Media | 44.8062 | -0.81496 | -0.57333 | 3.81×10^{-3} |

Los coeficientes α_{i2} , α_{i3} , y α_{i4} miden el cambio que sufre la irradiación media en el año i cuando, manteniendo constantes dos de las tres coordenadas geográficas, aumenta o disminuye en una unidad la tercera.

Se observa que todos los valores de α_{i2} , y α_{i3} son negativos lo que significa que a medida que aumenta la latitud o la longitud, la irradiación disminuye. Al estudiar una región situada al oeste del meridiano de Greenwich, las longitudes se han considerado negativas y aumentando de oeste a este, por lo que la irradiación aumenta de este a oeste. Al ser α_{i4} positiva, al aumentar la altitud también aumenta la irradiación.

El distinto orden de magnitud de α_{i2} , y α_{i3} con respecto a α_{i4} tiene sentido puesto que los valores de la latitud varían entre 40.5° y 42.9° y los de la longitud entre -6.72° y -2.10° , mientras que la altitud varía entre 467 y 1063 metros.

La Tabla 4.26, recoge el resumen del estudio realizado con el número de residuos atípicos y la relevancia de las variables independientes dada por el correspondiente “P-Valor”.

Tabla 4.26. Tabla resumen de los valores representativos de la relación entre la irradiación y las tres variables independientes, latitud, longitud y altitud

| | R^2 | Mayor P-valor | Nº de Residuos atípicos | R. estudentizado > 3 |
|-------|---------|----------------|-------------------------|----------------------|
| 2007 | 56.13 % | altitud 0.0016 | 3 | 1 |
| 2008 | 64.89 % | altitud 0.0001 | 3 | 1 |
| 2009 | 64.25 % | latitud 0.0002 | 3 | 1 |
| 2010 | 63.57 % | latitud 0.0010 | 1 | 1 |
| 2011 | 72.80 % | 0 | 3 | 1 |
| 2012 | 61.71 % | 0 | 2 | 1 |
| 2013 | 48.50 % | altitud 0.0047 | 4 | 1 |
| Media | 70.69 % | 0 | 3 | 1 |

4.4 ESTUDIO DE LA NORMALIDAD DE LOS DATOS

Una vez analizados los resultados obtenidos en los estudios anteriores se ha decidido comprobar si los datos de irradiación global media de 7 años seguían una distribución normal, y para ello se prescinde de las estaciones BU070 y BU101 que son las que presentan valores anómalos y que geográficamente tienen características diferentes a las del resto de las estaciones, con la particularidad geográfica que ambas estaciones no pertenecen a la cuenca del Duero.

Se trabaja, por lo tanto, con 42 estaciones y se ha realizado el estudio con ayuda de Statgraphics, como en el caso anterior. En la Tabla 4.27 se recogen los valores estadísticos correspondientes a la media de los siete años.

Tabla 4.27. Resumen Estadístico para Irradiación Media 7 años

| | |
|--------------------------|----------------------------|
| Recuento | 42 |
| Promedio | 16.4871 MJ·m ⁻² |
| Desviación Estándar | 0.5153 |
| Coeficiente de Variación | 3.13 % |
| Mínimo | 15.4434 |
| Máximo | 17.4742 |
| Rango | 2.0308 |
| Sesgo Estandarizado | -0.4364 |
| Curtosis Estandarizada | -0.7968 |

De particular interés en esta tabla son los valores del sesgo y la curtosis estandarizada, que al estar comprendidos entre -2 y 2 se encuentran en el rango esperado para datos provenientes de una distribución normal. De no haber sido sí, hubiera supuesto desviaciones significativas de la normalidad, lo que tendería a invalidar cualquier prueba estadística con referencia a la desviación estándar.

Si se analizan los resultados por medio de la gráfica de “Caja y Bigotes”, Figura 4.8, se observa que no aparece ningún valor atípico en oposición a lo que ocurría en el caso de trabajar con las 44 estaciones.

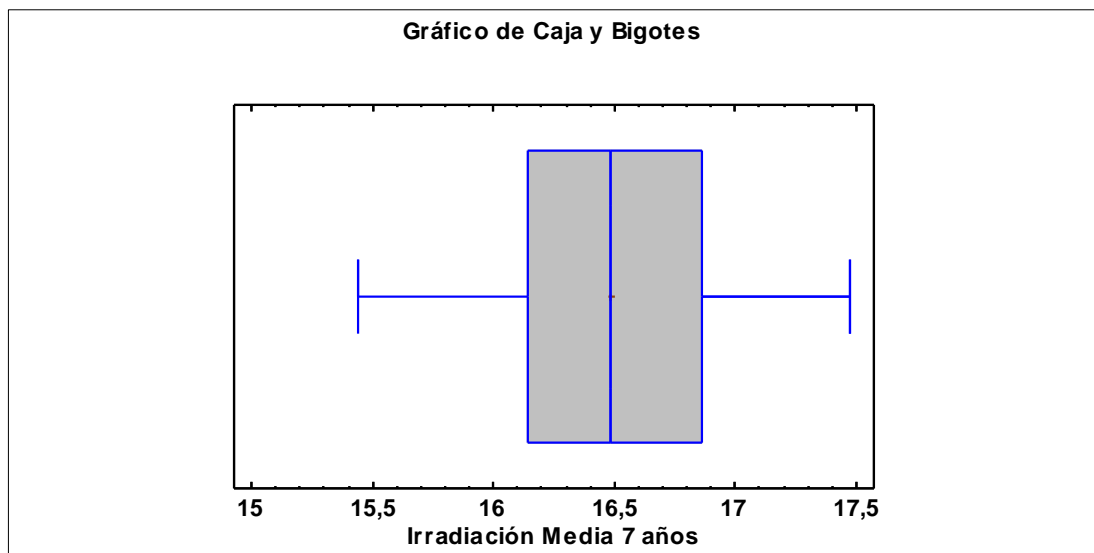


Figura 4.8. Gráfica de Caja y Bigotes para 42 estaciones ($\text{MJ}\cdot\text{m}^{-2}$)

Se han calculado los intervalos de confianza para la media y la desviación estándar, obteniéndose los siguientes resultados:

* Intervalo de confianza del 95 % para la media (16.4871 ± 0.1606) $\text{MJ}\cdot\text{m}^{-2}$

* Intervalo de confianza del 95 % para la desviación estándar [0.4240, 0.6571] $\text{MJ}\cdot\text{m}^{-2}$

Estos resultados son fiables en el caso de que los datos provengan de una distribución normal, y no lo son en el caso de la desviación estándar, si los datos no se ajustan a una distribución normal. Para asegurar que se estaba en el caso de una distribución normal se ha representado el gráfico de

probabilidad normal (Figura 4.9) y se ha comprobado que los datos se ajustaban correctamente.

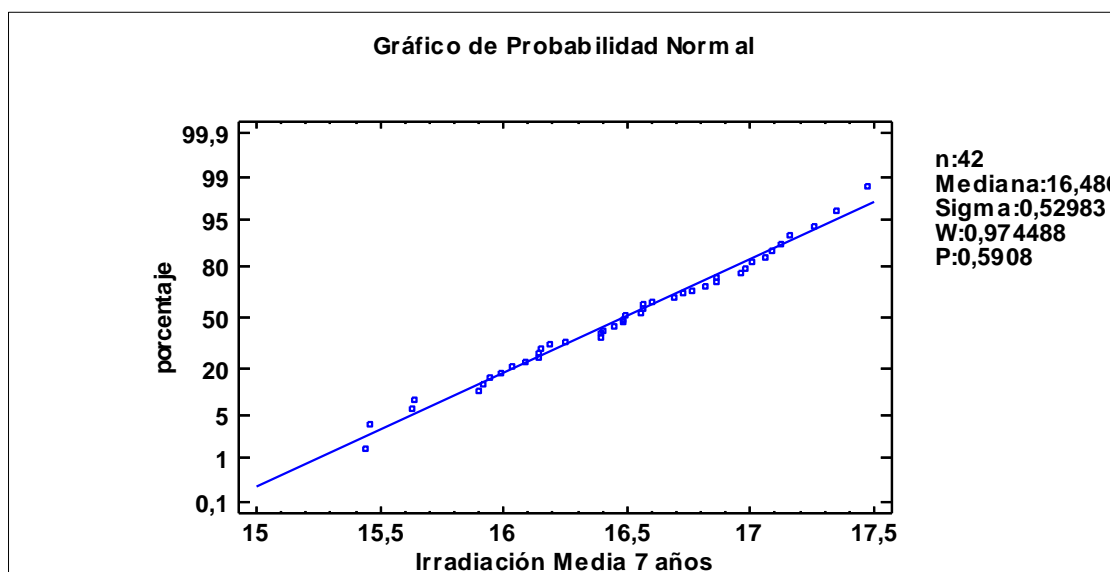


Figura 4.9. Gráfico de Probabilidad Normal

Los resultados de la prueba de Kolmogorov-Smirnov, se recogen en la Tabla 4.28, para la bondad de ajuste con los siguientes resultados: se ajustan 42 valores de “Datos No Censurados” con rango desde 15.4434 MJ·m⁻² a 17.4742 MJ·m⁻². Los datos presentan una media de 16.4871 MJ·m⁻² y una desviación estándar de 0.5153 MJ·m⁻².

Tabla 4.28. Resultados de la Prueba de Kolmogorov-Smirnov

| | <i>Normal</i> |
|---------|---------------|
| DMAS | 0.05418 |
| DMENOS | 0.06880 |
| DN | 0.06880 |
| P-Valor | 0.98866 |

Debido a que el “P-Valor” más pequeño de las pruebas realizadas es mayor o igual a 0.05, no se puede rechazar la idea de que la irradiación media de 7 años proviene de una distribución normal con 95 % de confianza. La representación de la Traza de Densidad en la Figura 4.10 y en la Figura 4.11 el histograma de los valores, confirman esta hipótesis

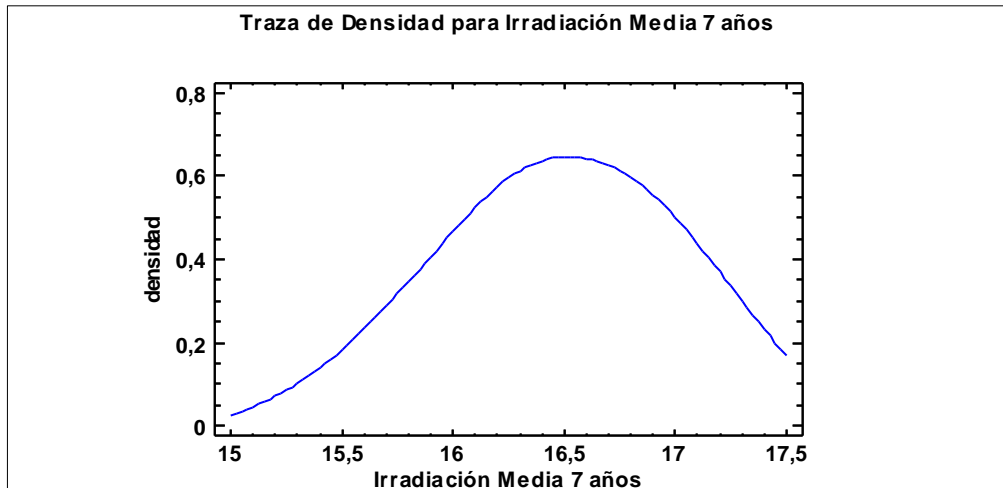


Figura 4.10. Traza de Densidad de los valores de Irradiación Global Media de 7 años con las 42 estaciones que garantizan una distribución normal

A la vista de los resultados se ha considerado que los datos de irradiación media anual de las 42 estaciones siguen una distribución normal, lo que se ha tenido en cuenta para elegir el número de puntos para realizar la validación de los resultados.

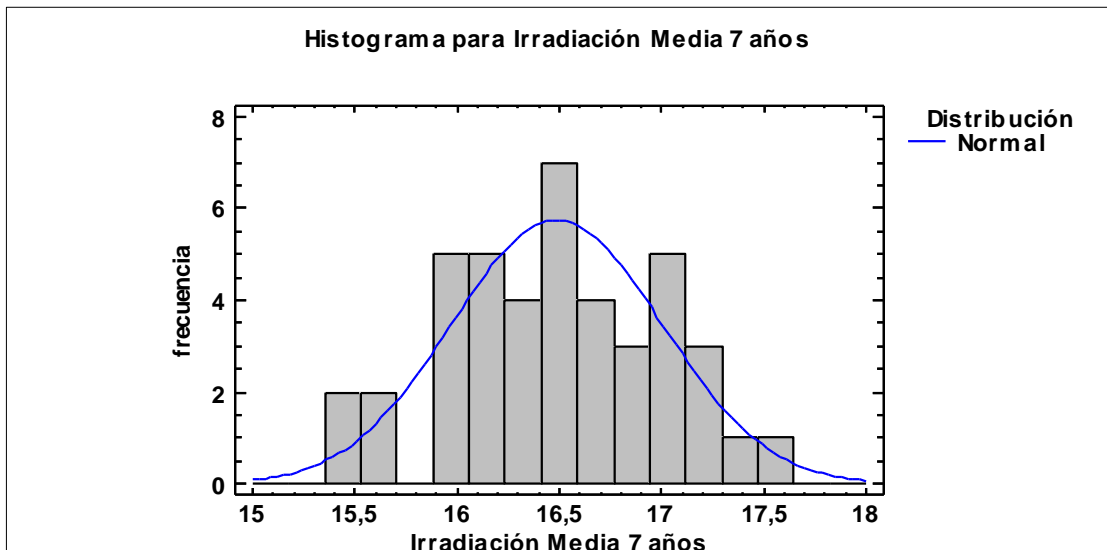


Figura 4.11. Histograma de los valores de Irradiación Global Media de 7 años con las 42 estaciones que garantizan una distribución normal

Una vez aceptada la hipótesis de que los valores de irradiación global media de las 42 estaciones siguen una distribución normal, se ha realizado una vez más un análisis de regresión múltiple, con tres variables independientes (latitud, longitud y altitud) para comprobar si los resultados se ajustan mejor, al no intervenir las estaciones “anómalas”, y se ha comprobado que el R^2 es de

49.39 % y el estadístico R^2 ajustado es 45.39 %. Los resultados no son mucho mejores que los conseguidos con las 44 estaciones originales, pero el error absoluto medio (MAE) pasa de 0.3488 a un 0.2924, corroborando en este caso las tesis de Willmott (*Willmott y Matsuura, (2005), Willmott et al., (2009)*) de que MAE puede ser más apropiado para evaluar los errores que otros estimadores.

4.5 ESTUDIO DE LA VALIDACIÓN DE RESULTADOS

Para elegir el mejor método de interpolación hay que validar los resultados obtenidos con cada uno de ellos. Para ellos se establece una serie de estaciones para realizar los cuatro tipos de interpolación escogidos y se reserva el resto como estaciones de validación. En número de estaciones en cada serie (interpolación y control) se toma considerando que en el estudio estadístico anterior, salvo las dos estaciones correspondientes a la Cuenca del Ebro, el resto de estaciones, 42, siguen una distribución normal. El mínimo número de puntos necesarios para realizar la interpolación con un determinado índice de confianza, en este caso un 95 %, y con un error admisible menor del 5 % se ha establecido utilizando la fórmula de (*Hines y Montgomery, (1990)*):

$$n = \frac{z_{\alpha}^2 \sigma^2}{d^2}$$

Esta relación, cuando se conoce el número de estaciones (población), se transforma en:

$$n = \frac{N z_{\alpha}^2 \sigma^2}{d^2 (N - 1) + z_{\alpha}^2 \sigma^2}$$

expresión en la que N es el número total de datos disponible, d es la cota de error que se admite, z_{α} el valor asociado al 95 % de confianza, σ^2 es la varianza de la irradiación global media de 7 años para las 42 estaciones.

Las dos estaciones del noreste de Burgos (BU070 y BU101), se incluyen en la base para la interpolación, pero no se consideran como puntos de control, ya que el resultado de la interpolación puede ser distorsionado por los valores de

esas estaciones. Por lo tanto, para el cálculo del número de estaciones necesarias para hacer la interpolación, se ha tomado $N=42$ estaciones, $d=0.05$ (ya que se admite un error del 5 %) $z_{\alpha} = 1.96$ y $\sigma^2 = 0.259$.

Con los cálculos anteriores se ha obtenido un valor de $n=38.07$, por lo que se ha decidido tomar para el primer control 38 estaciones de las 42 consideradas, a las que se deben añadir para la interpolación las dos del noreste de Burgos y dejar cuatro estaciones de las 42 para el control. Como se ha comentado en capítulos previos, se ha completado el conjunto de puntos de control con una estación que no pertenece a la red SIAR de Castilla y León, pero está situada en el noreste de Burgos, muy cerca de las dos estaciones antes mencionadas, para ver el comportamiento de la interpolación en zonas no homogéneas en sus características geográficas y de irradiación. Por lo tanto en ese primer control el número de puntos reservados para la validación ha sido de cinco.

En los siguientes controles no se incluyen en los conjuntos de puntos de validación ninguna estación fuera de Castilla y León ni las estaciones BU 070 y BU101. Se ha estudiado el posible aumento de error cometido al tomar 37 estaciones, de 42 posibles, más las dos anteriormente citadas como puntos de interpolación, y dejando cinco puntos de validación. En este caso, se ha obtenido que el error cometido sería del 5.7 %. Al repetir los cálculos considerando la situación más desfavorable (las 44 estaciones con igual consideración con lo que la varianza de la irradiación media aumenta tomando σ^2 un valor de 0.764), el error sube hasta un 9.4 %, que se ha considerado asumible y se ha decidido utilizar 39 puntos para la interpolación y dejar cinco puntos para realizar la validación.

A esta forma de realizar el control de validez se la denomina "Split-Sample validation" (Demirhan et al., (2013), Smith et al., (2005)). Otros autores (Apaydin et al., (2004), Güler, (2014), Price et al., (2000)) la incluyen como una validación cruzada, que en su forma más común se realiza haciendo la interpolación usando $n-1$ puntos de los n posibles y viendo el error que se comete en la estimación sobre el punto n . Se repite el proceso con todos los puntos y se comparan los resultados.

Una vez fijados el número de puntos necesarios para realizar la interpolación, se ha procedido a preparar los datos para emplear los métodos de interpolación seleccionados tomando como puntos base de interpolación los de 40 ó 39 estaciones de la red SIAR y 26 ó 27 de las estaciones complementarias manteniendo el número de estaciones para la validación en cinco. El proceso es similar a los anteriores: hay que establecer dos nuevas bases de datos: una con las estaciones empleadas para la interpolación y la otra con las estaciones utilizadas para el control. Dada la distribución geográfica de las estaciones de la red SIAR, se ha decidido realizar diferentes controles para estudiar la influencia de la ubicación y proximidad de las mismas. En total se han realizado cinco controles diferentes. En cada uno de los controles se ha calculado el error cometido con cada modelo de interpolación y para cada estación.

Posteriormente se ha realizado otra validación, utilizando datos que no se han usado en la interpolación y que se habían obtenido de forma independiente, para estudiar el error cometido en la interpolación. En este caso se han tomado como base de interpolación las 44 estaciones de la red SIAR y las 27 complementarias y el control se ha realizado con los datos de cuatro estaciones de la AEMET. Esta es una variación del método Jak-knifing (*Erdogan, (2009), Robinson y Metternicht, (2005)*) ampliamente utilizado para la validación de resultados, considerando que un número suficiente de puntos en el conjunto independiente de validación minimiza el riesgo de sesgo en el resultado. Por último se ha realizado un control complementario.

Willmott y Matsuura, (2005) y otros, otorgan mayor fiabilidad al indicador estadístico MAE frente a otros parámetros para el estudio del error, como pudiera ser el RMSE. MAE es la medida más natural del error promedio, ya que es una medida inequívoca mientras que RMSE está influida por otros errores. El estudio estadístico de los resultados se basa en los cuatro parámetros estadísticos clásicos: RMSE, MBE, MAE y MAPE, todos ellos en %.

CAPÍTULO 5.

RESULTADOS

5.1 INTRODUCCIÓN

En este capítulo se exponen todos los resultados obtenidos en el desarrollo de este trabajo que confirman la consecución de los objetivos planteados al inicio. El capítulo se ha dividido en dos partes. La primera parte se dedica a la obtención de los mapas de irradiación global horizontal media anual diaria en Castilla y León a partir de métodos de interpolación con la base de datos construida y validada en capítulos anteriores. La selección del método de interpolación óptimo se ha realizado a partir de los indicadores de calidad estadísticos habituales y previamente definidos, garantizando la fiabilidad de los mapas de radiación en la región objeto de estudio.

La segunda parte está dedicada a la obtención de irradiación global sobre superficie vertical en las cuatro orientaciones cardinales, norte, sur, este y oeste. Para este trabajo se ha recurrido a los modelos matemáticos de irradiación sobre superficie inclinada a partir de valores medidos de irradiación global, directa y difusa sobre superficie horizontal. La validación de los modelos propuestos, aplicados a los datos de irradiación horizontal conseguidos en el procedimiento anterior, se ha realizado con respecto a los datos experimentales obtenidos en la instalación de la Escuela Politécnica Superior de Burgos. De esta forma, se ha conseguido la caracterización del recurso solar sobre superficies verticales en las cuatro direcciones cardinales en cualquier punto de la comunidad de Castilla y León y la posibilidad de realizar estimaciones más exactas del potencial energético sobre fachadas, de gran importancia para la eficiencia energética en la edificación.

5.2 ANÁLISIS DE LOS MÉTODOS DE INTERPOLACIÓN

En el capítulo 3.2 se ha hecho una descripción del sistema de validación realizado sobre los diferentes métodos de interpolación utilizados, identificando en cada caso las estaciones utilizadas en la base de interpolación y las empleadas para la validación. En este capítulo se concretan qué estaciones forman parte de cada grupo en cada una de las pruebas realizadas, que llamaremos control. Los datos de irradiación de las estaciones de la red SIAR de la región siguen una distribución normal (cuando se prescinde de las estaciones BU070 y BU101), como se ha comprobado en el capítulo 4. Los cinco primeros controles descritos emplean únicamente estaciones de la red SIAR de la comunidad y estaciones fronterizas, tanto en la base de interpolación como en la validación. Se ha realizado un sexto control con los mismos métodos de interpolación que en los casos anteriores, utilizando como estaciones para la comprobación las de la red AEMET disponibles en la región y por último un análisis empleando como control estaciones de la red SIAR y las de la red AEMET. Estos controles intentan garantizar que la bondad del método de interpolación escogido para generar los mapas de irradiación no depende de las estaciones elegidas como control. La calidad de cada una de las interpolaciones en cada control se ha establecido según los parámetros estadísticos habituales (RMSE, MBE, MAE y MAPE (en %)). Los resultados (Control 1,..., Control 7) se presentan, para cada uno de los controles utilizados de la siguiente forma:

- a) Se indican las estaciones utilizadas como control.
- b) Se muestran los parámetros estadísticos RMSE, MBE, MAE y MAPE (%) calculados mediante la comparación de los datos experimentales (medidos en las estaciones de validación) con los datos calculados a través de la interpolación, para cada uno de los métodos de interpolación utilizados (29 en total).
- c) Se muestran los mapas de irradiación global sobre plano horizontal obtenidos con el método de interpolación que ha presentado el menor error.

5.2.1 CONTROL 1 (C1)

En la Tabla 5.1 se recogen las coordenadas geográficas (longitud, latitud y altitud) y el valor medio anual diario de irradiación global horizontal (GHI) de las estaciones que forman la base de control.

Tabla 5.1. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C1

| Estación | latitud ($^{\circ}$) | longitud ($^{\circ}$) | altitud (m s.n.m.) | GHI ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) |
|----------|------------------------|-------------------------|--------------------|---|
| BI030 | 43.163 | -3.28 | 300 | 11.7667 |
| LE020 | 42.506 | -5.43 | 791 | 16.3909 |
| LE060 | 42.460 | -5.90 | 835 | 16.6910 |
| PA080 | 42.344 | -4.28 | 798 | 16.1487 |
| VA070 | 41.643 | -4.27 | 756 | 16.4895 |

Como se observa en la Tabla 5.1, la estación BI030, perteneciente a la red fronteriza del País Vasco, pero situada en la provincia de Burgos, tiene un valor medio anual diario de GHI muy distinto del resto de estaciones incluidas en esta base de control. Se ha incluido esta estación en el C1 para analizar la influencia de la zona limítrofe, y comparar los resultados obtenidos en ese punto con los datos de la estación de la red SIAR BU101, muy próxima a ésta y, que como se ha indicado en los capítulos 3 y 4, muestra un comportamiento diferente al resto de las estaciones de Castilla y León, hecho atribuido a sus características geográficas (no pertenece a la Cuenca del Duero).

La Tabla 5.2 recoge todos los resultados obtenidos en el C1 con los 29 métodos de interpolación empleados. Todos los errores muestran resultados semejantes, dando valores pequeños. El método que presenta menor valor del error cuadrático medio (RMSE) es el método Kriging Ordinario Lineal con 35 puntos (KOL35), con un valor de 2.11 %. Por el contrario el mayor valor RMSE (18.48 %), lo muestra el método Kriging Ordinario Gaussiano con 35 puntos (KOG35). De todos los métodos de interpolación utilizados, 21 de ellos (un 72 %) tiene valores RMSE inferiores al 5 % y de ellos nueve inferiores al 2 %. Las diferencias entre los métodos de interpolación son mínimas, salvo el método Kriging Gaussiano de 12 y 35 puntos.

Con respecto al error medio, MBE, todos los modelos sobreestiman el valor de la irradiación, excepto el método IDW (12 y 35 puntos) que subestima este valor dando valores MBE negativos.

Tabla 5.2. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C1(BI030, LE020, LE060, PA080 y VA070), con todos los métodos de interpolación

| CONTROL 1 | | | | | | | |
|-------------------|--------------|-------------|----------|--------|--------|---------|------|
| | | | RMSE (%) | MBE(%) | MAE(%) | MAPE(%) | |
| KRIGING 12 PUNTOS | ORDINARIO | Esférico | 2.45 | 1.64 | 2.05 | 1.96 | |
| | | Circular | 2.45 | 1.66 | 2.07 | 1.98 | |
| | | Exponencial | 2.45 | 1.36 | 2.19 | 2.14 | |
| | | Gaussiano | 7.54 | 5.61 | 6.35 | 6.96 | |
| | | Lineal | 2.43 | 1.70 | 2.10 | 2.02 | |
| | UNIVERSAL | Lineal | 3.84 | 3.07 | 3.18 | 3.25 | |
| | | Cuadrático | 3.56 | 2.50 | 3.05 | 3.24 | |
| KRIGING 35 PUNTOS | ORDINARIO | Esférico | 2.13 | 1.74 | 1.77 | 1.67 | |
| | | Circular | 2.16 | 1.63 | 1.92 | 1.86 | |
| | | Exponencial | 2.32 | 1.56 | 2.12 | 2.10 | |
| | | Gaussiano | 18.49 | 8.62 | 9.71 | 12.24 | |
| | | Lineal | 2.12 | 2.10 | 2.10 | 2.14 | |
| | UNIVERSAL | Lineal | 6.62 | 4.75 | 5.97 | 5.86 | |
| | | Cuadrático | 3.65 | 2.66 | 3.36 | 3.60 | |
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 5.00 | 4.08 | 4.08 | 4.62 |
| | | | 0.01 | 4.28 | 3.50 | 3.49 | 3.79 |
| | | | 0.001 | 4.28 | 3.49 | 3.49 | 3.79 |
| | TENSIÓN | PESO | 1 | 3.19 | 2.43 | 2.43 | 2.48 |
| | | | 5 | 3.07 | 2.24 | 2.24 | 2.23 |
| | | | 10 | 2.99 | 2.09 | 2.16 | 2.14 |
| SPLIN 35 PUNTOS | REGULARIZADO | PESO | 0.1 | 6.03 | 4.81 | 4.81 | 5.74 |
| | | | 0.01 | 5.09 | 4.16 | 4.16 | 4.77 |
| | | | 0.001 | 5.09 | 4.16 | 4.16 | 4.77 |
| | TENSIÓN | PESO | 1 | 3.42 | 2.82 | 2.83 | 3.07 |
| | | | 5 | 5.17 | 3.70 | 3.83 | 4.71 |
| | | | 10 | 3.11 | 2.55 | 2.55 | 2.67 |
| IDW | 12 PUNTOS | | 3.24 | -0.43 | 2.59 | 2.90 | |
| | 35 PUNTOS | | 4.33 | -0.54 | 3.17 | 3.65 | |
| VECINO PROXIMO | | 12 PUNTOS | | 3.17 | 2.96 | 2.96 | 3.05 |

La Figura 5.1, representa el modelo que tiene menor valor RMSE (KOL35) y el que de menor valor MAE (KOEf35).

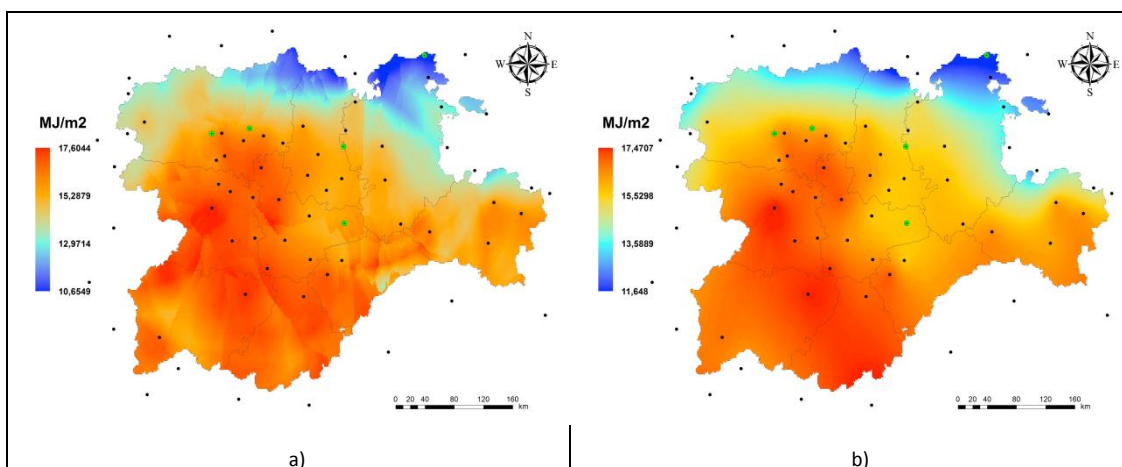


Figura 5.1. Mapas de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenidos para el control C1 con los metodos de interpolación: a) KOL35 y b) KOEf35

5.2.2 CONTROL 2 (C2)

En la tabla 5.3 se muestran las coordenadas geográficas (longitud, latitud y altitud) y el valor medio anual diario de irradiación global (GHI) de las estaciones que forman el control C2. Se repiten dos estaciones con respecto al control C1 (LE020 y PA080). La Tabla 5.4 recoge todos los resultados obtenidos en el C2 con los 29 métodos de interpolación empleados. En general se observa una disminución de los valores de los errores en comparación con el control C1.

Tabla 5.3. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C2

| Estación | latitud ($^{\circ}$) | longitud ($^{\circ}$) | altitud (m s.n.m.) | GHI ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}$) |
|----------|------------------------|-------------------------|--------------------|--|
| LE020 | 42.506 | -5.43 | 791 | 16.3909 |
| PA080 | 42.344 | -4.28 | 798 | 16.1487 |
| SO020 | 41.566 | -3.22 | 855 | 16.5615 |
| VA060 | 41.487 | -5.00 | 658 | 16.4831 |
| ZA060 | 41.934 | -5.67 | 738 | 17.0064 |

Un 93 % de los casos tienen un valor RMSE inferior al 4 %. El método KOG35 también es el de mayor desviación, aunque se mejora la interpolación con respecto al control C1. El método KOL35 repite como el

más aproximado pero con un valor de RMSE ligeramente inferior al calculado en el control C1. El método IDW sobreestima los valores de irradiación en este control (en el control C1 lo subestimaba) y el método SR(0.1)35 es el único que subestima la irradiación al interpolar.

Tabla 5.4. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C2 (LE020, PA080, SO020, VA060 y ZA060), con todos los métodos de interpolación

| CONTROL 2 | | | | | | | |
|-------------------|--------------|-------------|----------|--------|--------|---------|-------|
| | | | RMSE (%) | MBE(%) | MAE(%) | MAPE(%) | |
| KRIGING 12 PUNTOS | ORDINARIO | Esférico | 2.55 | 1.03 | 2.03 | 2.04 | |
| | | Circular | 2.55 | 1.04 | 2.04 | 2.04 | |
| | | Exponencial | 2.55 | 1.02 | 2.05 | 2.06 | |
| | | Gaussiano | 2.21 | 0.61 | 1.79 | 1.8 | |
| | | Lineal | 2.53 | 1.03 | 2.03 | 2.04 | |
| | UNIVERSAL | Lineal | 2.49 | 1.10 | 2.01 | 2.03 | |
| | | Cuadrático | 3.37 | 1.34 | 2.27 | 2.27 | |
| KRIGING 35 PUNTOS | ORDINARIO | Esférico | 2.20 | 1.11 | 1.74 | 1.75 | |
| | | Circular | 2.10 | 1.07 | 1.70 | 1.71 | |
| | | Exponencial | 2.36 | 1.21 | 1.21 | 1.86 | |
| | | Gaussiano | 14.85 | 8.88 | 8.95 | 8.96 | |
| | | Lineal | 1.88 | 1.14 | 1.74 | 1.75 | |
| | UNIVERSAL | Lineal | 5.33 | 4.16 | 4.95 | 4.97 | |
| | | Cuadrático | 2.53 | 1.04 | 1.06 | 2.35 | |
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.77 | 1.67 | 2.28 | 2.30 |
| | | | 0.01 | 3.70 | 1.66 | 2.17 | 2.19 |
| | | | 0.001 | 3.49 | 1.56 | 2.01 | 2.02 |
| | TENSIÓN | PESO | 1 | 3.30 | 1.47 | 1.95 | 1.964 |
| | | | 5 | 3.23 | 1.44 | 1.94 | 1.95 |
| | | | 10 | 3.18 | 1.42 | 1.93 | 1.94 |
| SPLIN 35 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.93 | -1.76 | 2.09 | 2.11 |
| | | | 0.01 | 3.87 | 1.73 | 2.04 | 2.06 |
| | | | 0.001 | 3.63 | 1.62 | 1.93 | 1.94 |
| | TENSIÓN | PESO | 1 | 3.34 | 1.50 | 1.85 | 1.86 |
| | | | 5 | 3.28 | 1.47 | 1.83 | 1.83 |
| | | | 10 | 3.25 | 1.45 | 1.81 | 1.82 |
| IDW | 12 PUNTOS | | 2.87 | 1.12 | 2.42 | 2.46 | |
| | 35 PUNTOS | | 3.20 | 1.78 | 2.49 | 2.50 | |
| VECINO PROXIMO | 12 PUNTOS | | 2.51 | 2.95 | 2.15 | 2.16 | |

Willmott et al. (2005), asegura que el error absoluto medio (MAE) es mejor parámetro estadístico para validar un procedimiento. Si se analiza este parámetro, el método de mejor ajuste es el KOL35, con un valor MAE (1.73 %) ligeramente inferior al RMSE.

En la Figura 5.2 se representan los mapas de irradiación para el menor error cuadrático medio (RMSE), para este control, el método KOL35, y el obtenido por interpolación con el método que tiene el menor MAE, KUC35.

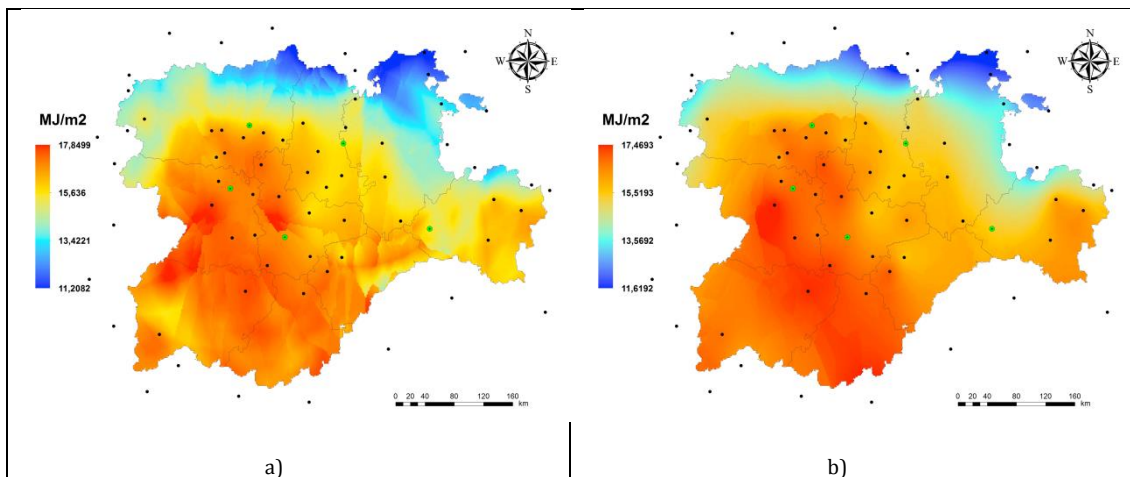


Figura 5.2. Mapas de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenidos para el control C2 con los métodos de interpolación: a) KOL35 y b) KUC35

5.2.3 Control 3 (C3)

Igual que en los controles anteriores se muestran las coordenadas geográficas y la irradiación de las estaciones de validación en la Tabla 5.5. En ella se aprecian escasas diferencias en los valores de GHI, y una mayor variabilidad en los parámetros geográficos, sobre todo la estación BU050.

Tabla 5.5. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C3

| Estación | latitud ($^{\circ}$) | longitud ($^{\circ}$) | altitud (m s.n.m.) | GHI ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) |
|----------|------------------------|-------------------------|--------------------|---|
| LE020 | 42.506 | -5.43 | 791 | 16.3909 |
| BU050 | 41.637 | -3.58 | 870 | 16.1454 |
| SA030 | 40.990 | -5.48 | 815 | 17.3493 |
| VA080 | 41.861 | -5.07 | 739 | 17.0884 |
| ZA080 | 41.508 | -5.37 | 650 | 16.8129 |

Los resultados estadísticos de los métodos de interpolación se recogen en la Tabla 5.6. En este caso los parámetros estadísticos para todos los métodos de interpolación, son, en general, menores que los obtenidos en los controles C1 y C2.

Tabla 5.6. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C3 (LE020, BU050, SA030, VA080 Yza080), con todos los métodos de interpolación

| CONTROL 3 | | | | | | | |
|-------------------|--------------|-------------|--------|---------|--------|-------|------|
| | | | RMSE | MBE | MAE | MAPE | |
| KRIGING 12 PUNTOS | ORDINARIO | Esférico | 1.65 | - 0.01 | 1.44 | 1.44 | |
| | | Circular | 1.64 | - 0.03 | 1.43 | 1.43 | |
| | | Exponencial | 1.60 | 0.17 | 1.42 | 1.41 | |
| | | Gaussiano | 1.45 | 0.88 | 1.20 | 1.20 | |
| | | Lineal | 1.62 | - 0.06 | 1.37 | 1.37 | |
| | UNIVERSAL | Lineal | 1.99 | 1.10 | 1.81 | 1.81 | |
| | | Cuadrático | 1.87 | - 0.29 | 1.63 | 1.64 | |
| KRIGING 35 PUNTOS | ORDINARIO | Esférico | 1.61 | 0.26 | 1.46 | 1.46 | |
| | | Circular | 1.69 | 0.19 | 1.53 | 1.53 | |
| | | Exponencial | 1.58 | 0.57 | 1.41 | 1.40 | |
| | | Gaussiano | 19.09 | - 10.91 | 13.18 | 13.31 | |
| | | Lineal | 2.406 | 0.45 | 2.20 | 2.21 | |
| | UNIVERSAL | Lineal | 3.72 | 3.07 | 3.07 | 3.07 | |
| | | Cuadrático | 1.26 | 0.16 | 1.07 | 1.08 | |
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 1.65 | - 0.12 | 1.39 | 1.38 |
| | | | 0.01 | 1.67 | - 0.06 | 1.36 | 1.36 |
| | | | 0.001 | 1.68 | - 0.05 | 1.34 | 1.35 |
| | TENSIÓN | PESO | 1 | 1.62 | 0.02 | 1.24 | 1.25 |
| | | | 5 | 1.58 | 0.11 | 1.19 | 1.20 |
| | | | 10 | 1.56 | 0.17 | 1.17 | 1.18 |
| SPLIN 35 PUNTOS | REGULARIZADO | PESO | 0.1 | 1.40 | 0.22 | 1.21 | 1.21 |
| | | | 0.01 | 1.43 | 0.21 | 1.12 | 1.13 |
| | | | 0.001 | 1.55 | 0.18 | 1.16 | 1.17 |
| | TENSIÓN | PESO | 1 | 1.59 | 0.17 | 1.20 | 1.22 |
| | | | 5 | 1.59 | 0.20 | 1.25 | 1.26 |
| | | | 10 | 1.59 | 0.23 | 1.27 | 1.29 |
| IDW | 12 PUNTOS | 1.86 | 0.73 | 1.65 | 1.64 | | |
| | 35 PUNTOS | 2.29 | 1.63 | 1.85 | 1.83 | | |
| VECINO PROXIMO | 12 PUNTOS | 1.75 | - 0.22 | 1.63 | 1.62 | | |

El 86 % de los errores es menor del 2 %. El método con menor error cuadrático medio es el KUC35 (1.25 %), que coincide con el de menor valor de MAE (1,07 %). El KOG35 sigue siendo el método que peor se ajusta a la zona en estudio, dando el mayor error cuadrático medio (19.09 %). Resulta interesante destacar que el método KOG12 tiene el mismo comportamiento que el resto de los métodos, dando un error cuadrático medio pequeño (1.44 %). Hay nueve modelos que subestiman la irradiación, pero con diferencias mínimas respecto al valor experimental (inferiores a 1 %), salvo el método KOG35 que presenta un valor de -10.90 %. El resto de métodos sobreestiman el valor de irradiación.

Como en los controles anteriores la Figura 5.3 representa el mapa de irradiación para el modelo KOC35.

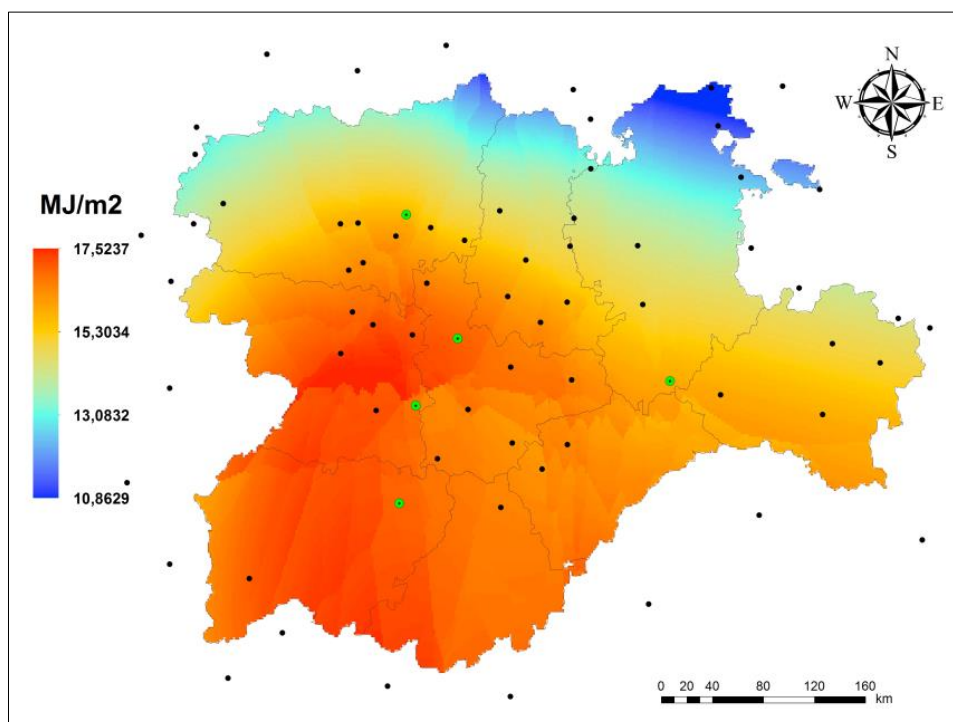


Figura 5.3. Mapa de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenido para el control C3 con los métodos de interpolación KOC35

5.2.4 CONTROL 4 (C4)

La Tabla 5.7 muestra las características de las estaciones elegidas como control C4. El control C4 está caracterizado por tener valores homogéneos de coordenadas geográficas y de irradiación global media.

Tabla 5.7. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C4

| Estación | latitud ($^{\circ}$) | longitud ($^{\circ}$) | altitud (m s.n.m.) | GHI ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) |
|----------|------------------------|-------------------------|--------------------|---|
| LE020 | 42.506 | -5.43 | 791 | 16.3909 |
| PA030 | 42.081 | -4.72 | 744 | 16.4520 |
| SG010 | 41.301 | -4.30 | 870 | 15.9209 |
| VA020 | 41.228 | -5.21 | 793 | 17.2612 |
| ZA040 | 41.880 | -5.39 | 701 | 16.8603 |

En este caso, se observa que todos los valores tienen un RMSE menor del 6 %, y, en general, mayores que en el control anterior (C3), pero sin diferencias destacables. El método matemático de interpolación que presenta el mayor error cuadrático medio, es el KOG35 pero mucho menor que en los controles anteriores (5.68 %). El método de interpolación que muestra el menor valor de RMSE, es el método KUL12 (1.95 %), que también coincide con el que presenta el menor error absoluto medio, MAE (1.30 %).

Si se analiza el error medio, la mayoría de los modelos subestiman el valor de la irradiación y sólo cuatro la sobreestiman. El MBE, en general, es bajo y, como en el control C3, los valores son menores que uno. Cabe destacar la diferencia que presenta el modelo que peor se ajusta a la zona en estudio, el KOG35, con un MBE de -4.12 %, el más elevado en comparación con los restantes métodos.

Todos los valores de los parámetros estadísticos obtenidos con la aplicación de los 29 métodos matemáticos de interpolación con este control, están reflejados en la Tabla 5.8.

Tabla 5.8. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C4 (LE020, PA030, SG010, VA020 Y ZA040), con todos los métodos de interpolación

| CONTROL 4 | | | | | | | |
|-------------------|--------------|-------------|-------|-------|-------|------|------|
| | | | RMSE | MBE | MAE | MAPE | |
| KRIGING 12 PUNTOS | ORDINARIO | Esférico | 2.80 | -0.75 | 2.32 | 2.36 | |
| | | Circular | 2.80 | -0.75 | 2.32 | 2.35 | |
| | | Exponencial | 2.74 | -0.75 | 2.28 | 2.31 | |
| | | Gaussiano | 4.02 | 0.41 | 3.18 | 3.21 | |
| | | Lineal | 2.79 | -0.74 | 2.31 | 2.34 | |
| | UNIVERSAL | Lineal | 1.95 | -0.47 | 1.30 | 1.32 | |
| | | Cuadrático | 2.52 | -0.71 | 2.33 | 2.35 | |
| KRIGING 33 PUNTOS | ORDINARIO | Esférico | 2.85 | -0.63 | 2.29 | 2.32 | |
| | | Circular | 2.89 | -0.62 | 2.32 | 2.35 | |
| | | Exponencial | 2.74 | -0.51 | 2.23 | 2.26 | |
| | | Gaussiano | 5.68 | -4.12 | 4.18 | 4.26 | |
| | | Lineal | 4.85 | -1.58 | 3.40 | 3.48 | |
| | UNIVERSAL | Lineal | 4.15 | 2.19 | 3.80 | 3.82 | |
| | | Cuadrático | 2.84 | -1.17 | 2.14 | 2.17 | |
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.77 | -0.01 | 3.13 | 3.10 |
| | | | 0.01 | 3.65 | -0.01 | 3.01 | 2.97 |
| | | | 0.001 | 3.49 | 0.01 | 2.82 | 2.78 |
| | TENSIÓN | PESO | 1 | 3.37 | -0.01 | 2.69 | 2.65 |
| | | | 5 | 3.32 | -0.01 | 2.66 | 2.62 |
| | | | 10 | 3.27 | -0.01 | 2.63 | 2.59 |
| SPLIN 33 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.74 | -0.01 | 3.05 | 3.00 |
| | | | 0.01 | 3.62 | -0.01 | 2.95 | 2.90 |
| | | | 0.001 | 3.40 | -0.01 | 2.75 | 2.70 |
| | TENSIÓN | PESO | 1 | 3.24 | -0.01 | 2.60 | 2.55 |
| | | | 5 | 3.21 | -0.01 | 2.58 | 2.53 |
| | | | 10 | 3.17 | -0.01 | 2.56 | 2.52 |
| I DW | 12 PUNTOS | | 2.50 | -0.50 | 2.10 | 2.13 | |
| | 35 PUNTOS | | 2.41 | 0.12 | 1.97 | 1.98 | |
| VECINO PROXIMO | | 12 PUNTOS | 2.68 | -0.35 | 2.15 | 2.18 | |

En la Figura 5.4. se representa el mapa de irradiación media anual diaria, obtenido con el método que menor RMSE y MAE presenta en este control, el KUL12.

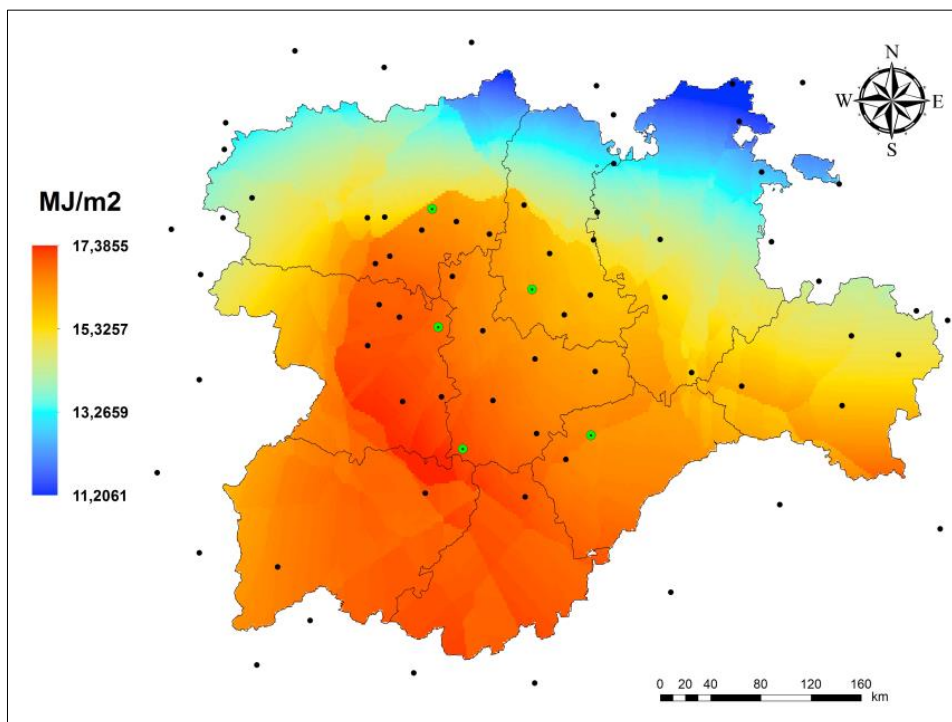


Figura 5.4. Mapa de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenido para el control C4 con los métodos de interpolación KUL12.

5.2.5 CONTROL 5 (C5)

Las coordenadas geográficas y la irradiación media anual diaria de las estaciones empleadas para este control se han recogido en la Tabla 5.9. En este caso, de las cinco estaciones con las que se trabaja, tres ya han sido utilizadas en otros controles, la LE020 (en todos los anteriores), SA030, usada en el control C3 y la estación VA020 utilizada en el control C4. En este control C5, todas las estaciones tienen parecidos valores de longitud y una menor diferencia de altitud que en los controles anteriores (150 m). Los valores de irradiación son elevados y semejantes. Todas las estaciones están situadas en la zona oeste de la región con diferencia de latitud entre ellas.

Tabla 5.9. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C5

| Estación | latitud ($^{\circ}$) | longitud ($^{\circ}$) | altitud (m s.n.m.) | GHI ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) |
|----------|------------------------|-------------------------|--------------------|---|
| LE020 | 42.506 | -5.43 | 791 | 16.3909 |
| SA030 | 40.990 | -5.48 | 815 | 17.3493 |
| VA020 | 41.228 | -5.21 | 793 | 17.2612 |
| ZA060 | 41.934 | -5.67 | 738 | 17.0065 |
| ZA080 | 41.508 | -5.37 | 650 | 16.8129 |

La Tabla 5.10. recoge los valores de los errores obtenidos al aplicar todos los métodos matemáticos de interpolación para este control C5, como en los casos anteriores. Se observa que los dos métodos Kriging Gaussiano, KOG12 y KOG35, presentan los errores cuadráticos medios más elevados (17.65 % y 4.32 %). El método que mejor se ajusta a la zona de estudio es el KUC35 con un RMSE de 1.36 % y un MAE de 1.19 %. Este método no coincide con el de menor error absoluto medio (MAE), que es el SR(0.001)12, con un valor de 1.04 %.

En este control hay 22 valores con un RMSE menor del 2 %, lo que representa un 75.8 % y otros cuatro con error menor del 3 %. Se observa que todos los métodos sobrestiman el valor de la irradiación media anual diaria (valores MBE positivos), destacan sobre todo los métodos KOG35 (10.49 %) y KUL35 (3.88 %).

Tabla 5.10. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C5 (LE020, SA030, VA020, ZA060 y VA080), con todos los métodos de interpolación

| CONTROL 5 | | | | | | | |
|-------------------|--------------|-------------|-------|------|-------|------|---------|
| | | | RMSE | MBE | MAE | MAPE | |
| KRIGING 12 PUNTOS | ORDINARIO | Esférico | 1.65 | 0.74 | 1.38 | 1.37 | |
| | | Circular | 1.62 | 0.72 | 1.36 | 1.35 | |
| | | Exponencial | 1.79 | 0.83 | 1.47 | 1.45 | |
| | | Gaussiano | 4.32 | 2.57 | 3.71 | 3.68 | |
| | | Lineal | 1.61 | 0.72 | 1.35 | 1.34 | |
| | UNIVERSAL | Lineal | 1.78 | 0.77 | 1.27 | 1.26 | |
| | | Cuadrático | 1.78 | 0.77 | 1.27 | 1.26 | |
| KRIGING 33 PUNTOS | ORDINARIO | Esférico | 1.53 | 0.91 | 1.23 | 1.22 | |
| | | Circular | 1.53 | 0.90 | 1.26 | 1.25 | |
| | | Exponencial | 1.73 | 1.10 | 1.35 | 1.33 | |
| | | Gaussiano | 17.65 | 1.05 | 11.69 | 1.15 | |
| | | Lineal | 1.69 | 0.74 | 1.50 | 1.49 | |
| | UNIVERSAL | Lineal | 4.35 | 3.88 | 3.88 | 3.90 | |
| | | Cuadrático | 1.36 | 0.42 | 1.19 | 1.18 | |
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 2.47 | 1.64 | 1.94 | 1.98 |
| | | | 0.01 | 1.90 | 1.28 | 1.43 | 1.44 |
| | | | 0.001 | 1.57 | 1.05 | 1.05 | 1.06 |
| | TENSIÓN | PESO | 1 | 1.59 | 1.04 | 1.10 | 1.10 |
| | | | 5 | 1.61 | 1.06 | 1.14 | 1.15 |
| | | | 10 | 1.62 | 1.060 | 1.16 | 1.17 |
| SPLIN 33 PUNTOS | REGULARIZADO | PESO | 0.1 | 2.20 | 1.60 | 1.78 | 1.81 |
| | | | 0.01 | 1.20 | 1.44 | 1.55 | 1.57 |
| | | | 0.001 | 1.78 | 1.25 | 1.25 | 1.27 |
| | TENSIÓN | PESO | 1 | 1.70 | 1.19 | 1.22 | 1.23 |
| | | | 5 | 1.71 | 1.19 | 1.24 | 1.25 |
| | | | 10 | 1.71 | 1.18 | 1.25 | 1.26276 |
| IDW | 12 PUNTOS | | 2.06 | 0.99 | 1.74 | 1.72 | |
| | 35 PUNTOS | | 2.41 | 1.75 | 1.90 | 1.87 | |
| VECINO PROXIMO | | 12 PUNTOS | | 1.62 | 1.21 | 1.21 | 1.37 |

En la Figura 5.5. se representan los mapas de irradiación obtenidos para los métodos que han dado valores pequeños de RMSE y MAE (KUC35) y el menor valor de MAE, Splin Regularizado con peso 0.001 y 12 puntos.

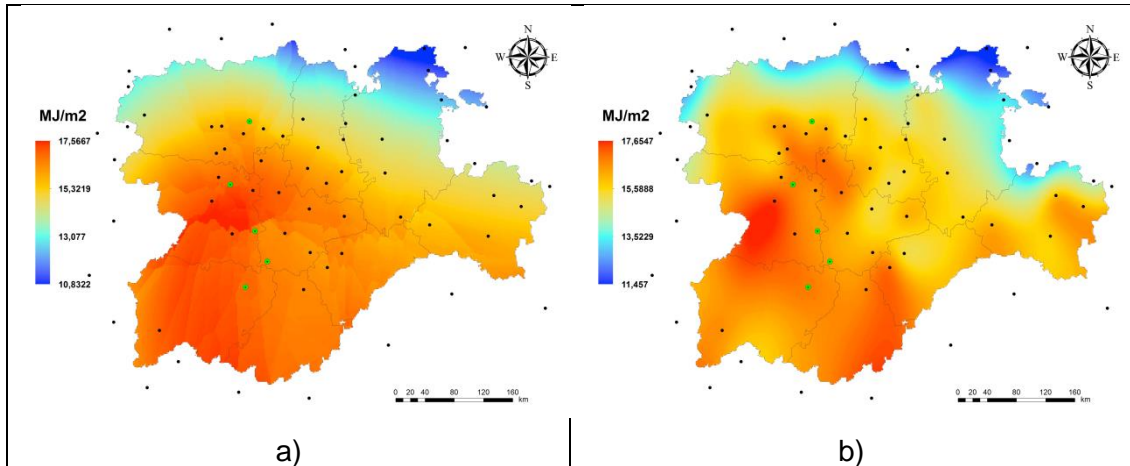


Figura 5.5. Mapas de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenidos para el control C5 con los métodos de interpolación: a) KUC35 y b) SR(0.001)12

5.2.6 CONTROL AEMET (C6)

Los cinco controles analizados hasta el momento se han realizado con estaciones de la red SIAR, en este se va a trabajar con estaciones de la AEMET. Existen dos grandes diferencias entre este control C6 y los anteriores. La primera es el número de estaciones empleadas como base de interpolación, hasta ahora se ha trabajado con 66 y en este control con 71 (toda la red SIAR en Castilla y León y red fronteriza). La segunda es que el origen de los datos de las de estaciones de validación son totalmente diferentes a los de la base de interpolación, ya que las estaciones de la AEMET están situadas en entornos ciudadanos y las de la red SIAR son mayoritariamente rurales, y los datos de ambas redes son tratados de forma diferente.

Las estaciones empleadas en el control C6 son las de la AEMET en Castilla y León situadas en Salamanca, León, Valladolid y Soria. Las coordenadas geográficas y los valores de irradiación media anual diaria se muestran en Tabla 5.11.

Tabla 5.11. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C6

| Estación | latitud ($^{\circ}$) | longitud ($^{\circ}$) | altitud (m s.n.m.) | GHI ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) |
|------------|------------------------|-------------------------|--------------------|---|
| LEÓN | 42.511 | -5.64 | 916 | 16.5401 |
| VALLADOLID | 41.641 | -4.75 | 873 | 16.6198 |
| SALAMANCA | 40.959 | -5.50 | 790 | 16.8676 |
| SORIA | 41.775 | -2.48 | 1082 | 16.0446 |

Las estaciones están distribuidas por el territorio con lo que las características geográficas son menos homogéneas que en los casos anteriores, aunque los valores de irradiación global media anual diaria son semejantes, con diferencias máximas de $0.8 \text{ MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$. Los errores obtenidos al comparar los valores calculados mediante los métodos de interpolación con los medidos en las estaciones se muestran en la Tabla 5.12.

Se observa que el método de interpolación que da el mejor resultado es el KUC35, con un RMSE de 1.43 %, comparable con el obtenido en el control 5 (1.36 %). Coincide con el método de menor valor de MAE (1.26 %). El mayor error cuadrático medio lo presenta el método KOG35, como en la mayoría de los controles anteriores. Cabe destacar que en este caso, como en el C5, el método KOG12 con un RMSE de error 3.86 % mejora resultados con respecto a métodos anteriores y el método KUL35 con 5.1 % es el más desfavorable.

El 93 % de los métodos empleados tienen valores de error inferior al 4 %. Todos los métodos, excepto tres subestiman el valor de irradiación. Sólo los métodos, KUL12, KUC12 y KUL35 sobreestiman el valor de la irradiación.

Tabla 5.12. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C6 (León, Salamanca, Soria y Valladolid), con todos los métodos de interpolación

| CONTROL 6 (AEMET (44 SIAR)) | | | | | | | |
|-----------------------------|--------------|-------------|-------|-------|-------|------|------|
| | | | RMSE | MBE | MAE | MAPE | |
| KRIGING 12 PUNTOS | ORDINARIO | Esférico | 2.65 | -0.59 | 2.46 | 2.45 | |
| | | Circular | 2.65 | -0.58 | 2.45 | 2.44 | |
| | | Exponencial | 2.62 | -0.58 | 2.43 | 2.42 | |
| | | Gaussiano | 3.86 | -2.20 | 3.19 | 3.21 | |
| | | Lineal | 2.64 | -0.58 | 2.45 | 2.44 | |
| | UNIVERSAL | Lineal | 3.36 | 1.69 | 2.64 | 2.65 | |
| | | Cuadrático | 2.38 | 0.92 | 2.13 | 2.14 | |
| KRIGING 35 PUNTOS | ORDINARIO | Esférico | 2.62 | -0.58 | 2.43 | 2.42 | |
| | | Circular | 2.69 | -0.65 | 2.51 | 2.50 | |
| | | Exponencial | 2.62 | -0.53 | 2.49 | 2.43 | |
| | | Gaussiano | 5.70 | -2.74 | 4.57 | 4.62 | |
| | | Lineal | 2.60 | -0.20 | 2.34 | 2.33 | |
| | UNIVERSAL | Lineal | 5.10 | 3.19 | 3.67 | 3.70 | |
| | | Cuadrático | 1.43 | -0.08 | 1.26 | 1.25 | |
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.31 | -1.40 | 2.98 | 2.96 |
| | | | 0.01 | 3.19 | -1.22 | 2.94 | 2.95 |
| | | | 0.001 | 3.08 | -1.09 | 2.87 | 2.87 |
| | TENSIÓN | PESO | 1 | 3.00 | -1.01 | 2.80 | 2.80 |
| | | | 5 | 3.01 | -0.98 | 2.82 | 2.82 |
| | | | 10 | 3.01 | -0.96 | 2.83 | 2.83 |
| SPLIN 35 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.41 | -1.35 | 3.12 | 3.13 |
| | | | 0.01 | 3.34 | -1.23 | 3.11 | 3,11 |
| | | | 0.001 | 3.21 | -1.11 | 3.01 | 3,01 |
| | TENSIÓN | PESO | 1 | 3.07 | -1.02 | 2.89 | 2,90 |
| | | | 5 | 3.07 | -0.10 | 2.89 | 2,90 |
| | | | 10 | 3.07 | -0.98 | 2.90 | 2,90 |
| IDW | 12 PUNTOS | | 2.38 | -1.26 | 2.02 | 2.01 | |
| | 35 PUNTOS | | 2.23 | -1.06 | 1.69 | 1.68 | |
| VECINO PROXIMO | 12 PUNTOS | | 2.69 | -0.08 | 2.42 | 2.40 | |

El método KUC35 muestra los menores valores de RMSE y MAE y se representa en la figura 5.6.

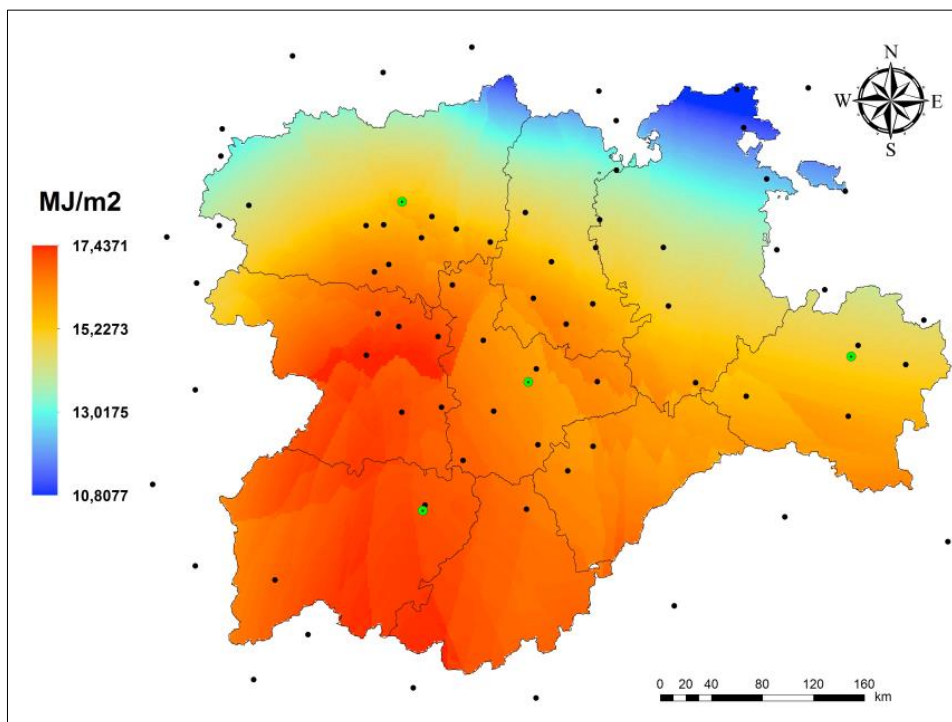


Figura 5.6. Mapa de GHI Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) obtenido para el control C6 con los métodos de interpolación KUC35

En este momento se puede afirmar que el método de interpolación que mejor se adapta para la zona en estudio es el Kriging Universal Cuadrático con 35 puntos. En la mayoría de los casos el comportamiento del Kriging Universal es mejor que el del Kriging Ordinario, salvo en alguna excepción en la que el Kriging Lineal se comporta mejor que el Cuadrático.

Para asegurarnos del ajuste de los métodos de interpolación obtenidos en el control C6 y conociendo que las estaciones empleadas para el control son totalmente diferentes a las que se han empleado como base de interpolación, se ha realizado un último control con ocho estaciones, cuatro de la AEMET y cuatro de la red SIAR (Control C7).

5.2.7 CONTROL AEMET Y SIAR (C7)

Las estaciones empleadas para este control junto con sus coordenadas geográficas y los valores de irradiación media anual diaria (valores experimentales), se han recogido en la Tabla 5.13

Tabla 5.13. Coordenadas geográficas e Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) de las estaciones del control C7

| Estación | latitud ($^{\circ}$) | longitud ($^{\circ}$) | altitud (m s.n.m.) | GHI ($\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$) |
|------------|------------------------|-------------------------|--------------------|---|
| LEON | 42.511 | -5.64 | 916 | 16.5401 |
| VALLADOLID | 41.641 | -4.75 | 735 | 16.6200 |
| SALAMANCA | 40.959 | -5.50 | 790 | 16.8676 |
| SORIA | 41.775 | -2.48 | 1082 | 16.0450 |
| LE080 | 42.374 | -5.02 | 856 | 16.3068 |
| PA020 | 41.946 | -4.49 | 750 | 16.0206 |
| VA030 | 41.310 | -4.69 | 750 | 16.7009 |
| ZA010 | 42.000 | -5.81 | 709 | 16.8512 |

Los valores de irradiación global media anual diaria están comprendidos entre [16.02, 16.86] con una diferencia entre los valores extremos del 5 %. Las coordenadas geográficas permiten cubrir uniformemente la zona en estudio. Los errores obtenidos comparando los valores interpolados con los medidos en las estaciones del control C7 están recogidos en la Tabla 5.14.

El método de interpolación con menor valor RMSE (1.51 %) es el IDW35, coincide con el de menor valor MAE (1.14 %). El resultado más discordante es el conseguido con el método KUL35, con un RMSE de 5.25 %, y un MAE de 3.71 %, este resultado es distinto de los obtenidos en controles C1 a C5. En general los valores obtenidos para los errores son semejantes a los conseguidos en el control C6. Si se analiza el MBE, tanto en este control como en control C6, los métodos de interpolación subestiman los valores de irradiación en 26 casos, coincidentes en ambos controles, y sobrestiman en el resto.

Tabla 5.14. Resultados estadísticos (RMSE, MBE, MAE, MAPE, %) para el control C7 (León, Salamanca, Soria y Valladolid, LE080, PA020, VA030 y ZA010), con todos los métodos de interpolación

| CONTROL C7 (AEMET (40 SIAR)) | | | | | | | |
|------------------------------|--------------|-------------|-------|-------|-------|------|------|
| | | | RMSE | MBE | MAE | MAPE | |
| KRIGING 12 PUNTOS | ORDINARIO | Esférico | 2.64 | -0.61 | 2.43 | 2.42 | |
| | | Circular | 2.63 | -0.61 | 2.43 | 2.42 | |
| | | Exponencial | 2.61 | -1.22 | 2.40 | 2.39 | |
| | | Gaussiano | 3.90 | -1.68 | 3.45 | 3.48 | |
| | | Lineal | 2.62 | -0.60 | 2.41 | 2.40 | |
| | UNIVERSAL | Lineal | 3.23 | 1.56 | 2.62 | 2.63 | |
| | | Cuadrático | 2.69 | 1.09 | 2.31 | 2.32 | |
| KRIGING 35 PUNTOS | ORDINARIO | Esférico | 2.65 | -0.62 | 2.45 | 2.44 | |
| | | Circular | 2.73 | -0.63 | 2.53 | 2.52 | |
| | | Exponencial | 2.61 | -0.57 | 2.41 | 2.40 | |
| | | Gaussiano | 4.70 | -3.21 | 3.53 | 3.57 | |
| | | Lineal | 3.21 | -0.39 | 3.00 | 3.00 | |
| | UNIVERSAL | Lineal | 5.25 | 3.34 | 3.71 | 3.74 | |
| | | Cuadrático | 2.13 | -0.38 | 1.99 | 1.98 | |
| SPLIN 12 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.29 | -1.44 | 2.92 | 2.93 |
| | | | 0.01 | 3.16 | -1.26 | 2.89 | 2.89 |
| | | | 0.001 | 3.05 | -1.12 | 2.83 | 2.83 |
| | TENSIÓN | PESO | 1 | 2.98 | -1.04 | 1.39 | 2.76 |
| | | | 5 | 2.99 | -1.01 | 2.80 | 2.80 |
| | | | 10 | 2.99 | -0.98 | 2.81 | 2.80 |
| SPLIN 35 PUNTOS | REGULARIZADO | PESO | 0.1 | 3.34 | -1.49 | 2.99 | 2.99 |
| | | | 0.01 | 3.28 | -1.36 | 2.98 | 2.9 |
| | | | 0.001 | 3.16 | -1.18 | 2.93 | 2.94 |
| | TENSIÓN | PESO | 1 | 3.05 | -1.05 | 2.85 | 2.85 |
| | | | 5 | 3.05 | -1.03 | 2.86 | 2.86 |
| | | | 10 | 3.04 | -1.01 | 2.86 | 2.86 |
| IDW | 12 PUNTOS | | 1.64 | -0.44 | 1.25 | 0.44 | |
| | 35 PUNTOS | | 1.51 | -0.05 | 1.14 | 1.14 | |
| VECINO PROXIMO | 12 PUNTOS | | 2.67 | -0.11 | 2.39 | 2.37 | |

5.2.8 DISCUSIÓN DE LOS RESULTADOS

A la vista de los resultados mostrados en los apartados anteriores, se puede considerar el método de interpolación KUC35 como el más aconsejable para conseguir valores de irradiación media anual diaria en la Comunidad de Castilla de León. El error que presenta está por debajo del 3 %. No obstante, se puede generalizar que cualquier método de los empleados en este trabajo, podría ser utilizado para esta labor, excepto el método KOG, en las dos versiones, 12 y 35 puntos. Los errores están en todos los casos por debajo del 7 %.

El último control (C7) se ha realizado para analizar la influencia de las estaciones que se emplean como base de interpolación. En la Tabla 5.15 se han recogido los valores experimentales de irradiación global media anual diaria de las cuatro estaciones de la AEMET, los valores interpolados para estas mismas estaciones con el método seleccionado (KUC35), utilizando la base de interpolación correspondiente al control C6 y los valores interpolados utilizando la base de interpolación del control C7.

Tabla 5.15. Irradiación Global Media Anual Diaria ($\text{MJ}\cdot\text{m}^{-2}$) experimental e interpolada en los controles C6 y C7

| | GHI/ $\text{MJ}\cdot\text{m}^{-2}$ | Interpolación C6 | Interpolación C7 |
|------------|------------------------------------|------------------|------------------|
| Valladolid | 16.6197 | 16.4604 | 17.0119 |
| León | 16.5400 | 16.3059 | 16.1202 |
| Salamanca | 16.8676 | 17.2652 | 17.2589 |
| Soria | 16.0445 | 15.4295 | 15.9342 |

Se observa que el valor obtenido en la instalación de Salamanca no se ve influido por la base de interpolación formada por estaciones de la red SIAR. Las cuatro estaciones utilizadas en el control están suficientemente alejadas y en zonas con alta densidad de estaciones, por eso no influyen en la de interpolación. En el caso de la instalación de Soria, el error es sensiblemente inferior al aplicar la base de interpolación del control C7, base con menor número de estaciones que el control C6. Al haber menos estaciones en la zona central la variación es más suave.

En el caso de Valladolid y León al estar las dos en una zona con muchas estaciones en la base de interpolación y muy próximas, la influencia de las cuatro estaciones que se quitan para el control C7 es importante, ya que las cuatro omitidas están todas próximas a LE080 y VA 030, por eso empeoran los resultados.

La Tabla 5.6 muestra un resumen de los resultados incluyendo los métodos con menores valores RMSE y MAE en cada control.

Tabla 5.16. Resumen de los métodos de interpolación con menores RMSE y MAE calculados en los siete controles.

| | C1 | C2 | C3 | C4 | C5 | C6 | C7 |
|----------|--------|-------|-------|-------|-------------|-------|-------|
| Modelo | KOL35 | KOL35 | KUC35 | KUL12 | SR(0.01)35 | KUC35 | IDW35 |
| RMSE (%) | 2.12 | 1.88 | 1.26 | 1.95 | 1.20 | 1.43 | 1.51 |
| Modelo | KOEF35 | KUC35 | KUC35 | KUL12 | SR(0.001)12 | KUC35 | IDW35 |
| MAE (%) | 1.77 | 1.06 | 1.07 | 1.30 | 1.05 | 1.26 | 1.14 |

A la vista de los resultados (fijándonos en el control C6) el método Kriging Universal Cuadrático de 35 puntos es el que mejor se adapta a la zona, con menor RMSE y MAE. En general el método Kriging presenta buenos resultados en la mayoría de los controles.

Analizando todos los controles efectuados, el método *KRIGING* está en primera posición en el 71 % de casos (menor error), tanto analizando el parámetro RMSE como el MAE. Un 70 % de los casos mejores se obtienen con el *UNIVERSAL* y de éstos el 57 % con el *CUADRÁTICO* y todos ellos con 35 puntos. En resumen, consideramos que el método KUC35 es el óptimo para la labor propuesta.

De la misma forma, en la Tabla 5.17, se han reflejado los métodos con peor comportamiento en la zona en estudio.

Se observa en la tabla que, con todos los controles aplicados, el método con errores superiores (MAE y RMSE) es el *KRIGING ORDINARIO GAUSSIANO* de 35 puntos.

En todos los casos el MAE es menor que el RMSE, corroborando la hipótesis de Willmott et al. (2005): los errores cuadráticos medios acumulan las diferencias aumentando considerablemente el error.

Tabla 5.17. Resumen de los métodos de interpolación con mayores RMSE y MAE calculados en los siete controles

| | C1 | C2 | C3 | C4 | C5 | C6 | C7 |
|----------|-------|-------|-------|-------|-------|-------|-------|
| Modelo | KOG35 | KOG35 | KOG35 | KOG35 | KOG35 | KOG35 | KOC35 |
| RMSE (%) | 18.49 | 14.85 | 19.09 | 5.68 | 17.65 | 5.70 | 5.25 |
| Modelo | KOG35 | KOG35 | KOG35 | KOG35 | KOG35 | KOG35 | KOG35 |
| MAE (%) | 9.71 | 8.95 | 13.18 | 4.18 | 11.69 | 4.57 | 3.70 |

Se va a analizar las estaciones de la red SIAR que se han empleado en más de un control. Estas estaciones son las que aparecen en la Tabla 5.18, en ella se muestra el error obtenido (diferencia entre valor experimental e interpolado), para el mismo método de interpolación, que se ha determinado como el mejor para esta zona, KUC35.

Tabla 5.18. Errores observados (diferencia entre valor experimental e interpolado) para las estaciones de la red SIAR empleadas en más de un control

| RMSE (KUC35) | C1 | C2 | C3 | C4 | C5 |
|--------------|------|------|-------|------|-------|
| LE020 | 0.30 | 0.25 | 0.13 | 0.12 | 0.04 |
| PA080 | 0.64 | 0.65 | X | X | X |
| ZA060 | X | 0.20 | X | X | -0.18 |
| ZA080 | X | X | -0.02 | X | -0,14 |
| SA030 | X | X | 0.22 | X | 0.28 |
| VA020 | X | X | X | 0.28 | 0.36 |

La estación LE020: se observa una disminución del error en cada control, la razón es la estación próxima LE060 utilizada en el control C1 y no en el resto. En el control C2 se ha sustituido LE060 por ZA060, estación relativamente próxima a la anterior, razón por la que la diferencia de errores entre C1 y C2 es mínima. En el caso de PA080: la situación de esa estación en los dos controles no cambia, la base de interpolación es prácticamente la misma en la zona próxima a la estación, no hay una influencia significativa en la interpolación. Para las estaciones SA030 y VA020 la situación es

semejante, los errores aumentan al tener una estación en el control próxima a ellas. Las estaciones ZA060 y ZA080: en ambos casos en el control C5, hay estaciones próximas por lo que se empeora el resultado dando valores subestimados de irradiación.

5.3 ANALISIS DE LOS MODELOS DE RADIACIÓN

Conocidos los valores de irradiación global media anual diaria sobre plano horizontal a partir de los mapas de irradiación calculados en el apartado anterior para toda la Comunidad de Castilla y León, se trata ahora de calcular la irradiación global media anual diaria sobre plano vertical en cada una de las direcciones cardinales. Para hacer la conversión de los datos sobre plano horizontal a vertical, se ha recurrido a cuatro de los modelos habituales en la bibliografía: *Circunsolar* (Iqbal, (1983)), *Isotrópico* ((Iqbal, (1983)), *Klucher* (Klucher, (1979)) y *Hay* (Hay y Davies, (1980)). Los datos experimentales necesarios para contrastar los resultados de los modelos han sido recogidos en la estación radiométrica situada en la Escuela Politécnica Superior de la Universidad de Burgos. La bondad de los modelos se ha establecido en función de los parámetros estadísticos habituales, RMSE, MAE, MBE.

Los modelos seleccionados cubren toda la variedad de tipo de cielo. El modelo *Circunsolar* se refiere a cielos claros y despejados. El modelo *Isotrópico* supone que el cielo se comporta de forma uniforme, asumiendo el cielo cubierto. Los otros dos modelos, *Hay* y *Klucher*, son modelos anisotrópicos que consideran la influencia de la componente difusa de la irradiación de manera distinta. El modelo de *Klucher* se basa en las condiciones de cielo claro y añade las de cielo cubierto, mientras que el modelo *Hay* interpreta la anisotropía como suma de dos componentes, la componente circunsolar directa del sol y la componente isótropa de la bóveda celeste.

Los valores utilizados como referencia han sido los recogidos a lo largo del año 2014 con los filtros de calidad ya comentados. El número total de datos

considerados válidos para ser utilizados es de 347 valores diarios de irradiación global diaria vertical en las cuatro orientaciones cardinales (norte, sur, este y oeste) acompañados de los correspondientes datos de irradiación horizontal global, difusa y directa.

5.3.1 RESULTADOS EN LA DIRECCIÓN NORTE

En esta orientación, el Sol es visto por el sensor solamente durante unos días del año. Este hecho obliga a trabajar con aquellos valores correspondientes a los días en los que la órbita solar es mayor de 90° tomando el sur como origen (para el este y el oeste) que son un total de 172 días.

Con cada uno de los modelos seleccionados (*Circunsolar* (C), *Isotrópico* (I), *Klucher* (K) y *Hay* (H)) se ha calculado el valor medio diario anual de irradiación vertical en la dirección Norte, comparando este resultado con el valor registrado por el sensor en la misma orientación. Los resultados de los errores de los tres estadísticos se han recogido en la Tabla 5.19

Tabla 5.19. Errores anuales para los cuatro modelos de radiación vertical dirección Norte

| | C | I | K | H |
|----------|-------|--------|-------|--------|
| RMSE (%) | 63,80 | 27.66 | 34.80 | 18.31 |
| MBE (%) | 61.53 | -24.49 | -6.47 | -13.49 |
| MAE (%) | 61.53 | 24.68 | 29.15 | 15.45 |

El modelo que tiene menor error cuadrático medio y menor error absoluto medio es el modelo *Hay* (H) (18.31 % y 15.45 % respectivamente). Este modelo subestima la irradiación media anual diaria (-13.49 %). El modelo *Isotrópico* (I), presenta un error ligeramente mayor en los dos casos (RMSE 27.66 % y MAE 24.68 %) y también subestima la irradiación.

El modelo *Klucher* (K), basado en condiciones de cielo claro incorporando las condiciones de cielo cubierto, se comporta peor que los anteriores, dando errores superiores con los dos estadísticos (34.80 % y 29.15 %) y también subestima el valor de la irradiación. El comportamiento del modelo *Circunsolar* (C), es el peor, como cabía esperar, ya que está definido

únicamente para días claros y sobrestima el valor de la variable. Este hecho, en primera aproximación puede parecer lógico, ya que este modelo identifica el cielo para días claros y despejados, y esa situación no es la más habitual en Burgos.

Se ha realizado un estudio adicional en forma mensual distribuyendo los 172 valores disponibles a lo largo del año. La figura 5.7 muestra los resultados del cálculo del RMSE con los cuatro modelos.

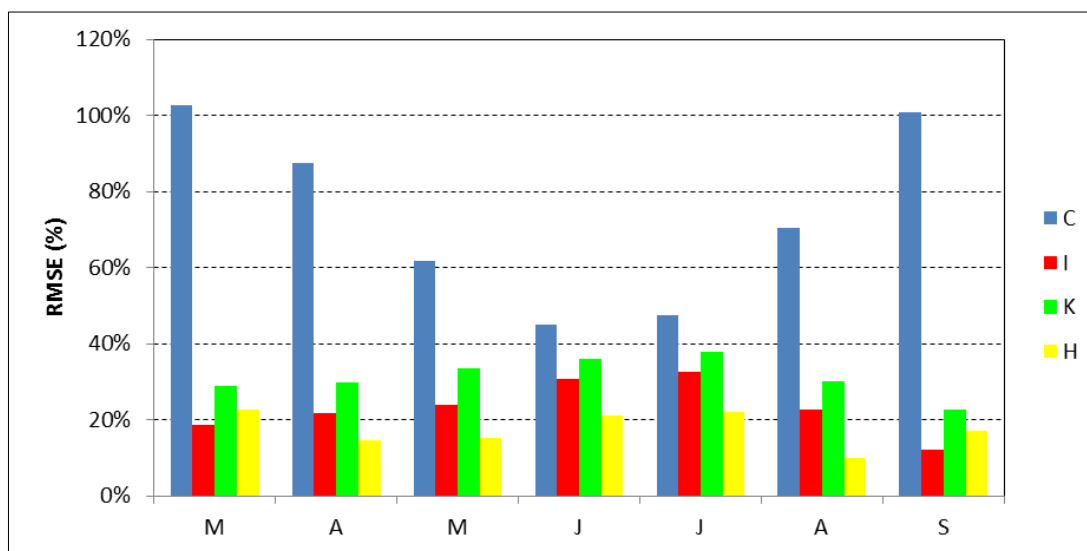


Figura 5.7. Valores de RMSE de media mensual diaria en superficies orientadas hacia el Norte

Analizando esta figura se aprecia que el modelo *Hay* (H) presenta el menor valor de RMSE durante los meses centrales del año mientras que el modelo *Isotrópico* (I), es el más aconsejable en los meses de Marzo y Septiembre. Por el contrario, el modelo de *Klucher* (K) se comporta peor en los meses de Junio y Julio y el *Circunsolar* (C) presentan valores RMSE excesivamente elevados en Marzo y Septiembre, coincidiendo con la tónica general de tipo de cielo en la ciudad durante esos meses. En la Tabla 5.20 se aprecian los resultados del cálculo del parámetro estadístico MAE para la dirección Norte calculado con los cuatro modelos, considerando los 172 datos distribuidos en los meses correspondientes. El modelo *Circunsolar* (C), muestra los errores más elevados todos los meses y el modelo de *Hay* (H), es el más aproximado, aunque empeora en los días en los que se espera que el cielo se encuentre más despejado.

Tabla 5.20. Valores de MAE mensuales para los cuatro modelos de irradiación para la orientación Norte.

| MAE (%) | M | A | M | J | J | A | S |
|---------|--------|-------|-------|-------|-------|-------|-------|
| C | 108.84 | 86.19 | 60.31 | 42.05 | 45.94 | 69.31 | 98.09 |
| I | 19.98 | 19.75 | 23.65 | 29.37 | 31.94 | 20.92 | 9.13 |
| K | 44.15 | 26.54 | 28.84 | 31.30 | 32.96 | 25.61 | 18.70 |
| H | 15.31 | 12.48 | 13.81 | 19.49 | 21.69 | 7.76 | 13.92 |

Se ha realizado también, un estudio por intervalos de irradiación, con el objetivo de ver el comportamiento de los cuatro modelos. La Figura 5.8 recoge los valores del RMSE para los cinco intervalos en los que se ha clasificado el valor de la irradiación (media anual diaria), medida por el sensor vertical en esta orientación.

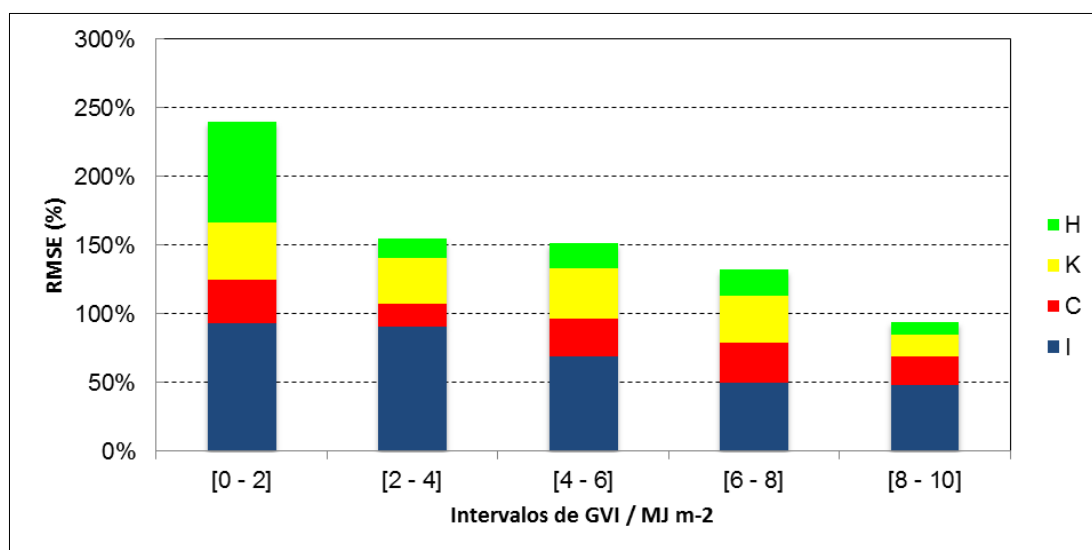


Figura 5.8. Valores de RMSE de irradiación media diaria ($GVI/ MJ \cdot m^{-2}$) sobre superficies vertical orientado hacia el Norte

Esta clasificación se ha realizado exclusivamente para los días que el sensor ve el Sol. Se observa que cuanto mayor es el valor de irradiación, mejor se comportan los modelos obteniendo con todos ellos el menor valor de RMSE dentro del intervalo de irradiación más elevada. El modelo más efectivo en todos los intervalos, excepto el de menores valores de irradiación, es el modelo de *Hay* (H), con errores comprendidos entre el 10 % y el 20 %. El modelo *Isotrópico* (I), se comporta muy bien en el intervalo de [2, 4] $MJ \cdot m^{-2}$, correspondiente a días cubiertos.

La hipótesis de que valores bajos de irradiación se corresponden con días cubiertos debe ser contrastada en función de la duración del día solar.

5.3.2 RESULTADOS EN LA DIRECCIÓN SUR

El sensor situado verticalmente orientado hacia el sur “ve” el sol durante todos los días del año, aunque hay días en los que la órbita descrita por el sol tiene un ángulo superior a 90°. Esto hace que el sensor no reciba radiación en algún momento del día que coincide con horas del amanecer y anochecer.

En la Tabla 5.21 se recogen los valores calculados de RMSE anual de los cuatro modelos empleados. En general los errores observados son muy elevados mostrando el modelo *Isotrópico* (I), el valor más bajo, pero comparable con los obtenidos por los dos modelos anisotrópicos y siendo el modelo *Circunsolar* (C), el de error más elevado.

Tabla 5.21. Errores anuales para los cuatro modelos de radiación vertical dirección Sur

| SUR | C | I | K | H |
|----------|--------|-------|-------|-------|
| RMSE (%) | 113.76 | 85.96 | 87.93 | 94.93 |

En el estudio de distribución mensual del error, mostrado en la Figura 5.9. según el valor del error absoluto medio (MAE) calculado, se observa que en general todos los modelos reducen considerablemente los errores en los meses de primavera y verano y que presentan valores del error mucho mayores en los meses de invierno y otoño. En el mes de agosto se observa que los cuatro modelos presentan el menor error absoluto medio (*Isotrópico* 7.56 %, *Hay* 9.04 %, *Klucher* 14.86 % y *Circunsolar* 25.89 %). Es una circunstancia curiosa, ya que el resultado debía ser diferente para el modelo *Circunsolar* (C) y el *Isotrópico* (I) (ambos consideran los cielos distintos, uno cubiertos y el otro claro).

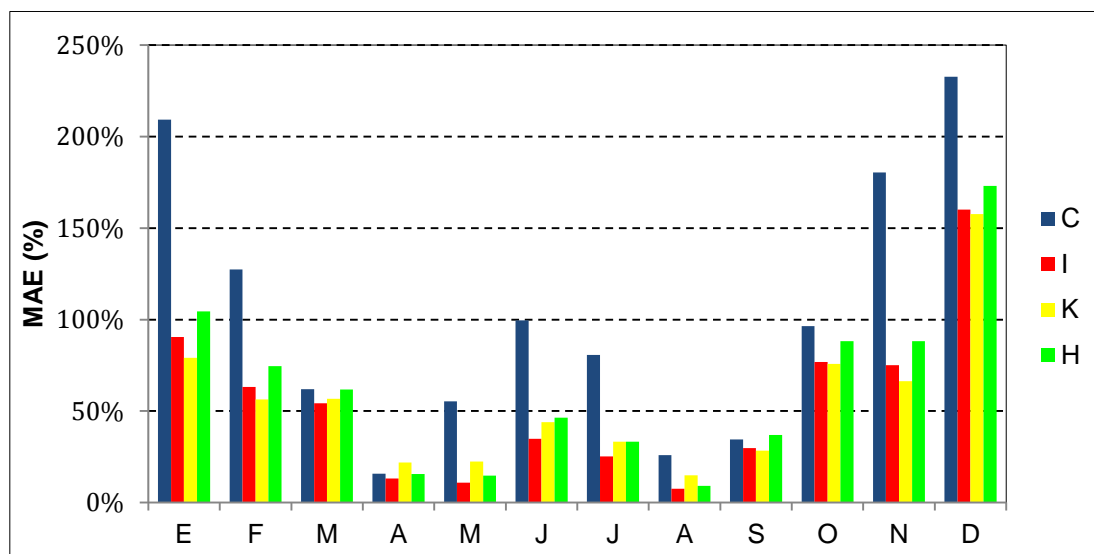


Figura 5.9. Valores de MAE de media mensual diaria en superficies orientadas hacia el Sur

Para ver cuál es la razón de ese comportamiento de los modelos *Isotrópico* (I) y *Circunsolar* (C), se ha hecho un estudio de los días del mes de agosto. Se ha observado la siguiente distribución:

1. 6 días muy nublados (difusa aproximadamente igual a global)
2. 7 días con bastantes nubes pero no completamente cubierto (global y difusa separadas)
3. 5 días que se pueden considerar como parcialmente cubiertos
4. 7 días bastante claros, pero con alguna nube (medio día claro y medio día cubierto)
5. 6 días totalmente despejados

Según esta distribución diaria, se puede decir que el comportamiento del modelo *Circunsolar* (C) es lógico, hay 11 días con un cielo prácticamente claro (6 días totalmente despejados y cinco días parcialmente cubiertos). Por otra parte el modelo *Isotrópico* (I) dará muy buenos resultados los 13 días cubiertos o muy cubiertos que hay en este mes, además de otros cinco días en los cuales el cielo está parcialmente cubierto, en ellos el resultado será también bueno.

El mes de abril presenta un comportamiento muy similar. Se ha observado que en este caso el número de días claros es de 20, por lo que el modelo *Circunsolar* (C) da mejor resultado que en el mes de agosto, y 10 días

cubiertos, por lo que el *Isotrópico* (I), da mayor error en abril que en agosto. Hay que tener en cuenta que este mes tiene menos horas de sol.

Los otros dos modelos, *Klucher* (K) y *Hay* (H), tienen un error similar en los meses de abril y agosto (en abril 25.93 % y 20.45 % y en agosto, 17.48 % y 10.81 %, respectivamente). El modelo de *Klucher* (K) se parece más al *Circunsolar* (C) en ambos meses, y el modelo *Hay* (H) se comporta más como un modelo *Isotrópico* (I) en los meses de verano y como un *Circunsolar* (C), en los meses de primavera. Esto se puede deber a la longitud del día, es decir al número de horas en las que hay irradiación.

El mes con mayor valor de MAE en todos los modelos es diciembre (*Circunsolar*, 233 %; *Isotrópico*, 160 %; *Klucher*, 158 %; *Hay*, 173 %). El modelo de *Klucher* (K) funciona mejor debido fundamentalmente a las condiciones de cielo cubierto que incorpora el autor. La mayoría de los días de diciembre son cubiertos o parcialmente cubiertos. Analizando los valores de irradiación (global, frente a difusa y directa) del mes de diciembre, se ha observado que 20 días se corresponden con condiciones de cielo cubierto, otros cuatro son parcialmente cubiertos (la difusa y la global son muy parecidas) y solo hay siete días claros y tres poco cubiertos, lo que justifica el error del modelo *Circunsolar* (C).

Igual que en la dirección Norte, se han clasificado los resultados por intervalos de irradiación, estudio que se muestra en la Figura 5.10 donde se ven los valores del MAE para los dos modelos que mejores resultados ofrecen, el modelo *Isotrópico* (I) y el modelo *Klucher* (K).

En esta figura se observa que el intervalo que presenta menor error es el comprendido entre 12 y 15 MJ·m⁻² y a partir de esos valores de irradiación el error va aumentando de forma considerable (de 20 % a 40 % y 80 % aproximadamente). Para valores de irradiación pequeños el error no es excesivamente grande. En este caso también parece lógico el comportamiento de los modelos: cuanto mayor es el valor de irradiación, más claros son los días, por lo que los modelos anisotrópicos se comportaran peor y aparece mayor error en los intervalos de mayor

irradiación. En los intervalos de poca irradiación, habituales en invierno, hay pocas horas de insolación y generalmente cielos claros, con lo que los modelos anisotrópicos dan peores resultados.

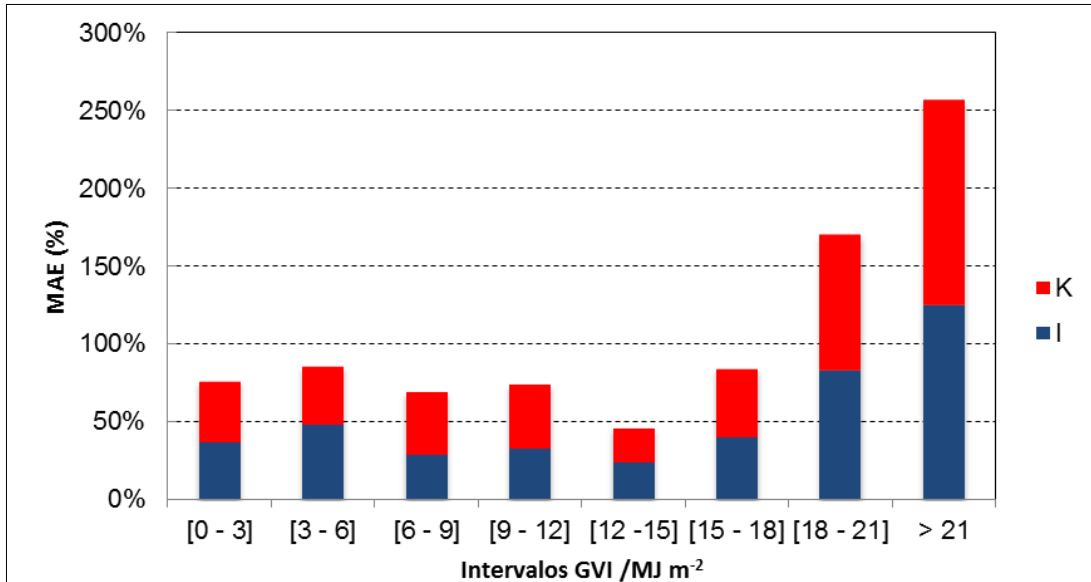


Figura 5.10. Valores de MAE de irradiación media diaria (GVI/ MJ·m⁻²) sobre superficie vertical orientado hacia el Sur

5.3.3 RESULTADOS EN LA DIRECCIÓN ESTE

Los parámetros de error se han calculado en este caso teniendo en cuenta que el ángulo acimutal es de 90°, por lo que hay gran parte del día en el que el sensor no “ve” el sol y por tanto hay valores muy pequeños de irradiación, no representativos, correspondientes a las horas de la tarde. En la Tabla 5.22 se recogen los valores del error cuadrático medio anual para los cuatro modelos empleados. Se observa que los valores del RMSE son parecidos para todos los modelos. El modelo *Circunsolar* (C), muestra los valores más bajos y el modelo *Hay* (H), los más elevados. En esta orientación los momentos en los que el sensor detecta irradiación son aquellos correspondientes a la mañana, desde el amanecer hasta mediodía, lo que influye en la respuesta de los modelos.

El valor medio anual calculado en la Tabla 5.22. es muy poco significativo, aún más que en las orientaciones ya estudiadas, ya que solo recorre el valor de irradiación correspondiente a medio día.

Tabla 5.22. Errores anuales para los cuatro modelos de radiación vertical dirección Este

| ESTE | C | I | K | H |
|------|-------|-------|-------|-------|
| RMSE | 57.13 | 66.13 | 68.22 | 71.51 |

El estudio mensual, cuyos resultados con respecto al valor del MAE se muestran en la Figura 5.11, muestra que todos los modelos arrojan muy buenos resultados durante el verano, fundamentalmente para los datos correspondientes al mes de agosto, con diferencias mínimas entre los modelos (*Circunsolar* 4.6 %, *Isotrópico* 6.3 %, *Klucher* 6.2 % y *Hay* 6.9 %).

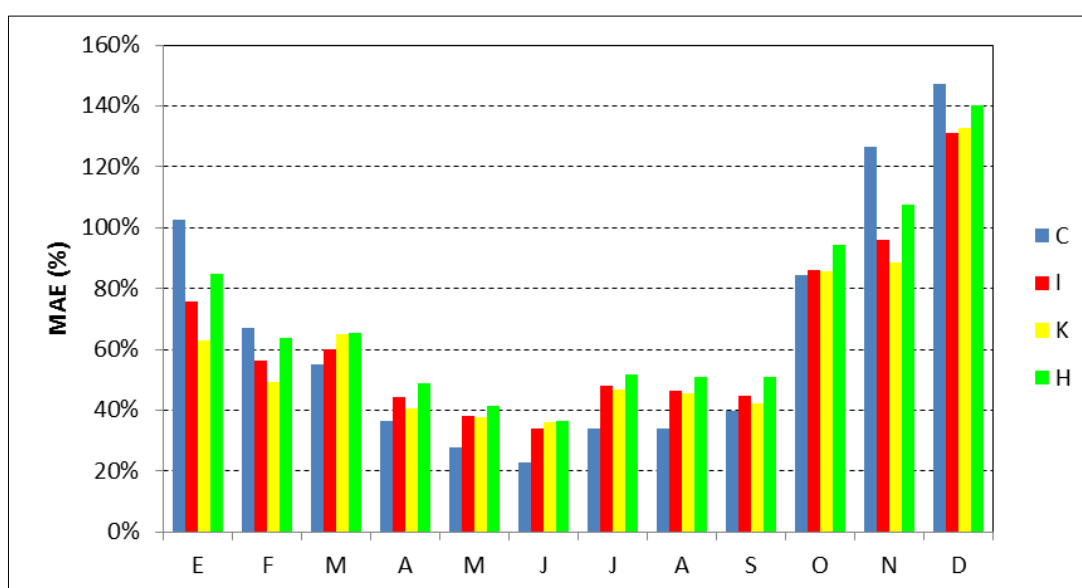


Figura 5.11. Valores de MAE de media mensual diaria en superficies orientadas hacia el Este

Los valores de los errores por intervalos de irradiación se recogen en la Figura 5.12. En ella se observa que para tres de los intervalos, correspondientes a valores bajos de irradiación, los errores de los cuatro modelos son aproximadamente iguales. Para los tres intervalos de mayor irradiación hay más diferencias entre los cuatro modelos, pero en general, los errores son menores. El modelo con menor error es el modelo *Circunsolar* (C) para el intervalo de mayor irradiación, tanto desde el punto de vista del RMSE como del MAE. El error mayor se obtienen para el modelo *Circunsolar* (C) en el intervalo de [0, 3] ($\text{MJ}\cdot\text{m}^{-2}$), respecto al valor del RMSE, pero respecto al valor del MAE el modelo que presenta el error mayor es el modelo *Hay* (H) en el intervalo intermedio [6, 9] ($\text{MJ}\cdot\text{m}^{-2}$).

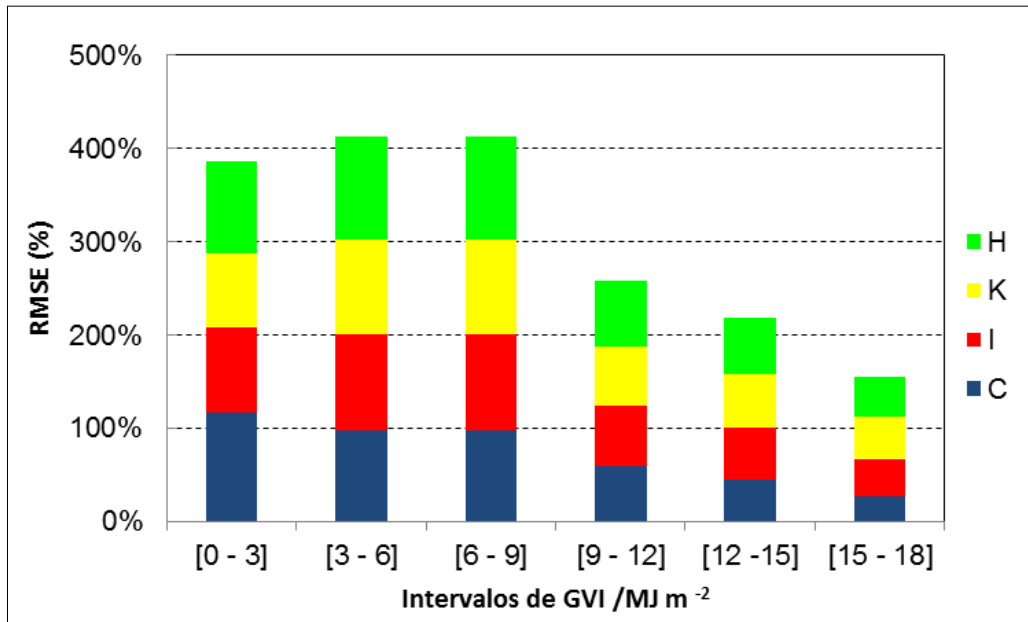


Figura 5.12 Valores de RMSE de irradiación media diaria (GVI/ MJ·m⁻²) sobre superficie vertical orientado hacia el Este

5.3.4 RESULTADOS EN LA DIRECCIÓN OESTE

Este caso es simétrico al presentado en la orientación Este, pero considerando el ángulo acimutal como -90°, lo que hace que el sensor no “vea” el Sol la parte del día correspondiente a la mañana. En la Tabla 5.23 se recogen los errores cuadráticos medios de los cuatro modelos empleados para la comparación con los valores experimentales. Para esta orientación, la respuesta es muy similar a la obtenida para la orientación Este. El modelo con errores menores es el *Circunsolar* (C) y el de errores más grandes es el modelo *Hay* (H), no apreciándose diferencias significativas entre los modelos *Klucher* (K) e *Isotrópico* (I).

Tabla 5.23. Errores anuales para los cuatro modelos de radiación vertical dirección Oeste

| OESTE | C | I | K | H |
|----------|-------|-------|-------|-------|
| RMSE (%) | 66.29 | 76.33 | 78.63 | 82.06 |

Los datos medios mensuales se grafican en función del valor del RMSE en la Figura 5.13. En este caso en los meses de verano, los cuatro modelos muestran un comportamiento comparable, con menores errores calculados

que en el resto de los meses y siendo el modelo *Circunsolar* (C), el que arroja menores valores del error RMSE.

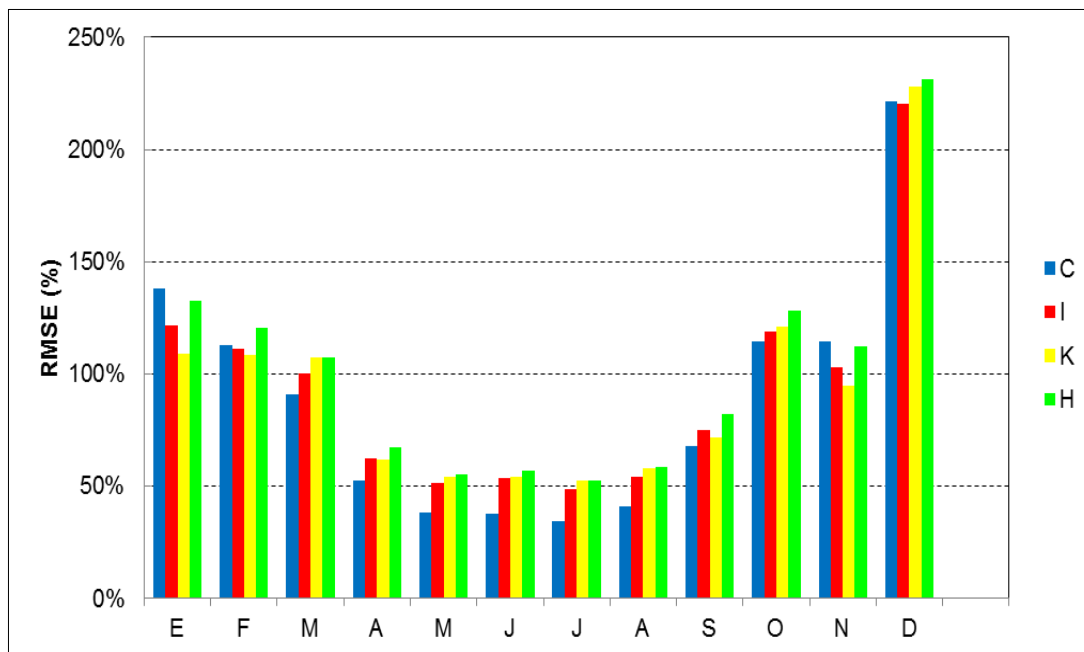


Figura 5.13. Valores de MAE de media mensual diaria en superficies orientadas hacia el Oeste

Sí que se observa una diferencia apreciable con los resultados obtenidos para la orientación Este. En el caso que nos ocupa (orientación Oeste), los meses centrales del año tienen un comportamiento parecido en los cálculos con los cuatro modelos, mientras que en la orientación Este el mes de agosto presentaba resultados sustancialmente diferentes al resto de los meses de verano, con errores mucho menores.

En la Figura 5.14 se muestran los valores del estadístico MAE de los modelos que mejor se ajustan a esta orientación. Se observa que el modelo *Klucher* (K), permite obtener menores valores de error en los intervalos de menor valor de irradiación, mientras que el modelo *Circunsolar* (C) funciona más adecuadamente cuando la irradiación es más alta. Siguiendo el razonamiento expuesto a lo largo de este apartado, los valores de irradiación bajos suelen corresponder con días cubiertos mientras que los días claros suelen ser días con valores elevados de irradiación, en los que el modelo *Circunsolar* (C), presenta un buen comportamiento.

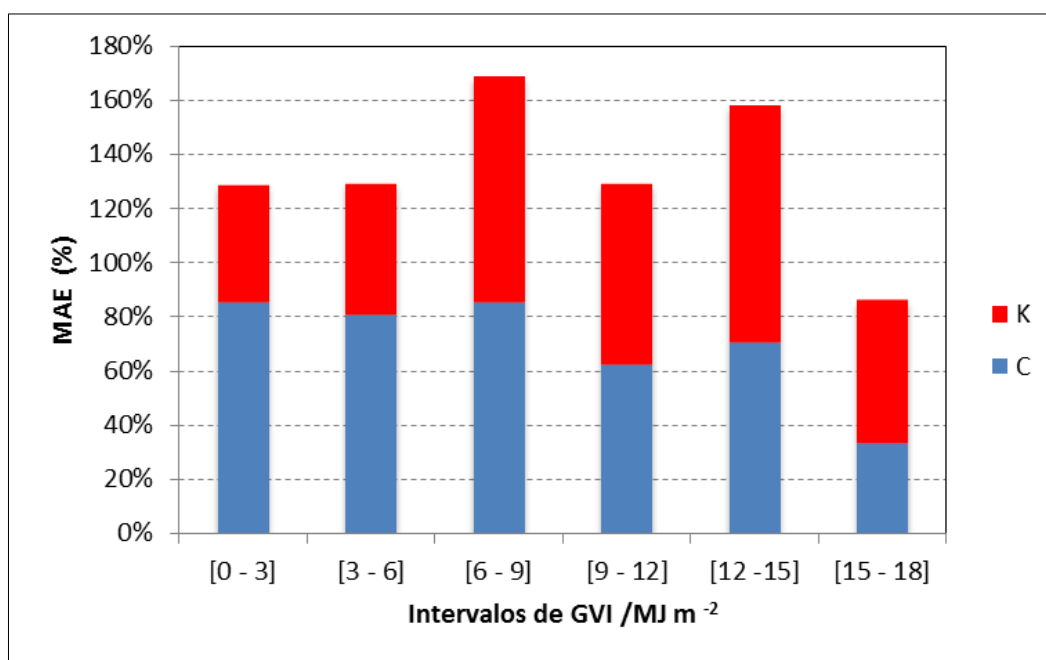


Figura 5.14. Valores de MAE de irradiación media diaria (GVI/ MJ·m⁻²) sobre Superficie vertical orientado hacia el Oeste

5.4 RESUMEN DE LOS RESULTADOS

La Tabla 5.24 refleja cuáles son los modelos considerados más adecuados para el cálculo de la irradiación media (anual, mensual o clasificados por intervalos de irradiación) para las cuatro orientaciones.

Tabla 5.24. Tabla resumen del comportamiento de los modelos considerados más adecuados para el cálculo de IGV en las direcciones cardinales

| | NORTE | | SUR | | ESTE | | OESTE | |
|------------------|----------|----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | RMSE (%) | MAE (%) | RMSE (%) | MAE (%) | RMSE (%) | MAE (%) | RMSE (%) | MAE (%) |
| Anual | H | H | I | I | C | C | C | C |
| Modelo/mes | H/Ab | I/Sept | I/Ab | I/Ab | C/JN | C/Ag | C/JI | C/JI |
| Clasif. por Int. | H (2-4) | H (8-10) | K (12-15) | I (12-15) | C (15-18) | C (15-18) | C (15-18) | C (15-18) |

En la Tabla 5.24 se observa que en la dirección Este y Oeste el mejor modelo es el *Circunsolar* (C), tanto en valores anuales como mensuales. Si se estudia el comportamiento por intervalos se aprecia que este modelo

sigue ofreciendo los mejores resultados, con menor error, en el intervalo de mayor irradiación cumpliendo la premisa de ser un modelo muy adecuado para cielos claros y con elevados valores de irradiación. En el caso de trabajar con valores medios mensuales el modelo también se ajusta bien en las épocas de altos valores de irradiación (meses de abril, junio y julio).

En el caso de la orientación Sur, el modelo *Isotrópico* (I) es el que mejor se ajusta a los valores experimentales, tanto en valores medios anuales como mensuales. En este caso el intervalo que ofrece menores errores es un intervalo intermedio con cielos cubiertos o parcialmente cubiertos, típicos de periodos primaverales (con valores de irradiación entre 12 y 15 MJ·m⁻²). Los modelos *Isotrópico* (I) y *Hay* (H), son los más adecuados para esta orientación.

Para la orientación Norte, el modelo *Hay* (H), resulta el más adecuado ofreciendo un buen ajuste para cielos cubiertos y/o parcialmente cubiertos. En este caso, el comportamiento del modelo *Hay* (H), es mejor al del modelo *Isotrópico* (I). Este hecho puede ser debido a un error sistemático detectado en las medidas experimentales atribuido a un reflejo recibido por el sensor en determinados momentos del día proveniente del perímetro metálico de la zona límite de la terraza donde se sitúa la estación radiométrica. Este efecto hace que el valor medido sea superior al real, influyendo sobre los resultados finales.

CONCLUSIONES

Para finalizar ponemos de manifiesto las conclusiones a las que hemos llegado y que se resumen en los siguientes puntos:

1. Se ha conseguido el objetivo fundamental de esta tesis que ha consistido en la estimación del recurso solar sobre superficies verticales. Este objetivo es de aplicación inmediata para: a) la localización adecuada de superficies verticales (fachadas) susceptibles de acoger de forma eficiente instalaciones fotovoltaicas para la generación distribuida de electricidad; b) el cálculo de la eficiencia energética de edificios a través de la determinación de cargas térmicas; c) la estimación de la iluminación natural disponible en un edificio y d) el cálculo de elementos de sombra para el mejor confort en un edificio
2. Se ha realizado una importante revisión bibliográfica que demuestra la utilidad de los métodos de interpolación para la obtención de valores locales de datos climatológicos a partir de un número relativamente escaso de estaciones terrestres repartidas en el territorio en estudio. Se ha realizado otra revisión bibliográfica para determinar los modelos teóricos que permiten trasladar los datos de radiación solar global sobre plano horizontal a plano inclinado (incluso vertical), determinando que los modelos más aplicables para el estudio y que cubren toda la tipología posible de cielo son los modelos Isotrópico, Circunsolar, Klucher y Hay.
3. Se ha localizado una red de estaciones meteorológicas con datos de irradiación solar global horizontal que verifica dos condiciones fundamentales: elevada calidad de los valores registrados y suficiente número de estaciones en la zona de estudio (Castilla y León), para aplicar métodos matemáticos de interpolación con resultados significativos. Seleccionada para tal fin la red de estaciones de SIAR, se ha realizado un primer análisis de los datos disponibles en un periodo de siete años que permite garantizar la calidad de los registros y determinar

las características en cuanto a irradiación de la zona estudiada. En este sentido se ha concluido:

- La provincia con más irradiación media anual diaria es Zamora y la de menor irradiación, Burgos. La diferencia entre ellas de un 13 %
- En el periodo de siete años estudiado se ha detectado un año energéticamente “bueno” (2009), es decir, con una irradiación global horizontal media anual diaria mayor que la media y otro año energéticamente “malo” (2013) con un valor por debajo de la media.
- La mayor parte de las estaciones tiene un valor medio anual diario comprendido entre 15 a 18 MJ·m⁻². Esta distribución se repite casi todos los años.
- Burgos tiene un 18 % de los días cubiertos y en Zamora el 50 % de los días son claros. La media del índice de claridad en la zona es de 0.56 (corresponde a días parcialmente cubiertos).
- No hay diferencias estadísticamente significativas para los valores de irradiación entre los años analizados. Se ha contrastado la homogeneidad de las varianzas, mediante la prueba de Levene de verificación de varianzas.
- Se han encontrado dos estaciones, BU101 y BU070, con valores de irradiación extremos globales, que presentan todos los años valores de irradiación significativamente menores que el resto de las estaciones. Estas dos no pertenecen a la cuenca del Duero.
- Se han detectado dos estaciones, una de Valladolid y otra de Salamanca, con valores extremos locales, que estando dentro del intervalo normal son significativamente diferentes de las de su entorno.
- Se ha encontrado una relación entre los valores de irradiación con la latitud, longitud y altitud.

4. La herramienta que ha permitido la obtención de los mapas de irradiación solar global horizontal en Castilla y León ha sido la plataforma ArcGis (sistema de análisis de la información geográfica). El primer paso ha consistido en la selección de los métodos matemáticos de interpolación. Posteriormente se ha realizado una validación cruzada de los resultados obtenidos con cada uno de los métodos empleando como control diferentes grupos de estaciones. Por último, se aplican los métodos de interpolación sobre la red de estaciones SIAR incluyendo un grupo de control adicional formado por las estaciones de la red de la Agencia Estatal de Meteorología, determinando el método de interpolación más preciso para la zona de trabajo. En este sentido, los resultados más sobresalientes han sido:

- El método de interpolación que mejor se ajusta a la zona de estudio es el Kriging Universal, resultado coherente con las características geográficas de Castilla y León que suponen un importante gradiente en los valores de radiación en la dirección del Duero.
- El Kriging gaussiano no es adecuado para la zona en estudio, el error obtenido con todos los métodos de interpolación es considerablemente mayor que el resto.
- La distribución de las estaciones cercanas al punto de interpolación es determinante para el valor obtenido.
- Los errores obtenidos son inferiores al 7 % salvo en el Kriging gaussiano.

5. Para obtener los valores de irradiación sobre superficie vertical en las cuatro orientaciones norte, sur, este y oeste se han utilizado los datos experimentales recogidos en la estación de medida de la E.P.S. Esta instalación ha sido completamente implementada durante la realización de esta tesis y constituye parte del trabajo experimental realizado. Se han empleado los valores de irradiación global, difusa y directa sobre superficie horizontal para obtener los valores de irradiación vertical teóricos, aplicando los cuatro modelos seleccionados que luego se han

comparado con los medidos por los cuatro piranómetros verticales. Los resultados han sido los siguientes:

- En la orientación Norte: para valores medios anuales diarios de irradiación, el modelo que mejor se comporta es el Hay y el peor el Circunsolar. Para intervalos de valores altos de irradiación, el Hay se comporta bien como en los meses centrales del año (abril – agosto). Para intervalos de baja irradiación el modelo que peor se ajusta es el Circunsolar y para los meses de marzo y septiembre el Isotrópico.
- En la orientación Sur: para valores anuales el mejor es el Isotrópico y el peor el Circunsolar. Desde el punto de vista mensual, el peor modelo todos los meses es el Circunsolar, y dependiendo de la época se alternan el Isotrópico (marzo – agosto) y el Klucher el resto de los meses. Analizando la irradiación por intervalos, los dos mejores son alternativamente el Isotrópico y el Klucher, con valores muy próximos.
- En la orientación Este: para valores anuales el mejor comportamiento lo presenta el modelo Circunsolar y el peor el Hay. Para valores mensuales, el modelo Circunsolar se comporta muy bien en los meses de marzo a octubre y es el peor en los meses restantes. Desde el punto de vista de intervalos, todos los modelos disminuyen el error a medida que aumenta el valor de irradiación, dando los errores menores el Circunsolar.
- En la orientación Oeste: para valores anuales el mejor es el Circunsolar y el peor es Hay. En valores mensuales durante los meses de marzo a octubre el mejor también es el Circunsolar y en esos meses el peor es el Hay. Por intervalos, los mejores modelos son el Circunsolar y el Klucher, para valores bajos de irradiación se comporta mejor el Klucher y para altos el Circunsolar.

Futura líneas abiertas de investigación:

1. Aplicar los resultados a una instalación de investigación ubicada en Astudillo (Palencia) con paneles verticales en las cuatro orientaciones.
2. Aplicar los resultados a tres estaciones próximas rurales, situadas en Astudillo, Torquemada y Herrera de Valdecañas, comparando el valor interpolado de irradiación con la producción real a diferentes inclinaciones.
3. Ampliar el periodo de medidas experimentales para comprobar la validez de las conclusiones obtenidas, considerando otros modelos de radiación sobre superficie inclinada.
4. Estudiar la influencia de otros parámetros climáticos como temperatura, presión, humedad y velocidad de viento en la irradiación sobre plano horizontal con ArcGis.
5. Aplicar los resultados obtenidos a valores de iluminancia

BIBLIOGRAFÍA

- K. Allenbach et al., (2015) "Black Sea beaches vulnerability to sea level rise". *Environmental Science & Policy*, vol. 46, p. 95-109.
- H. Alsamamra, J. A. Ruiz-Arias, D. Pozo-Vázquez and J. Tovar-Pescador (2009). "A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain." *Agricultural and Forest Meteorology* 149(8): 1343-1357.
- O. Antoni, J. Križan, A. Marki and D. Bukovec (2001). "Spatio-temporal interpolation of climatic variables over large region of complex terrain using neural networks." *Ecological Modelling* 138(1-3): 255-263.
- H. Apaydin, F. Kemal Sonmez and Y. E. Yildirim (2004). "Spatial interpolation techniques for climate data in the GAP region in Turkey." *Climate Research* 28(1): 31-40.
- J. Appelbaum, (2016). Bifacial photovoltaic panels field. *Renewable Energy*, 2016, vol. 85, p. 338-343
- V. Badescu (2002). "3D isotropic approximation for solar diffuse irradiance on tilted surfaces." *Renewable Energy* 26(2): 221-233.
- V. Badescu, C. A. Gueymard, S. Cheval, C. Oprea, M. Baciú, A. Dumitrescu, F. Iacobescu, I. Milos and C. Rada (2012). "Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models." *Renewable and Sustainable Energy Reviews* 16(3): 1636-1656.
- V. Badescu, C. A. Gueymard, S. Cheval, C. Oprea, M. Baciú, A. Dumitrescu, F. Iacobescu, I. Milos and C. Rada (2013). "Accuracy and sensitivity analysis for 54 models of computing hourly diffuse solar irradiation on clear sky." *Theoretical and Applied Climatology* 111(3-4): 379-399.

- W. Bahr, (2014). "A comprehensive assessment methodology of the building integrated photovoltaic blind system". *Energy and Buildings*, vol. 82, p. 703-708.
- K. Bakirci (2012). "General models for optimum tilt angles of solar panels: Turkey case study." *Renewable and Sustainable Energy Reviews* 16(8): 6149-6159.
- I. Baklouti, Z. Driss and M. S. Abid (2012). Estimation of solar radiation on horizontal and inclined surfaces in Sfax, TUNISIA. 2012 1st International Conference on Renewable Energies and Vehicular Technology, REVET 2012, Nabeul.
- R. Beatson, W. Ong and I. Rychkov (2014). "Faster fast evaluation of thin plate splines in two dimensions." *Journal of Computational and Applied Mathematics* 261: 201-212.
- E. F. Cassiraga (1999). "Incorporación de información blanda para la cuantificación de la incertidumbre: aplicación a la hidrogeología". Tesis UPV.
- M. Chelbi, Y. Gagnon, J. Waewsak, (2015). "Solar radiation mapping using sunshine duration-based models and interpolation techniques: Application to Tunisia". *Energy Conversion and Management*, vol. 101, p. 203-215.
- R. Chen, et al.(2006). "Trends of the global radiation and sunshine hours in 1961–1998 and their relationships in China." *Energy Conversion and Management*, vol. 47, no 18, p. 2859-2866.
- S. Chirarattananon, P. Rukkwansuk, P. Chaiwiwatworakul and P. Pakdeepol (2007). "Evaluation of vertical illuminance and irradiance models against data from north Bangkok." *Building and Environment* 42(11): 3894-3904.
- U. Chung and J. I. Yun (2004). "Solar irradiance-corrected spatial interpolation of hourly temperature in complex terrain." *Agricultural and Forest Meteorology* 126(1-2): 129-139.
- S. D. Conte and C. De Boor (1972). *Elementary numerical analysis: an algorithmic approach*, McGraw-Hill.

- M. Cucumo, A. De Rosa, V. Ferraro, D. Kaliakatsos and V. Marinelli (2007). "Experimental testing of models for the estimation of hourly solar radiation on vertical surfaces at Arcavacata di Rende." *Solar Energy* 81(5): 692-695.
- A. Das, J. Park, J. Park, (2015). "Estimation of available global solar radiation using sunshine duration over South Korea". *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 134, p. 22-29.
- C. Demain, M. Journée and C. Bertrand (2013). "Evaluation of different models to estimate the global solar radiation on inclined surfaces." *Renewable Energy* 50: 710-721.
- H. Demirhan, T. Menteş and M. Atilla (2013). "Statistical comparison of global solar radiation estimation models over Turkey." *Energy Conversion and Management* 68: 141-148.
- J. A. Duffie and W. A. Beckman (2013). *Solar Engineering of Thermal Processes: Fourth Edition*.
- A. A. El-Sebaili, F. S. Al-Hazmi, A. A. Al-Ghamdi and S. J. Yaghmour (2010). "Global, direct and diffuse solar radiation on horizontal and tilted surfaces in Jeddah, Saudi Arabia." *Applied Energy* 87(2): 568-576.
- M. M. Elsayed (1989). "Optimum orientation of absorber plates." *Solar Energy* 42(2): 89-102.
- R. Enriquez, L. Zarzalejo, M. J. Jiménez and M. R. Heras (2012). "Ground reflectance estimation by means of horizontal and vertical radiation measurements." *Solar Energy* 86(11): 3216-3226.
- D. Erbs, S. Klein and J. Duffie (1982). "Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation." *Solar energy* 28(4): 293-302.
- S. Erdogan (2009). "A comparison of interpolation methods for producing digital elevation models at the field scale." *Earth surface processes and landforms* 34(3): 366.

R. A. Fisher (1925). *Statistical Methods For Research Workers*, Cosmo Publications. ISBN (9788130701332).

R. Franke (1982). "Scattered data interpolation: Tests of some methods." *Mathematics of computation* 38(157): 181-200.

R. Franke (1985). "Thin plate splines with tension." *Computer Aided Geometric Design* 2(1): 87-95.

L. S. Gandin and R. Hardin (1965). *Objective analysis of meteorological fields*, Israel program for scientific translations Jerusalem.

P. Goovaerts (2000). "Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall." *Journal of Hydrology* 228(1-2): 113-129.

C. Gueymard (1987). "An anisotropic solar irradiance model for tilted surfaces and its comparison with selected engineering algorithms." *Solar Energy* 38(5): 367-386.

C. A. Gueymard and D. R. Myers (2009). "Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling." *Solar Energy* 83(2): 171-185.

M. Güler (2014). "A comparison of different interpolation methods using the geographical information system for the production of reference evapotranspiration maps in Turkey." *Journal of the Meteorological Society of Japan* 92(3): 227-240.

F. V. Gutierrez-Corea, M. A. Manso-Callejo, M. P. Moreno-Regidor and J. Velasco-Gómez (2014). "Spatial estimation of sub-hour global horizontal irradiance based on official observations and remote sensors." *Sensors (Switzerland)* 14(4): 6758-6787.

P. A. Hancock and M. F. Hutchinson (2006). "Spatial interpolation of large climate data sets using bivariate thin plate smoothing splines." *Environmental Modelling and Software* 21(12): 1684-1694.

J. E. Hay and J. A. Davies (1980). Calculation of the solar radiation incident on an inclined surface. Proc. of First Canadian Solar Radiation Data Workshop (Eds: JE Hay and TK Won), Ministry of Supply and Services Canada.

J. E. Hay and D. C. McKay (1985). "Estimating solar irradiance on inclined surfaces: A review and assessment of methodologies." *International Journal of Solar Energy* 3(3): 203-240.

R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones and A. Jarvis (2005). "Very high resolution interpolated climate surfaces for global land areas." *International Journal of Climatology* 25(15): 1965-1978.

W. W. Hines and D. C. Montgomery (1990). *Probability and statistics in engineering and management science*, J. Wiley. ISBN (9780471600909).

R. Hoodonk et al. (2014). "Opposite latitudinal gradients in projected ocean acidification and bleaching impacts on coral reefs". *Global change biology*, vol. 20, no 1, p. 103-112.

H. C. Hottel and B. B. Woertz (1942). "Performance of flat-plate solar-heat collectors." *Trans. ASME* 64(4): 91-104.

M. F. Hutchinson, (2000). ANUDEM Software. Centre for Resource and Environmental Studies, Australian National University, Canberra.

M. F. Hutchinson,(2004). ANUSPLIN Version 4.3. Centre for Resource and Environmental Studies, Australian National University.

Hutchinson, M. F. (1989). "A new procedure for gridding elevation and stream line data with automatic removal of spurious pits". *Journal of Hydrology*, 106: 211–232.

P. Ineichen (2008). "Comparison and validation of three global-to-beam irradiance models against ground measurements." *Solar Energy* 82(6): 501-512.

M. Iqbal (1983). *An introduction to solar radiation*, Academic Press. ISBN (9780123737526).

S. J. Jeffrey, J. O. Carter, K. B. Moodie and A. R. Beswick (2001). "Using spatial interpolation to construct a comprehensive archive of Australian climate data." *Environmental Modelling and Software* 16(4): 309-330.

H. D. Kambezidis, B. E. Psiloglou and C. Gueymard (1994). "Measurements and models for total solar irradiance on inclined surface in Athens, Greece." *Solar Energy* 53(2): 177-185.

A. Karim and B. S. M. Singh (2014). "Global Solar Radiation Modeling Using Polynomial Fitting." *Applied Mathematical Sciences* 8(8): 367-378.

Z. Kebaili Bargaoui and A. Chebbi (2009). "Comparison of two kriging interpolation methods applied to spatiotemporal rainfall." *Journal of Hydrology* 365(1-2): 56-73.

T. Khatib, A. Mohamed, K. Sopian and M. Mahmoud (2012). "Assessment of artificial neural networks for hourly solar radiation prediction." *International journal of Photoenergy* 2012.

T. M. Klucher (1979). "Evaluation of models to predict insolation on tilted surfaces." *Solar Energy* 23(2): 111-114.

P. S. Koronakis (1986). "On the choice of the angle of tilt for south facing solar collectors in the Athens basin area." *Solar Energy* 36(3): 217-225.

D. Krige (1951). "A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand." *Journal of the Chemical, Metallurgical and Mining Society of South Africa* 52(6): 119-139.

R. Kumar, L. Umanand (2005), "Estimation of global radiation using clearness index model for sizing photovoltaic system." *Renewable Energy*, vol. 30, no 15, p. 2221-2233.

N. S.-N. Lam (1983). "Spatial interpolation methods: a review." *The American Cartographer* 10(2): 129-150.

S.K. Lauvset et al. (2015). "Trends and drivers in global surface ocean pH over the past 3 decades". *Biogeosciences*, vol. 12, no 5, p. 1285-1298.

D. H. Li, J. C. Lam and C. C. Lau (2002). "A new approach for predicting vertical global solar irradiance." *Renewable Energy* 25(4): 591-606.

F. Li et al. (2014) "Research on the optimum design of building photovoltaic based on the radiation model evaluation on vertical surface in Beijing". *Electronics and Application Conference and Exposition (PEAC), 2014 International*. IEEE, 2014. p. 1212-1216.

S. Li, et al. (2015). "Evaluation of the threat from sea-level rise to the mangrove ecosystems in Tieshangang Bay, southern China". *Ocean & Coastal Management*, vol. 109, p. 1-8.

B. Y. H. Liu and R. C. Jordan (1962). "Daily insolation on surfaces tilted toward the equator." *ASHRAE Journal* 3(10).

G. López, M. A. Rubio and F. J. Batlles (2000). "Estimation of hourly direct normal from measured global solar irradiance in Spain." *Renewable Energy* 21(2): 175-186.

P. Loutzenhiser, H. Manz, C. Feltsmann, P. Strachan, T. Frank and G. Maxwell (2007). "Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation." *Solar Energy* 81(2): 254-267.

E. L. Maxwell (1987). *A quasi-physical model for converting hourly global horizontal to direct normal insolation*, Solar Energy Research Inst., Golden, CO (USA).

L. Mitáš and H. Mitášová (1988). "General variational approach to the interpolation problem." *Computers & Mathematics with Applications* 16(12): 983-992.

S.H. Monger, et al.(2016). "Applying the kriging method to predicting irradiance variability at a potential PV power plant". *Renewable Energy*, vol. 86, p. 602-610.

T. Muneer (1997). *Solar Radiation and Daylight Models for the Energy Efficient Design of Buildings*.

- I. A. Nalder and R. W. Wein (1998). "Spatial interpolation of climatic Normals: Test of a new method in the Canadian boreal forest." *Agricultural and Forest Meteorology* 92(4): 211-225.
- N. K. Newlands, A. Davidson, A. Howard and H. Hill (2011). "Validation and inter-comparison of three methodologies for interpolating daily precipitation and temperature across Canada." *Environmetrics* 22(2): 205-223.
- G. Notton, P. Poggi and C. Cristofari (2006). "Predicting hourly solar irradiances on inclined surfaces based on the horizontal measurements: Performances of the association of well-known mathematical models." *Energy Conversion and Management* 47(13-14): 1816-1829.
- OMM (2010). "Guía de Instrumentos y Métodos de Observación Meteorológicos". OMM-Nº 8.
- OMM (2013). "Guía del Sistema Mundial de Observación". OMM-Nº 488
- M. S. Okundamiya and A. N. Nzeako (2011). "Influence of orientation on the performance of a photovoltaic conversion system in Nigeria." *Research Journal of Applied Sciences, Engineering and Technology* 3(12): 1384-1390.
- K. Orehounig, S. Dervishi and A. Mahdavi (2014). "Computational derivation of irradiance on building surfaces: An empirically-based model comparison." *Renewable Energy* 71: 185-192.
- J. F. Orgill and K. G. T. Hollands (1977). "Correlation equation for hourly diffuse radiation on a horizontal surface." *Solar Energy* 19(4): 357-359.
- C.K. Pandey, A.K. Katiyar (2014). "Hourly solar radiation on inclined surfaces". *Sustainable Energy Technologies and Assessments*, vol. 6, p. 86-92.
- P. J. Pérez-Higueras, P. Rodrigo, E. F. Fernández, F. Almonacid and L. Hontoria (2012). "A simplified method for estimating direct normal solar irradiation from global horizontal irradiation useful for CPV applications." *Renewable and Sustainable Energy Reviews* 16(8): 5529-5534.

R. Perez, P. Ineichen, E. Maxwell, R. Seals and A. Zelenka (1991). Dynamic models for hourly global-to-direct irradiance conversion. 1991 Solar world congress: Proceedings of the biennial congress of the international solar energy society.

R. Perez, P. Ineichen, K. Moore, M. Kmiecik, C. Chain, R. George and F. Vignola (2002). "A new operational model for satellite-derived irradiances: Description and validation." *Solar Energy* 73(5): 307-317.

R. Perez, P. Ineichen, R. Seals, J. Michalsky and R. Stewart (1990). "Modeling daylight availability and irradiance components from direct and global irradiance." *Solar Energy* 44(5): 271-289.

R. Perez, J. Scott and R. Stewart (1983). "An anisotropic model for diffuse radiation incident on slopes of different orientations, and possible applications to CPCs." *Proceedings of ASES*: 883-888.

R. Perez, R. Seals, P. Ineichen, R. Stewart and D. Menicucci (1987). "A new simplified version of the Perez diffuse irradiance model for tilted surfaces." *Solar energy* 39(3): 221-231.

R. Perez, R. Stewart, C. Arbogast, R. Seals and J. Scott (1986). "An anisotropic hourly diffuse radiation model for sloping surfaces: description, performance validation, site dependency evaluation." *Solar energy* 36(6): 481-497.

J. Polo (2015). "Solar global Horizontal and direct normal irradiation Maps in Spain derived from geostationary Satellites". *Journal of Atmospheric and Solar-Terrestrial Physics*.

D. T. Price, D. W. McKenney, I. A. Nalder, M. F. Hutchinson and J. L. Kesteven (2000). "A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data." *Agricultural and Forest Meteorology* 101(2-3): 81-94.

C. Raichijk, F. Taddei (2012). "Estudio comparativo de modelos empíricos de estimación de la radiación solar directa horaria." *Avances en Energías Renovables y Medio Ambiente*, 16.D.

T. Reindl, W. A. Beckman and J. A. Duffie (1990a). "Diffuse fraction correlations." *Solar Energy* 45(1): 1-7.

D. T. Reindl, W. A. Beckman and J. A. Duffie (1990b). "Evaluation of hourly tilted surface radiation models." *Solar Energy* 45(1): 9-17.

A. Riccio (2005). "A Bayesian approach for the spatiotemporal interpolation of environmental data." *Monthly weather review* 133(2): 430-440.

T. Robinson and G. Metternicht (2005). "Comparing the performance of techniques to improve the quality of yield maps." *Agricultural Systems* 85(1): 19-41.

J,A. Ruiz-Arias, et al.(2015). "Optimal combination of gridded and ground-observed solar radiation data for regional solar resource assessment". *Solar Energy*, vol. 112, p. 411-424.

A. Sanchez-Lorenzo, J. Calbó, M. Wild, (2013). "Global and diffuse solar radiation in Spain: Building a homogeneous dataset and assessing their trends". *Global and Planetary Change*, vol. 100, p. 343-352.

G. Saraf and F. A. W. Hamad (1988). "Optimum tilt angle for a flat plate solar collector." *Energy Conversion and Management* 28(2): 185-191.

Y. Sasaki, A. Texas, M. U. System, O. Department of and Meteorology (1960). *An objective analysis for determining initial conditions for the primitive equations*. College Station, Tex., Texas A & M University, Dept. of Oceanography.

Z. Şen, A. Öztopal and A. D. Şahin (2004). "Solar Irradiation Estimation from Sunshine Duration by Geno-Fuzzy Partial Approach." *Energy Sources* 26(4): 377-386.

Z. Şen and A. D. Şahin (2001). "Spatial interpolation and estimation of solar irradiation by cumulative semivariograms." *Solar Energy* 71(1): 11-22.

R. Sibson (1981). "A brief description of natural neighbour interpolation." *Interpreting multivariate data* 21: 21-36.

A. Skartveit and J. Asle Olseth (1986). "Modelling slope irradiance at high latitudes." *Solar Energy* 36(4): 333-344.

S. Smith, D. Holland and P. Longley (2005). "Quantifying interpolation errors in urban airborne laser scanning models." *Geographical Analysis* 37(2): 200-224.

G. W. Snedecor and W. G. Cochran (1991). *Statistical Methods*, Wiley. ISBN (9780813815619).

K. Stahl, R. D. Moore, J. A. Floyer, M. G. Asplin and I. G. McKendry (2006). "Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density." *Agricultural and Forest Meteorology* 139(3-4): 224-236.

C. Stanciu and D. Stanciu (2014). "Optimum tilt angle for flat plate collectors all over the World - A declination dependence formula and comparisons of three solar radiation models." *Energy Conversion and Management* 81: 133-143.

M.T. Taye, P. Willems, P. Block (2015). "Implications of climate change on hydrological extremes in the Blue Nile basin: A review". *Journal of Hydrology: Regional Studies*, vol. 4, p. 280-293.

R. C. Temps and K. L. Coulson (1977). "Solar radiation incident upon slopes of different orientations." *Solar Energy* 19(2): 179-184.

D. Terzopoulos (1988). "The computation of visible-surface representations." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 10(4): 417-438.

Y. Q. Tian, R. J. Davies-Colley, P. Gong and B. W. Thorrold (2001). "Estimating solar radiation on slopes of arbitrary aspect." *Agricultural and Forest Meteorology* 109(1): 67-74.

H. B. Tolabi, M. H. Moradi and S. B. M. Ayob (2014). "A review on classification and comparison of different models in solar radiation estimation." *International Journal of Energy Research* 38(6): 689-701.

M. Utrillas, J. Martinez-Lozano and A. Casanovas (1991). "Evaluation of models for estimating solar irradiation on vertical surfaces at Valencia, Spain." *Solar Energy* 47(3): 223-229.

G. Wahba and J. Wendelberger (1980). "Some new mathematical methods for variational objective analysis using splines and cross validation." *Monthly weather review* 108(8): 1122-1143.

Y. Wang, S. Zhou, H. Huo, (2014). "Cost and CO₂ reductions of solar photovoltaic power generation in China: Perspectives for 2020". *Renewable and Sustainable Energy Reviews*, vol. 39, p. 370-380.

D. Watson and G. Philip (1985). "Comment on "a nonlinear empirical prescription for simultaneously interpolating and smoothing contours over an irregular grid" by F. Duggan." *Computer methods in applied mechanics and engineering* 50(2): 195-198.

A.P. Williams et al. (2012). "Recent summer precipitation trends in the Greater Horn of Africa and the emerging role of Indian Ocean sea surface temperature". *Climate Dynamics*, vol. 39, no 9-10, p. 2307-2328.

J. Wilcox, D. Makowski, (2014). "A meta-analysis of the predicted effects of climate change on wheat yields using simulation studies". *Field Crops Research*, vol. 156, p. 180-190.

C. J. Willmott and K. Matsuura (2005). "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance." *Climate research* 30(1): 79.

C. J. Willmott, K. Matsuura and S. M. Robeson (2009). "Ambiguities inherent in sums-of-squares-based error statistics." *Atmospheric Environment* 43(3): 749-752.

A. M. Wilson and J. A. Silander (2014). "Estimating uncertainty in daily weather interpolations: A Bayesian framework for developing climate surfaces." *International Journal of Climatology* 34(8): 2573-2584.

W. Wu, X. P. Tang, C. Yang, N. J. Guo and H. B. Liu (2013). "Spatial estimation of monthly mean daily sunshine hours and solar radiation across mainland China." *Renewable Energy* 57: 546-553.

D. Yan et al. (2015). " Hydrological response to climate change: The Pearl River, China under different RCP scenarios". *Journal of Hydrology: Regional Studies*, vol. 4, p. 228-245.

D. Yang, Z. Dong, A. Nobre, Y. S. Khoo, P. Jirutitijaroen and W. M. Walsh (2013). "Evaluation of transposition and decomposition models for converting global solar irradiance from tilted surface to horizontal in tropical regions." *Solar Energy* 97: 369-387.

D. Yang, Z. Dong, T. Reindl, P. Jirutitijaroen and W. M. Walsh (2014). "Solar irradiance forecasting using spatio-temporal empirical kriging and vector autoregressive models with parameter shrinkage." *Solar Energy* 103: 550-562.

T. Zhang, P. W. Stackhouse, Jr., W. S. Chandler and D. J. Westberg (2014). "Application of a global-to-beam irradiance model to the NASA GEWEX SRB dataset: An extension of the NASA Surface meteorology and Solar Energy datasets." *Solar Energy* 110: 117-131.

Direcciones web consultadas:

<http://www.europarl.europa.eu/portal/es>

<http://datos.bancomundial.org/indicador/SP.URB.TOTL.IN.ZS>

[http://www.eia.gov/forecasts/ieo/pdf/0484\(2013\).pdf](http://www.eia.gov/forecasts/ieo/pdf/0484(2013).pdf)

<http://ec.europa.eu/eurostat/data/database>

<http://www.ipcc.ch/>

<http://www2.meteogalicia.es/galego/observacion/estacions/listaEstacions.asp>

<http://www.fae.es/nueva/controler.php?id=48&idDetalle=3382&idIdioma=ES>

<http://www.airecantabria.com>

<http://www.euskalmet.euskadi.net>

<https://www.larioja.org/npRioja/default/defaultpage.jsp?idtab=442821>

<http://eportal.magrama.gob.es/websiar/Inicio.aspx>

<http://helios.ies-def.upm.es/consulta.aspx>

<http://redarexplus.gobex.es/RedarexPlus/index.php?modulo=agrometeorologia&camino=Agrometeorolog%EDa>

<http://www.lneg.pt/servicos/328/2263/>

http://www.inforiego.org/opencms/opencms/info_meteo/index.html

Este trabajo de tesis doctoral ha sido realizado en el marco del Proyecto de Investigación ENE 2011-27511 del MINISTERIO DE CIENCIA E INNOVACIÓN y del Proyecto BU 358A12-2 de la JUNTA DE CASTILLA Y LEÓN.

Los datos de radiación de la AEMET han sido proporcionados por la Agencia Estatal de Meteorología.