



Universidad Internacional de La Rioja
Escuela Superior de Ingeniería y Tecnología

Máster Universitario en Astrofísica y Técnicas de Observación
en Astronomía

Simulador de espectros estelares a partir de parámetros estelares por regresión simbólica

Trabajo fin de estudio presentado por:	Alberto Gutiérrez Morán
Tipo de trabajo:	Investigación en astronomía y astrofísica
Línea de trabajo:	Desarrollo software/matemático para el análisis de datos astronómicos Evolución estelar
Director:	Guillermo Torralba Elipe
Fecha:	14/07/2025

Agradecimientos

El trabajo aquí presente no hubiera sido posible sin la ayuda del director del mismo, Guillermo, a quién quiero agradecer no solo por proponer la idea inicial, sino por guiarme y aconsejarme a lo largo de todo el desarrollo, el cual comencé con una noción muy limitada sobre el campo que abarca y para el que me hubiera costado mucho encontrar la dirección correcta sin su ayuda. Por supuesto también quiero agradecer a mi familia, a mis primos, a mis hermanas y a mis padres, con los que sin su apoyo no hubiera podido completar el trabajo ni hubiera tenido el ánimo para invertir todo el tiempo que requirió.

Este trabajo se ha desarrollado en el marco del proyecto Astronomía no supervisada e interpretable con técnicas avanzadas (ANITA/PP-2024-08), que se financia mediante la convocatoria “Financiación de Proyectos Precompetitivos de Investigación UNIR 2024”. Dicho proyecto ha financiado parcialmente el desarrollo de este trabajo.

Resumen

En el presente estudio se pretende probar la efectividad de la construcción de espectros estelares mediante técnicas de regresión simbólica. El método predominante para generar espectros sintéticos es siguiendo modelos atmosféricos, aquí se evalúa la eficiencia de la programación genética para hallar ecuaciones que relacionen los parámetros físicos y abundancias químicas de una estrella con su flujo espectral, sin ayuda de ningún modelo ni elemento adicional. Estas ecuaciones funcionarán como modelo, permitiendo producir espectros mediante su aplicación.

En concreto se generan dos ecuaciones dependientes de tres parámetros estelares que permiten computar una distribución gaussiana por cada posible línea de absorción en un rango espectral que abarca desde los 846 nm a 870 nm, rango coincidente con el de los espectros de alta resolución RVS de la misión Gaia, los cuales se usan para validar el algoritmo. El rango espectral escogido permite considerar el espectro como continuo, por lo que el resultado final es formado ajustando cada línea de absorción existente en ese rango y añadiéndola al continuo.

Las funciones finales permiten sintetizar espectros con un error relativamente grande con respecto a cómo deberían ser, con ciertas líneas de absorción perdidas o mal ajustadas pero que sin embargo muestran las más predominantes correctamente, por lo que al menos permiten atisbar los elementos químicos que contribuyen a la absorción con fidelidad.

El ciclo de desarrollo creado se puede interpretar como una primera aproximación a la generación de espectros sintéticos mediante el uso de regresión simbólica, que, con un desarrollo más exhaustivo, y con mayores recursos y tiempo, puede producir resultados más satisfactorios y aplicables en el campo de la astronomía.

Palabras clave:

Regresión simbólica, espectros estelares, simulador, aprendizaje automático

Abstract

The present study aims to test the performance of symbolic regression techniques in computing a synthetic stellar spectrum. The main used method for synthetic spectrum generation is by following atmospheric models, this study evaluates the efficiency of genetic programming in discovering equations that link astrophysical parameters and chemical abundances of a star with its spectral flux, without any external help or model. These functions will work as the model, allowing to compute synthetic spectra by resolving them.

Two functions depending on three stellar parameters that provide a gaussian distribution will be generated for each possible absorption line in a specific spectral range, spanning from 846 nm to 870 nm, the same range as the high resolution RVS spectra from the mission Gaia, which are used to validate the algorithm. The chosen spectral range allows to consider the spectrum as a continuum, so the final result is created by computing each absorption line existing in the spectrum range and adding it to the continuum.

The final functions result in a spectrum synthesis with a relatively large error in comparison with how it should be, having some absorption lines missing or with a bad fitting, but which nevertheless shows the most predominant ones correctly, so it allows to derive the chemical elements that contribute to the absorption faithfully.

The development cycle created can be seen as a first approach on spectrum synthesis using symbolic regression, that with a more exhaustive development, with more resources and time, can produce more realistic results applicable in the astronomy field.

Keywords:

Symbolic regression, stellar spectra, simulator, machine learning

Índice de contenidos

1. Introducción	1
1.1. Estructura de la memoria	2
2. Estado del arte y objetivos	4
2.1. Regresión simbólica	4
2.2. Espectros y parámetros	5
2.3. Objetivos.....	7
3. Desarrollo	9
3.1. Datos.....	9
3.2. Aproximación inicial	14
3.2.1. Generación de espectro sintético y tratamiento de los datos.....	14
3.2.2. Selección de líneas atómicas	19
3.2.3. Normalización del espectro y ajuste de líneas atómicas	20
3.2.4. Regresión simbólica	39
3.2.5. Síntesis del espectro	44
3.3. Entrenamiento.....	46
3.3.1. Primer entrenamiento	46
3.3.2. Segundo entrenamiento.....	48
3.3.3. Tercer entrenamiento	52
4. Resultados	59
4.1. Funciones.....	59
4.2. Validación	61
5. Conclusiones.....	67

5.1.	Conclusiones.....	67
5.2.	Dificultades encontradas y posibles líneas futuras de investigación	68
	Referencias bibliográficas.....	70
Anexo A.	Lista completa de funciones generadas	73
	Índice de siglas y acrónimos	94

Índice de figuras

Figura 1. <i>Anchura equivalente de una línea de absorción</i>	6
Figura 2. <i>Diagrama de Kiel para los parámetros de GSP Spec sobre espectros RVS</i>	13
Figura 3. <i>Espectro inicial generado con iSpec</i>	18
Figura 4. <i>Ajuste inicial fallido de Gauss</i>	21
Figura 5. <i>Espectro inicial generado con iSpec normalizado</i>	22
Figura 6. <i>Espectro normalizado con líneas de absorción tras primera selección</i>	23
Figura 7. <i>Límites de la línea de absorción del Si 1 en espectro sintético</i>	25
Figura 8. <i>Espectro normalizado con líneas de absorción y límites tras segunda selección</i>	25
Figura 9. <i>Frecuencias en línea de absorción descartada</i>	26
Figura 10. <i>Espectro normalizado con líneas de absorción y límites tras tercera selección</i>	28
Figura 11. <i>Límites erróneos de otra línea de absorción del Si 1 en espectro sintético</i>	29
Figura 12. <i>Límites corregidos de la otra línea de absorción del Si 1 en espectro sintético</i>	30
Figura 13. <i>Espectro normalizado con líneas de absorción y límites tras cuarta selección</i>	30
Figura 14. <i>Primer ajuste de gaussianas en espectro de iSpec</i>	32
Figura 15. <i>Segundo ajuste de gaussianas con offset en espectro de iSpec</i>	33
Figura 16. <i>Ajuste de gaussianas completo en espectro de iSpec</i>	34
Figura 17. <i>Ajuste de gaussianas en espectro sintético de estrella A7</i>	36
Figura 18. <i>Espectro inicial sintetizado sin regresión contra gaussianas halladas</i>	38
Figura 19. <i>Segundo espectro inicial sintetizado sin regresión</i>	38
Figura 20. <i>Espectro inicial sintetizado con regresión</i>	45
Figura 21. <i>Síntesis con primera generación de funciones por regresión</i>	48
Figura 22. <i>Síntesis con segunda generación de funciones por regresión</i>	49

Figura 23. <i>Gaussianas generadas con el segundo set de funciones de la regresión</i>	50
Figura 24. <i>Síntesis filtrada con segunda generación de funciones por regresión</i>	51
Figura 25. <i>Comparación de síntesis para segunda generación de funciones</i>	52
Figura 26. <i>Errores de síntesis para todos los espectros de entrenamiento</i>	54
Figura 27. <i>Espectro sintético erróneo de iSpec y ajuste</i>	55
Figura 28. <i>Comparación de síntesis para tercera generación de funciones</i>	56
Figura 29. <i>Comparación de síntesis con líneas verticales en tercera generación de funciones</i>	57
Figura 30. <i>Errores de síntesis para los espectros de validación de Gaia</i>	62
Figura 31. <i>Sintetizaciones de validación para 2 espectros de Gaia</i>	63
Figura 32. <i>Sintetizaciones de validación con desviación corregida para 2 espectros de Gaia</i> .	64
Figura 33. <i>Errores de síntesis corregida para los espectros de validación de Gaia</i>	65
Figura 34. <i>Mejor síntesis de validación obtenida para un espectro de Gaia</i>	65

Índice de tablas

Tabla 1. <i>Funciones finales para las líneas más relevantes</i>	59
--	----

1. Introducción

El presente Trabajo de Fin de Máster pretende probar un enfoque alternativo y no usado hasta la fecha para descifrar la relación intrínseca entre los parámetros atmosféricos de una estrella y su flujo espectral. El fuerte aumento en la última década de observaciones de espectrógrafos de alta resolución ha generado enormes catálogos, como el de Gaia ESO (Gilmore et al., 2012), APOGEE (Abdurro'uf, 2022) y LAMOST (Li et al., 2022), cada uno de ellos con millones de espectros asociados, lo que ha resultado en la necesidad de herramientas automáticas con la capacidad de inferir parámetros estelares a partir de sus espectros y así facilitar la clasificación. La mayor parte de las herramientas existentes por tanto están desarrolladas con objeto de realizar la operación inversa a la que se ejecuta en este trabajo, inferir parámetros a partir del espectro. Esta investigación no solo probará la generación de espectros a partir de los parámetros, sino que lo hará con una técnica muy poco explorada, la regresión simbólica.

La regresión simbólica es una técnica de aprendizaje automático no muy popular en comparación con otras alternativas del dominio. Este mecanismo se encarga, simplifícadamente, de crear expresiones matemáticas que mejor encajen una serie de variables de entrada con una de salida (Michael Affenzeller, 2024).

Al contrario que otros mecanismos más populares como puede ser la regresión lineal, la regresión simbólica no usa modelos predefinidos, lo que conlleva que se puedan encontrar relaciones más complejas, que quizás no sean lineales o predecibles por métodos tradicionales, a costa de necesitar un mayor tiempo de ejecución. Teniendo esto en cuenta son claros los beneficios que puede tener en campos como la física: de las ecuaciones generadas se podrán obtener nuevas perspectivas sobre los procesos fundamentales que hay en juego para la obtención de las variables de salida a partir de las de entrada.

La regresión simbólica tiene su origen en la programación genética, lo que significa que para descifrar las funciones se hace uso de algoritmos evolutivos, inspirados comúnmente en selección natural y genética. Iterativamente se van optimizando las expresiones matemáticas basándose en criterios como la precisión o la complejidad, con fin de encontrar el modelo más representativo.

La técnica original, sin modelos predefinidos como en otros métodos de regresión que usan aprendizaje automático, hace uso de fuerza bruta para descifrar las funciones: se prueban todas las posibles expresiones simbólicas, o bien dentro de una clase o en orden ascendente de complejidad. Este enfoque será capaz de hallar multitud de ecuaciones, pero si la complejidad continúa ascendiendo teóricamente el algoritmo podría tardar cientos de años en hallar la solución (Udrescu & Tegmark, 2020). Debido al enfoque adoptado esta técnica ha sido eclipsada en los últimos años con respecto a alternativas que hacen uso de redes neuronales y mecanismos más eficientes.

La relación entre parámetros atmosféricos estelares y el flujo espectral fue descrita hace años y plasmada en los denominados modelos atmosféricos, de los que hoy hacen uso las herramientas más destacadas para inferir parámetros estelares o para generar espectros sintéticos, como GSP Spec (Recio-Blanco et al., 2023), TurboSpectrum (Alvarez & Plez, 1998) o iSpec (Blanco-Cuaresma et al., 2014). El uso de la regresión simbólica descarta la posibilidad de utilizar dichos modelos, ya que la esencia del algoritmo es encontrar estas relaciones de forma autónoma. Sin embargo, es el objetivo de este trabajo facilitar al algoritmo el descubrimiento de estas relaciones tomando como referencia técnicas aplicadas por las herramientas actuales de generación espectros sintéticos.

1.1. Estructura de la memoria

A continuación, se expone brevemente la temática de los capítulos de los que se compone el siguiente documento.

El Capítulo 2 se divide en tres apartados que presentan el contexto del trabajo: se desarrolla el estado del arte de los campos en los que se basa el desarrollo presente: la regresión simbólica y la relación entre parámetros atmosféricos y sus espectros estelares. Se exploran herramientas existentes que usan técnicas de regresión y sus eficiencias, así como la elegida, y en el campo de los parámetros aquellas que son capaces de generar espectros a partir de parámetros y viceversa. El apartado final detalla los objetivos concretos del trabajo: cómo se pretende aplicar exactamente la regresión simbólica para enlazar los parámetros estelares

con sus espectros, es decir, qué fórmulas matemáticas ha de descifrar y cómo se usarán posteriormente para crear los espectros.

El Capítulo 3 describe en detalle el desarrollo completo llevado a cabo para probar la regresión entre parámetros y flujos espectrales. Se divide también en varios apartados. Se detalla la búsqueda de catálogos de datos de espectros y parámetros, cual se ha decidido usar y por qué. Se describe la aproximación inicial: el proceso completo realizado para inferir la relación entre un espectro sintético individual con sus parámetros, así como las decisiones que luego servirán para el desarrollo final (lista de líneas atómicas, selección, ajuste de Gauss). Se describen finalmente los tres entrenamientos llevados a cabo con multitud de datos espectrales y los resultados preliminares obtenidos.

El Capítulo 4 muestra los resultados del trabajo. Se prueba la eficiencia de las funciones halladas con regresión lineal seleccionando un número de espectros RVS de Gaia, sintetizándolos en base a sus parámetros con los métodos desarrollados y comparándolos con los espectros originales.

El Capítulo final compila las conclusiones del trabajo y describe posibles mejoras y avances en el futuro de la regresión simbólica en este campo.

2. Estado del arte y objetivos

En este capítulo se presenta el contexto del trabajo: el estado del arte de los campos en los que se basa el desarrollo (la regresión simbólica y la relación entre parámetros atmosféricos y sus espectros estelares) y los objetivos concretos del mismo.

2.1. Regresión simbólica

Hasta 2020, el software más fructuoso que había implementado regresión simbólica había sido Eureka (Dubcáková, 2011). Eureka se probó contra un bloque de 100 ecuaciones (Udrescu & Tegmark, 2020) de múltiples áreas de la física provenientes del libro *Feynman Lectures on Physics* (Feynman et al., 1964) y uno extra de 20 de mayor complejidad extraídas de distintos libros de física, como *Classical Mechanics* (Herbert Goldstein, 2002) o *Classical electrodynamics* (Jackson, 1999). El software consiguió una efectividad del 68% sobre el bloque principal y un 15% sobre el bloque extra.

En 2020 surgieron nuevos desarrollos inspirados en híbridos entre las técnicas tradicionales de fuerza bruta y enfoques más modernos como es el uso de redes neuronales. Al Feynman (Udrescu & Tegmark, 2020) combinó pasado y presente, y mediante el uso de redes neuronales consiguió reducir iterativamente el problema principal a otros más pequeños y abordables, eliminando variables dependientes mediante el descubrimiento de propiedades de las ecuaciones como simetrías o posibles separaciones. Este software se probó contra los dos bloques descritos en el párrafo anterior, y se consiguió una efectividad del 100% en el bloque de 100 ecuaciones y una del 90% en el bloque extra de 20 ecuaciones.

Aún más recientemente se encuentra PySR (Cranmer, 2023), en forma de librería de código abierto. Esta herramienta se apoya en tecnologías más modernas para aumentar su rendimiento: hace uso de un algoritmo evolutivo multi-población (computa múltiples evoluciones de forma asíncrona), un backend distribuido y múltiples integraciones con paquetes Python (Van Rossum, 2009) de aprendizaje automático. La librería se probó con un set distinto al de las ecuaciones del libro de Feynman, criticando que estas están listadas con las constantes físicas asociadas, mientras que el set en cuestión con el que se probó

(EmpiricalBench) cuenta con expresiones empíricas reales descubiertas por científicos a partir de datos experimentales y con ruido: la ley de Hubble, la tercera ley de Kepler, la ley de Newton de la gravitación universal, la ley de Planck... De las mencionadas, más seis adicionales, la librería fue capaz de generar al completo 7 de ellas.

2.2. Espectros y parámetros

Existen dos estrategias principales para inferir la dependencia de un flujo espectral con los parámetros atmosféricos y las abundancias químicas en la superficie de la estrella. El primero consiste en la comparación de los espectros observados con teóricos, es decir, sintéticos.

El módulo GSP Spec (Recio-Blanco et al., 2023), General Stellar Parametrised-spectroscopy, consiguió producir un catálogo de 5.6 millones de espectros estelares con parámetros físicos y químicos asociados, el más grande publicado hasta la fecha. GSP Spec estimó los parámetros a partir de espectros, mediante dos técnicas alternativas: un complejo algoritmo entrenado mediante regresión lineal, MATISSE (Recio-Blanco et al., 2006), o una red neuronal basada en aprendizaje automático. En cualquier caso, ambas técnicas fueron entrenadas a partir de un set de 51373 espectros sintéticos y sus parámetros fisicoquímicos. Esta cantidad de espectros se generó de manera que abarcasen todos los posibles parámetros de todas las poblaciones estelares galácticas, obteniéndose así una precisión muy alta al entrenar a los algoritmos y luego usarlos para inferir propiedades.

GSP Spec computó los espectros siguiendo el método descrito en el proyecto de AMBRE (Laverny et al., 2012), con una serie adicional de mejoras y técnicas de refinamiento. El proyecto AMBRE creó una extensa librería de espectros sintéticos que abarcaba un amplio rango de parámetros atmosféricos estelares y un ancho dominio espectral, además de contar con una alta resolución espectral, parámetros bien establecidos, y variaciones de abundancias químicas. La creación de los espectros se basó en los modelos atmosféricos MARCS (Gustafsson et al., 2008), que llevan en constante evolución durante 40 años: estos modelos asumen geometría esférica, equilibrio hidrostático y termodinámico local para todas las estrellas. Los modelos se usaron como base para generar los espectros mediante el código

El factor principal para hallar los parámetros en este método es, por tanto, la anchura de las líneas espectrales.

2.3. Objetivos

El trabajo presente pretende revertir el proceso usado en el método de la anchura equivalente, no obtener los parámetros atmosféricos a partir de un perfil gaussiano sino computar el perfil gaussiano a partir de ellos. Para llegar a este punto, se pretende complementar la distribución gaussiana para que se ajuste a cada línea de absorción concreta con una función adicional condicionada por los parámetros estelares, por lo cual, se obtendría una línea de absorción concreta y única para cada set de parámetros estelares. Un bloque de fórmulas computadas por regresión lineal y dependientes de los parámetros estelares (una por cada línea de absorción) acabaría formando el espectro completo: un continuo condicionado por múltiples líneas de absorción características.

Para hallar la función clave se debe aplicar la regresión lineal sobre un conjunto de espectros con sus respectivas propiedades bien afianzadas. El mecanismo de regresión lineal debe ajustar la distribución gaussiana de cada línea de absorción en función de los parámetros estelares y no de la media ni de la desviación típica, ya que serán parámetros desconocidos cuando se intente producir el espectro.

La función de densidad de probabilidad de una distribución gaussiana de absorción, considerando un espectro con continuo de 1, viene dada por (C., 2009):

$$f(x) = 1 - D \exp\left(\frac{-(x - x_0)^2}{2\sigma^2}\right) \quad (1)$$

siendo x_0 la media de la distribución, σ la desviación típica, D la altura en el pico de la curva y x la variable. Al producir el espectro usando la distribución x sería la variable a hallar, es decir: la intensidad en una frecuencia concreta dentro de una línea de absorción. La media, el punto de mayor profundidad de la distribución (dicha profundidad es a su vez la altura en el pico de la curva), es por tanto la frecuencia exacta de absorción del elemento cuya línea de absorción se intenta ajustar. La desviación es desconocida y ha de ser distinta para cada línea de

absorción y dependiente de los parámetros estelares, por tanto: la variable se representará por medio de una función, y será el trabajo de la regresión lineal hallarla. A mayores la profundidad de la gaussiana (la profundidad de la línea de absorción) también será desconocida a la hora de generar un espectro, por lo que la regresión habrá de generar otra función adicional que represente este parámetro también.

$$f(x) = 1 - f(params) \exp\left(\frac{-(x - x_0)^2}{2 [h(params)]^2}\right) \quad (2)$$

El resultado final tras el entrenamiento del mecanismo de regresión deben ser dos funciones por cada posible línea de absorción en un rango de frecuencias concreto, a partir de las cuales se obtendrán la desviación típica de cada línea y su profundidad, pudiendo por tanto representarla cómo una distribución gaussiana. El conjunto de líneas sobre un espectro continuo compondrá el espectro final.

Sobre el algoritmo de regresión simbólica, tras el análisis realizado se decide usar la herramienta PySR para aplicarlo debido a la modernidad, integración con nuevas tecnologías y eficiencia mostrada.

Todo el desarrollo, desde la descarga de los datos, ajuste de perfiles gaussianos, ejecución del algoritmo de regresión y evaluación del mismo se implementará en Python en forma de distintos módulos que realicen los distintos pasos, debido a su sencillez y principalmente a su cantidad de bibliotecas, entre las que se encuentran múltiples de astronomía y tratamiento de espectros.

3. Desarrollo

En este capítulo se detalla el desarrollo completo llevado a cabo para probar la regresión entre parámetros y flujos espectrales con el objeto de poder generar espectros a partir de las ecuaciones obtenidas.

3.1. Datos

Una búsqueda exhaustiva de los catálogos actuales más significativos de espectros estelares, con propiedades atmosféricas asociadas, determina que el más preciso y completo es la tercera publicación de datos obtenidos por la misión Gaia, de la Agencia Espacial Europea: Gaia Data Release 3 (Gaia Collaboration, 2023a), que cuenta con 220 millones de espectros de baja resolución y 1 millón de alta. El catálogo proporciona los espectros continuos y su versión calibrada (con la que se pierde algo de información respecto al original), procedentes del espectrómetro de baja resolución (variable, entre 30 y 100 $\lambda/\Delta\lambda$) presente en el satélite Gaia, que abarca las frecuencias desde 330 a 680 nm y desde 640 a 1050 nm. En su contrapartida, el satélite cuenta con el espectro RVS, Radial Velocity Spectrometer (Cropper et al., 2018), que cuenta con una resolución de $R \sim 11500$ y abarca las frecuencias desde 846 a 870 nm.

La sustancial diferencia en resolución permite obtener parámetros astrofísicos más fiables y de más alta calidad de los espectros RVS, además de que gracias al parametrizador GSP-Spec (Laverny et al., 2012), se consigue a mayores inferir la abundancia de 13 especies químicas a partir de estos espectros de alta resolución.

Se decide por tanto para el presente trabajo, validar los mecanismos de regresión simbólica con la dependencia entre espectros RVS del catálogo Gaia DR3 y sus parámetros astrofísicos, debido a que con el rango de tiempo que se tiene para completar el estudio y la capacidad del sistema para aplicar las técnicas de regresión, un millón de espectros son más que suficientes para validar los resultados. A mayores, usar estos espectros provee la posibilidad de posteriormente usar la regresión simbólica para relacionar los espectros con las abundancias químicas de la estrella.

Sin embargo, una de las principales ventajas del uso de los espectros RVS en este trabajo es que, como se mencionó en la introducción, la regresión lineal se usará para computar un espectro considerando este como un continuo compuesto por múltiples líneas de absorción condicionadas por los parámetros atmosféricos. El amplio rango de frecuencias que los espectros de baja resolución abarcan no permitiría considerar el espectro como un continuo, pero el estrecho rango de los RVS sí.

Para obtener los espectros, se usa preliminarmente la tabla principal del catálogo Gaia DR3 *gaia_source* (Gaia Collaboration, 2023a), que contiene una entrada por cada fuente observada. Las entradas se filtran por la columna *has_rvs*, que indica aquellos objetos que tienen un espectro de alta resolución asociado. De los espectros a usar, se extrae la columna principal *source_id* con el identificador del objeto, con el que se descargará el vector con el flujo completo del espectro (almacenados por separado debido a su tamaño, como un recurso *Datalink*) así como posteriormente los parámetros.

Respecto al acceso al catálogo y a sus tablas, Gaia presenta sus recursos de manera pública y proporciona dos alternativas principales para extraer los datos: la descarga directa desde el archivo Gaia, accesible vía web, o la descarga a través de su API con peticiones ADQL (Astronomical Data Query Language). Todo el desarrollo presente se va a implementar en Python, por lo que se usa la segunda opción debido a la ventaja de la automatización.

Sobre los parámetros estelares, se probarán los mecanismos con un set básico generados mediante GSP-Spec (Gaia Collaboration, 2023b) y relativos a espectros del espectrómetro de Gaia RVS, para hacer así un análisis inicial de la eficiencia de la regresión simbólica y las distintas herramientas existentes para aplicar la técnica, sin sobre complicarlo con parámetros excesivos. De la tabla de Gaia DR3, *astrophysical_parameters*:

- *teff_gspspec*: temperatura efectiva en Kelvin.
- *logg_gspspec*: logaritmo de la gravedad en la superficie estelar.
- *mh_gspspec*: metalicidad global (M/H).

Se obvian los parámetros de abundancias de 13 especies químicas que GSP-Spec es capaz de generar, así como el genérico de abundancia de elementos alpha respecto al hierro, de

importancia pero que añaden significativa complicación adicional para el algoritmo de regresión.

Será de uso sin embargo un parámetro adicional de la tabla *astrophysical_parameters* generado por GSP-Spec:

- *logchisq_gspspec*: logaritmo del nivel de ajuste de GSP-Spec.

El grado del ajuste es generado para cada espectro parametrizado por GSP-Spec. Comparan el espectro original y uno sintético generado correspondiente a los parámetros atmosféricos obtenidos a partir del original y almacenan el logaritmo del ajuste entre ambos espectros como el parámetro mencionado. Debido a que GSP-Spec proporciona parámetros para prácticamente el millón completo de espectros medios RVS del catálogo GDR 3, y que son demasiados para validar el algoritmo de regresión en el alcance del presente trabajo, se usará este parámetro de la eficiencia del ajuste para filtrar y usar solo aquellos espectros cuyos parámetros formen un espectro sintético más similar al original, de forma que se obvien aquellos espectros más caóticos y difíciles de ajustar.

A mayores se ha de tener en cuenta que el espectrógrafo de alta resolución de Gaia, como todos, tiene ruido en sus capturas, ruido que puede provocar la aparición de líneas de absorción adicionales irreales o alterar la existentes. Esto puede ser un grave problema para el mecanismo de regresión lineal, que no distinguirá de lo que es ruido y no y por tanto ajustará las funciones de las distribuciones gaussianas sobre este ruido de fondo, produciendo un error inevitable en el cómputo. Una forma de aligerar este fallo puede ser indicándole al mecanismo de alguna forma el ruido que puede tener el flujo del espectro que está ajustando para que lo tenga en cuenta, conseguible añadiendo un parámetro adicional a los tres intrínsecos de las estrella que se usarán para computar las funciones: la razón de señal a ruido del espectro de alta resolución. De la tabla *gaia_source*:

- *rvs_spec_sig_to_noise*: la razón de señal a ruido en el espectro RVS.

Sin embargo, como se explica más adelante, esta idea inicial de usar el ruido como parámetro adicional en la regresión simbólica se descarta inicialmente en el alcance del presente trabajo debido a que se usan espectros sintéticos (sin ruido) para entrenar el algoritmo.

De forma extra, la aplicación del estimador FLAME (Kordopatis et al., 2023), *Final Luminosity Age Mass Estimator*, sobre los parámetros producidos por GSP-Spec junto con la aplicación de modelos estelares y datos fotométricos permiten estimar otros parámetros primordiales como son el radio, la luminosidad, la masa o la edad de la estrella. FLAME se encuentra aplicado tanto a parámetros de GSP-Phot (que trabaja con fotometría de espectros de baja resolución de Gaia), en la tabla *astrophysical_parameters* mencionada previamente, como a parámetros de GSP-Spec. Siguiendo con la política establecida se consultan las estimaciones de FLAME a partir de los parámetros de GSP-Spec, que contiene muchas menos estrellas (FLAME computó parámetros de 300 millones de fuentes a partir de GSP-Phot y de 5 millones a partir de GSP-Spec), pero de mayor fiabilidad debido a que solo se usan espectros de altas resolución. Estas estimaciones se encuentran en la tabla *astrophysical_parameters_supp*. Inicialmente se usan solo los tres parámetros indicados con anterioridad, pero tras una primera batida se pueden ampliar los parámetros de entrada para la regresión, entrenando al algoritmo con espectros RVS de Gaia, con objeto de comprobar si la eficiencia aumenta o se reduce, usando el ruido mencionado y a mayores los siguientes parámetros adicionales de FLAME:

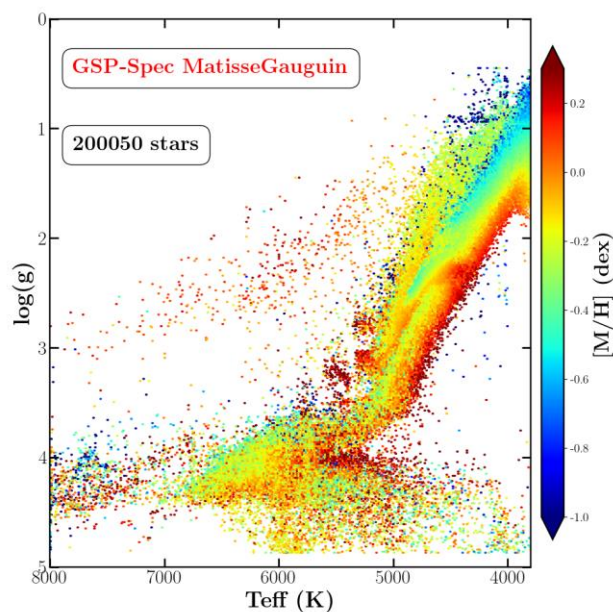
- *mass_flame_spec*: masa estelar en masas solares.
- *age_flame_spec*: edad estelar en giga años
- *lum_flame_spec*: luminosidad estelar en la banda G en función de la luminosidad solar.

La principal razón de descartar estos parámetros adicionales de clara relevancia en un inicio y mantener únicamente los tres elegidos (temperatura, gravedad y metalicidad) para inferir la relación entre espectro y parámetros reside en que la decisión tomada es de validar la eficacia del algoritmo con los espectros reales de RVS, pero no de entrenar al algoritmo con ellos, sino con espectros sintéticos. Esta decisión se apoya en una serie de razones: mayor facilidad de ajustar perfiles gaussianos en estos espectros, debido a la falta de ruido (un buen ajuste es esencial en los espectros con lo que se entrena la regresión, ya que conlleva que la línea de absorción sea lo más fiel a la realidad posible). Evitar cualquier fluctuación en las líneas de absorción debida a factores externos, que ensuciarían el trabajo del algoritmo ya que se crearían relaciones entre los parámetros de la estrella y un flujo que no se corresponde a ellos

con rigurosidad. Y principalmente por tener un control total del set con el que se entrena al algoritmo de regresión, en el cual se genera un abanico de espectros únicos que abarcan un rango completo de los tres parámetros de temperatura, gravedad y metalicidad elegidos, formándose multitud de combinaciones posibles para así lograr un entrenamiento rico y mixto que tiene en cuenta la mayor variedad de espectros posibles.

Los espectros generados han de tener los mismos parámetros técnicos que los RVS de Gaia, teniendo en cuenta que con estos se ha de validar el ajuste de la regresión, es decir: misma resolución y rango de frecuencias. En cuanto al abanico de rangos que abarcaran los tres parámetros elegidos, con los cuales se computarán los espectros, siguiendo la misma lógica son los mismos rangos que abarcan las estrellas capturadas el espectrómetro RVS. Estos rangos aparecen en el diagrama de Kiel (logaritmo de la gravedad en comparación con la temperatura efectiva de la estrella) en Recio-Blanco (2023), que muestra los dos parámetros generados por el parametrizador GSP-Spec, junto con una barra adicional que indica la metalicidad, para los 200050 espectros RVS parametrizados que tienen una señal a ruido superior a 150. Dicho diagrama se puede apreciar en la Figura 2.

Figura 2. Diagrama de Kiel para los parámetros de GSP Spec sobre espectros RVS



Fuente: (Recio-Blanco et al., 2023)

El diagrama muestra que los espectros de alta resolución parametrizados corresponden a una serie de estrellas que abarcan un rango aproximado de temperatura entre 3800K y 8000K, con logaritmos de gravedad variando entre 0.3 $\log(g)$ y 4.8 $\log(g)$, y metalicidades entre -1.1 [M/H] y 0.25 [M/H]. Estos valores serán los usados para generar los espectros sintéticos con los cuales se entrenará el algoritmo de regresión simbólica, intentando conseguir la mayor posibilidad de combinaciones entre ellos siempre y cuando la combinación sea físicamente posible.

3.2. Aproximación inicial

El ciclo automatizado de aprendizaje recorre cada espectro de entrenamiento cuyos datos vayan a usarse para entrenar al algoritmo de regresión simbólica es el siguiente: se normaliza el espectro de entrenamiento, se seleccionan las líneas atómicas del mismo, se ajustan a distribuciones gaussianas y finalmente se usan los parámetros de las distribuciones para hallar funciones mediante regresión simbólica que los calculen a partir de los parámetros atmosféricos de la estrella del espectro.

Este proceso, que se ha de realizar sobre multitud de espectros de manera secuencial para que el algoritmo de regresión simbólica tenga suficientes datos para generar funciones lo suficientemente genéricas y sólidas, se realiza en esta aproximación inicial sobre un único espectro sintético, para ir perfeccionando el proceso, que luego será automático, paso a paso y de manera controlada.

3.2.1. Generación de espectro sintético y tratamiento de los datos

Para probar la efectividad de la regresión en la batida inicial, como se ha mencionado, se decide no usar los espectros reales de RVS sino espectros sintéticos, ya que así, entre otras razones, se elimina el posible ruido que dificultaría la generación de las funciones (ya que las líneas de absorción se distancian de tener distribuciones gaussianas idóneas, porque se solapan con otras líneas producidas por el ruido).

Para ello se usa iSpec (Blanco-Cuaresma et al., 2014), una herramienta para el análisis de espectros estelares. Escrita en Python y de código abierto, iSpec surgió con el objetivo de

derivar parámetros atmosféricos y abundancias químicas estelares a partir de sus espectros mediante los dos métodos explicados con anterioridad: ajuste por comparación con espectros sintéticos generados en el momento o mediante el método de la anchura equivalente. iSpec pone a disposición del usuario su herramienta para generar los espectros sintéticos, en función de unos parámetros atmosféricos personalizables.

La herramienta de generación usa el código de transporte radiativo SPECTRUM (Gray & Corbally, 1994), que usa a su vez los modelos atmosféricos de Kurucz (Kurucz, *The Stellar Populations of Galaxies*, 1992), que con parámetros de entrada como la temperatura o las presiones computa, usando siete ecuaciones de equilibrio no lineales por cada nivel de la atmósfera: las densidades del Hidrógeno, Helio, Carbono, Oxígeno y Nitrógeno, sus iones, moléculas diatómicas que hayan podido surgir de estas especies y la densidad de electrones. En base a estos datos el código genera las opacidades sobre el continuo, obteniendo la absorción para las líneas atómicas de una lista concreta. iSpec es actualmente capaz de usar distintas listas de distintas fuentes, pero los mejores resultados se obtuvieron con dos fuentes (Blanco-Cuaresma et al., 2014):

- Dos listas de la base de datos VALD (Kupka & Dubernet, 2011): su última entrega, VALD-3, contiene 160 listas de líneas y un total de 66 millones de líneas atómicas. Las dos listas que incluye iSpec por defecto abarcan un amplio rango de frecuencias: desde los 300nm hasta los 1100nm y desde los 1100nm hasta los 24000nm.
- La lista de Gaia-ESO Survey (Heiter et al., 2021), que cubre un rango desde los 475nm hasta los 685nm y desde los 850 hasta los 895nm, proporciona una selección de 1300 líneas de media y alta calidad de 35 elementos.

iSpec indica en sus pruebas de desarrollo que ambas listas son muy similares, pero que encontraron una dispersión menor con la de Gaia. Este hecho junto a la homogeneidad que supondría usar la lista de Gaia, ya que se van a usar los espectros y parámetros de la misma misión para el trabajo presente, serían suficiente razón para decantarse por esta en vez de por la de VALD; sin embargo, la lista de Gaia cuenta con un significativo problema: el rango espectral no cubre el rango completo que abarcan los espectros RVS con los que se va a

ejecutar el algoritmo, de 846nm a 870nm. La lista no cubre los 4nm desde los 846nm hasta los 850nm¹, que, aunque parezcan insignificante representan un 16.7% del rango del espectro. Por esta razón, ya que no se pueden ignorar las líneas de absorción de más de un décimo del espectro, se decide computar los espectros sintéticos con la lista de líneas atómicas de VALD, y a su vez, se usará esta lista para obtener las líneas en las cuales el algoritmo generará funciones que acompañen a las posibles distribuciones gaussianas existentes.

Usando la herramienta proporcionada por iSpec, se deben asignar varios parámetros para generar el espectro:

- Código radiativo: cómo se mencionó, iSpec usó en Blanco-Cuaresma (2014) el código SPECTRUM, pero ofrece otros con TurboSpectrum (el usado por GSP Spec para sus espectros sintéticos) o MOOG (Snedden et al., 2012). En Blanco-Cuaresma (2019) el creador de iSpec muestra que cada uno tiene sus ventajas y desventajas, dependiendo de la estrella con la que se esté lidiando y de las líneas atómicas. Se decide usar TurboSpectrum al ser el que GSP Spec usó, en base a que para los resultados finales se usarán los parámetros generados por la misma herramienta.
- Modelo atmosférico: se presentan los modelos de MARCS (Gustafsson et al., 2008) y ATLAS9 (Kurucz, 2005). Siguiendo la lógica anterior, se selecciona MARCS al ser el usado por GSP Spec.
- Abundancias solares: los códigos radiativos necesitan las abundancias para escalar los valores en función de la metalicidad especificada para la estrella, iSpec ofrece siete distintos. La mejora del código TurboSpectrum (Gerber et al., 2023) para sintetizar líneas de elementos sin equilibrio termodinámico local, usó el modelo atmosférico MARCS con las abundancias solares de Grevesse de 2007 (Grevesse et al., 2007), presente en iSpec y por tanto el que se selecciona.

¹ El programa de iSpec muestra en la herramienta una lista de líneas de la Gaia-ESO Survey que abarca desde 420nm a 920nm, sin embargo, no se ha podido encontrar ningún artículo de la investigación Gaia que indique que la lista abarca este rango, toda referencia menciona que la lista abarca los dos rangos más estrechos mencionados con anterioridad. Por tanto, con el desconocimiento de cómo iSpec ha ampliado la lista para que abarque un rango bastante más amplio, se sigue decidiendo usar VALD cuyo origen si es conocido y validado.

- Lista de líneas atómicas: por las razones mencionadas, se usa la lista de VALD desde los 300nm hasta los 1100nm.

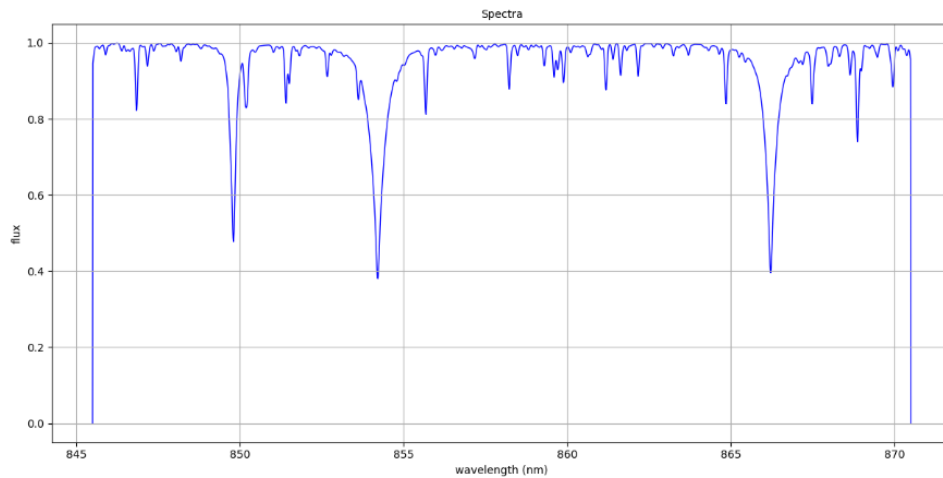
Con esto especificado, la herramienta tendría la teoría con la cual generar el espectro, ahora se deben indicar las características de la estrella cuyo espectro se quiere computar. Para una primera ejecución, se asignan los parámetros atmosféricos de una estrella de tipo G similar al sol del catálogo de Gaia DR3 que tuvo un ajuste muy bueno:

- Temperatura efectiva: 5837K
- Gravedad: 4.06 log(g)
- Metalicidad: 0.24 [M/H]

El resto de parámetros se mantienen por defecto (asignados en base a los valores recomendados por el set de estrellas comunes de calibración Gaia Benchmark Stars (Heiter et al., 2015)) excepto algunos que se ajustan a las propiedades de los espectros generados por el RVS de Gaia: la resolución a 11500 y longitudes de onda de 846 a 870 nm² en saltos de 0.01 nm.

² Una generación inicial descubre que iSpec tiene un fallo en su algoritmo en el cual no se genera flujo en las frecuencias límite inferior y superior, como se puede apreciar en la Figura 3. Por lo tanto, se aumentarán los límites ligeramente a la hora de generar y luego se recortarán para que cuadren con el rango de los espectros RVS de Gaia.

Figura 3. Espectro inicial generado con iSpec



Fuente: Elaboración propia

El espectro generado por iSpec se puede observar en la Figura 3 y sus correspondientes datos se almacenan en un archivo de texto con tres columnas: frecuencia, flujo y error. Los datos se tratan de la siguiente forma:

- El error del flujo se ignorará de momento, tanto en este espectro como el de los de Gaia, para no añadir más complejidad al algoritmo de regresión.
- Se eliminan las frecuencias adicionales que pueda haber, dejando únicamente el flujo entre 846nm y 870nm.
- Longitudes de onda: tanto la tabla de valores de frecuencias generada por iSpec como la de los espectros RVS de Gaia cuentan con valores que no progresan de forma regular (debería progresar de 0.01 nm en 0.01 nm), sino que ocasionalmente se desvían decimales a la potencia de -12, lo que conlleva el desvío de todas las frecuencias posteriores. Por ejemplo, estas son las frecuencias de las dos últimas líneas de la tabla del espectro generado con iSpec:
 - 869.9799999999782 - 869.9899999999782
- Esto es recurrente y ocurre en todos los espectros, por lo que antes de que se pasen los datos al algoritmo de regresión se aproximan todas las longitudes de onda a 2 decimales, consiguiendo así que cada una represente exactamente la diferencia de nm

que debería haber entre cada una (0.01 nm). Las dos últimas con redondeo por tanto quedarían de la siguiente manera:

- 869.98 – 869.99
- Valores: los flujos generados cuentan con una gran cantidad de decimales que añaden complejidad y dificultan el trabajo del algoritmo, por lo que se redondean a 4 decimales.
 - Se hará lo mismo con los valores de validación finales de los parámetros obtenidos de GSP Spec.

3.2.2. Selección de líneas atómicas

Con los datos del espectro con el cual se va a probar el algoritmo por primera vez, la siguiente tarea es concretar qué líneas de absorción se van a intentar ajustar. Como ya se mencionó, el algoritmo de regresión ha de generar dos funciones dependientes de los parámetros atmosféricos por cada línea de absorción del espectro, con objeto de que estas permitan ajustar un perfil gaussiano hallando la línea de absorción y completando el espectro.

La lista de líneas de absorción a usar será la de VALD desde los 300nm hasta los 1100nm. Se obtiene del propio código fuente de iSpec, que tiene multitud de listas de líneas atómicas almacenadas en archivos .tsv. La de VALD en concreto cuenta con 731377 líneas y 38 parámetros asociados a cada una de ellas. Por tanto, se adapta:

- Se mantienen únicamente aquellas líneas que abarquen desde los 846 nm hasta los 870 nm.
- Al igual que se hizo en el proceso de selección de líneas para la última ampliación de iSpec (Blanco-Cuaresma S. , 2019), solo se seleccionan aquellas líneas que tienen una profundidad teórica mayor a 0.01.

Con esta reducción se obtiene una lista final de 186 líneas atómicas. La precisión en las frecuencia de las líneas cada elemento en la lista es de 4 decimales, sin embargo, las frecuencias de los espectros RVS normalizados de Gaia con los que se validará solo cuentan con una precisión de 2 decimales (saltos de 0.01 nm). Por tanto, las frecuencias de las líneas se deben redondear a 2 decimales, lo que resulta en que de todas las frecuencias 10 de ellas

albergan dos líneas simultáneamente. Siguiendo la lógica aplicada en la selección de líneas para la ampliación de iSpec, en estos casos en los cuales dos líneas atómicas se pueden solapar en una única, se elige aquella con la mayor anchura equivalente teórica (es decir, la que más probabilidades tiene de ser la principal contribuyente a la absorción) y se descarta la otra. Finalmente, la lista obtenida cuenta con 176 elementos y líneas atómicas.

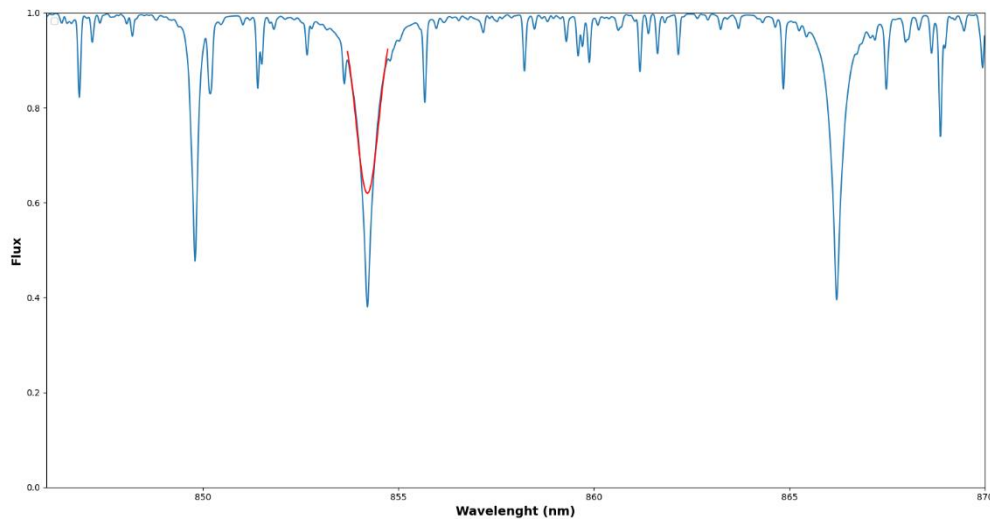
3.2.3. Normalización del espectro y ajuste de líneas atómicas

Antes de comenzar con la regresión simbólica, se van a intentar ajustar perfiles gaussianos a estas 176 líneas de absorción con el espectro con parámetros comunes a estrellas de tipo G sintetizado mediante iSpec, ya que para que el algoritmo encuentre las funciones incógnito que permitirán computar la desviación típica y la profundidad en función de los parámetros atmosféricos, se han de conocer previamente ambos para entrenarlo. Se intenta por tanto hallar estos parámetros para las distribuciones de las líneas presentes. Este proceso se realizará con cada espectro que se quiera usar para entrenar al mecanismo de regresión simbólica.

3.2.3.1. Normalización del espectro y primera selección de líneas atómicas

Un proceso inicial de ajuste de perfiles en las líneas más profundas del espectro mostró una dificultad para hallar desviaciones estándar adecuadas con la cuales se minimizara el error de éste, como se puede observar en la Figura 4 en el intento de ajuste de la línea de absorción del Ca II, en la longitud de onda de 854.21 nm del espectro sintético, que es la de mayor profundidad del mismo.

Figura 4. *Ajuste inicial fallido de Gauss*

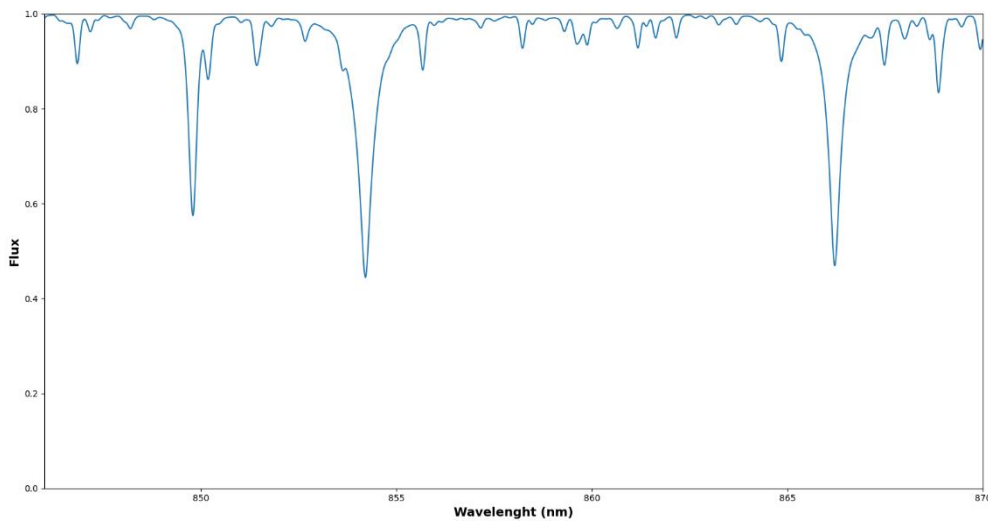


Fuente: Elaboración propia

Se prevé demasiada dificultad para ajustar las líneas de gran profundidad, así como para la gran cantidad conjunta que se observa, entre los 858 y los 863 nm en el espectro presente, siendo encima estas líneas cercanas muy estrechas y con caídas de profundidad verticales (lo que va en contra de la campana de Gauss y dificulta su ajuste), y se debe tener en cuenta a mayores que esto es un espectro sintético y que los reales de Gaia con los que se entrenará el algoritmo tendrán aún mayor ruido e incertidumbre. Todo esto conduce a la conclusión de que se han de normalizar los espectros para suavizarlos y facilitar el ajuste. Se deciden normalizar mediante convolución usando la librería Astropy (Astropy Collaboration, 2022), una de las librerías de Python más reconocidas para astronomía. Astropy proporciona distintos filtros (*kernels*) para normalizar listas de valores mediante convolución, en este caso se usará para normalizar la lista de valores de flujo a lo largo del espectro. Cada filtro está caracterizado por una función de respuesta, siendo dos disponibles, por ejemplo, el de Lorentz (que será de uso para suavizar datos que sigan curvas de Lorentz) o de Gauss. Naturalmente se decide usar el filtro de Gauss para suavizar el espectro. Se debe asignar la desviación estándar del filtro antes de usarse para convolucionar (es decir, el número de elementos de la lista de flujos, o píxeles, como lo denomina Astropy): una mayor desviación conlleva una

mayor normalización y por tanto mayor pérdida de detalles. Tras una serie de pruebas se decide usar una desviación estándar de 5 en el filtro para convolucionar, ya que se considera que se consigue un espectro mucho más suavizado y fácil de trabajar, pero sin llegar a perder líneas atómicas relevantes. El resultado de convolucionar el espectro generado (Figura 3) se puede apreciar en la Figura 5.

Figura 5. Espectro inicial generado con *iSpec* normalizado

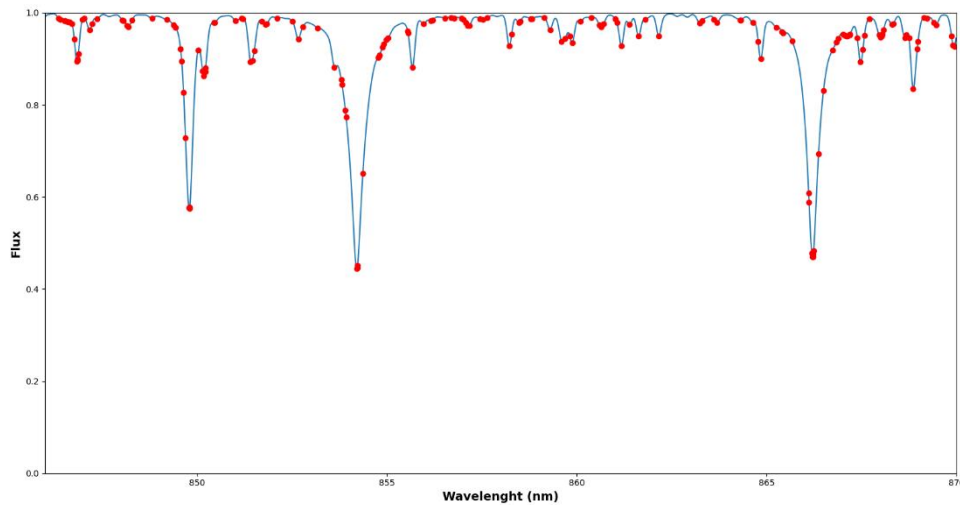


Fuente: Elaboración propia

El método implementado por *iSpec* no intentaba ajustar distribuciones en el espectro para aquellas líneas cuya profundidad fuera mayor que 1 o menor que 0.01, para el trabajo presente se aplica el mismo filtro.

En el caso del espectro sintético generado de las 176 frecuencias correspondientes a la lista de líneas 7 cuentan con una profundidad inferior a 0.01, por lo que se descartan y se mantienen las 168 restantes, que se pueden observar sobre el espectro normalizado en la Figura 6.

Figura 6. *Espectro normalizado con líneas de absorción tras primera selección*



Fuente: Elaboración propia

3.2.3.2. Segunda selección de líneas atómicas

Observando las frecuencias de absorción de todos los elementos de la lista sobre el espectro generado en el apartado anterior, en la Figura 6, se aprecia que muchas no se encuentran en un pico de máxima profundidad, representando una media de Gauss, sino a lo largo de una línea de absorción de mayor profundidad. Estas frecuencias no se van a poder ajustar como distribuciones de Gauss independientes ya que no representan la media de ninguna, sino que contribuyen a otra de mayor profundidad cuya frecuencia representativa sí que es la media de la distribución total, que abarcará las frecuencias de absorción de menor profundidad que no forman picos de absorción únicos (es decir, que no representan el punto más bajo de la distribución a la que pertenecen). Por tanto, ya que el algoritmo desarrollado funciona ajustando una distribución de Gauss por frecuencia de absorción, estas se descartan en el ámbito del presente trabajo y no se tienen en cuenta. A mayores se aprecia que varios puntos que representan las frecuencias de absorción se encuentran muy contiguos, siendo evidente que algunos representan la misma campana de absorción y por tanto solo se va a poder mantener y ajustar uno de ellos. Para descartar tanto estas frecuencias como las que no están en picos de absorción se desarrolla un método de limpieza adicional:

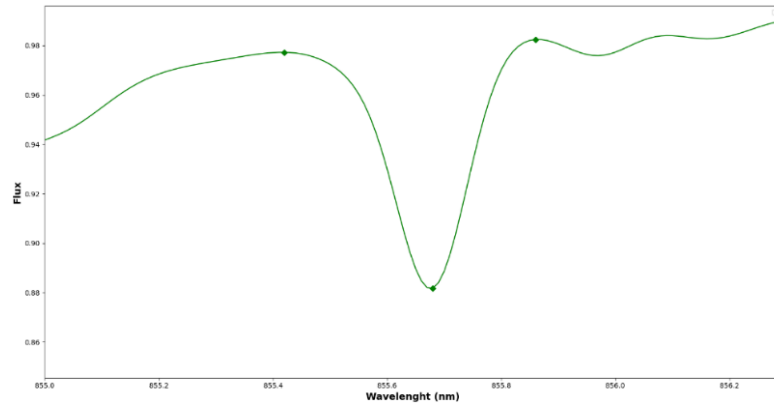
- Para cada línea de absorción presente, se extrae el flujo en la frecuencia concreta de absorción, es decir, en el pico de máxima profundidad.
- Simultáneamente se va recorriendo la lista de longitudes de onda hacia la izquierda y hacia la derecha, una a una, desde la longitud de onda del pico de absorción.
- Cuando se encuentra una longitud de onda cuyo flujo es menor que la de la última frecuencia comprobada, se apunta la última comprobada considerando que la línea de absorción ha terminado ahí ya que el flujo ha dejado de ascender.
 - Por ejemplo, se tiene un espectro con 100 puntos y estamos buscando el límite a la izquierda de la línea de absorción del Fe correspondiente a la posición 60. Si el flujo en la posición 52 es menor al de la posición 51, se marca 51 como la frecuencia límite izquierda para la línea de absorción del Fe.
 - Si se tiene en cuenta, como para los espectros presentes, que cada posición de la lista abarca 0.01 nm, consideraríamos que la línea de absorción del Fe tiene una anchura de 0.09 nm hacia la izquierda.

El mecanismo de limpieza calcula la anchura a la izquierda y a la derecha de cada frecuencia de absorción de la lista, y si alguna de las dos es menor a 0.03 nm se descarta, considerando que no forma una campana de Gauss válida para ajustar.

Este mecanismo a su vez proporciona los límites de cada línea de absorción permitiendo conocer cuáles son las frecuencias exactas que se han de ajustar mediante Gauss: todas aquellas que abarquen desde el límite izquierdo hasta el derecho.

Se pueden apreciar los límites encontrados por el mecanismo para la línea de absorción del Si I en la longitud de onda de 855.68 nm del espectro sintético generado con iSpec, en la Figura 7.

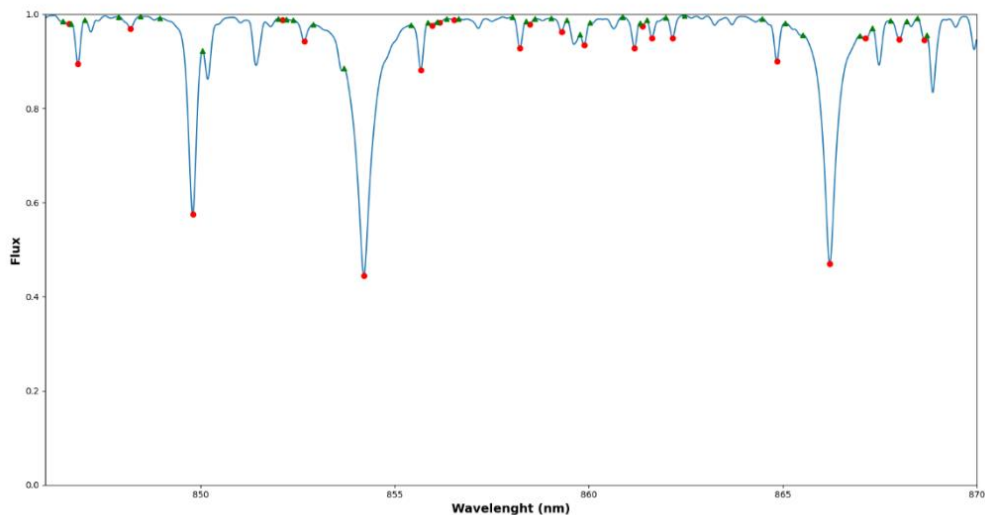
Figura 7. Límites de la línea de absorción del Si 1 en espectro sintético



Fuente: Elaboración propia

Tras ejecutar la herramienta sobre las 168 líneas presentes en la lista para el espectro sintético se encuentran 144 cuya anchura hacia la izquierda o hacia la derecha es menor a 0.03 nm, por tanto, se descartan, quedando 24 restantes que se pueden observar en la Figura 8 junto a sus límites.

Figura 8. Espectro normalizado con líneas de absorción y límites tras segunda selección



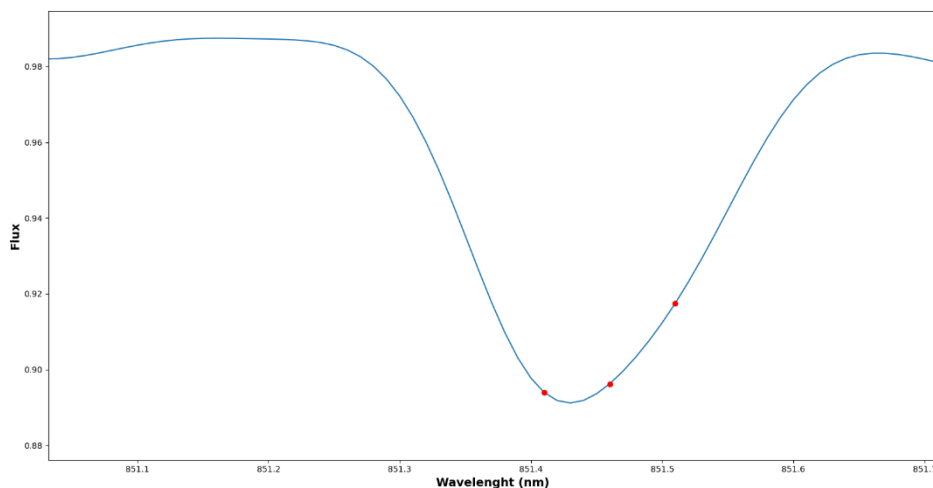
Fuente: Elaboración propia

3.2.3.3. Tercera selección de líneas atómicas

Al observar la Figura 8 se aprecia con claridad que la herramienta ha descartado frecuencias en demasía: se aprecia alguna longitud de onda de absorción con una profundidad visualmente relevante no identificada como una frecuencia presente en la lista, como es por ejemplo la existente entre 851.3 y 851.4 nm.

Revirtiendo la ejecución de la herramienta y aumentando esta línea de absorción para ver las frecuencias presentes antes de la limpieza, se aprecia con claridad en la Figura 9 la razón que ha producido que se descarten.

Figura 9. Frecuencias en línea de absorción descartada



Fuente: Elaboración propia

Las tres frecuencias observadas son respectivamente la del Fe 1 a 851.41nm, la del CN 1 a 854.46nm y de nuevo otra del Fe 1 a 851.51nm. La tercera se descartaría ya que está demasiado lejos del pico de máxima profundidad, sin embargo, cualquiera de las dos restantes podría ser la razón de que exista la línea de absorción y sin embargo ninguna de las longitudes de onda concretas de las dos coincide con el pico de máxima absorción de la línea, presente en 851.43nm en el espectro, por lo que la herramienta detecta que ni una ni otra está en un pico y los descarta.

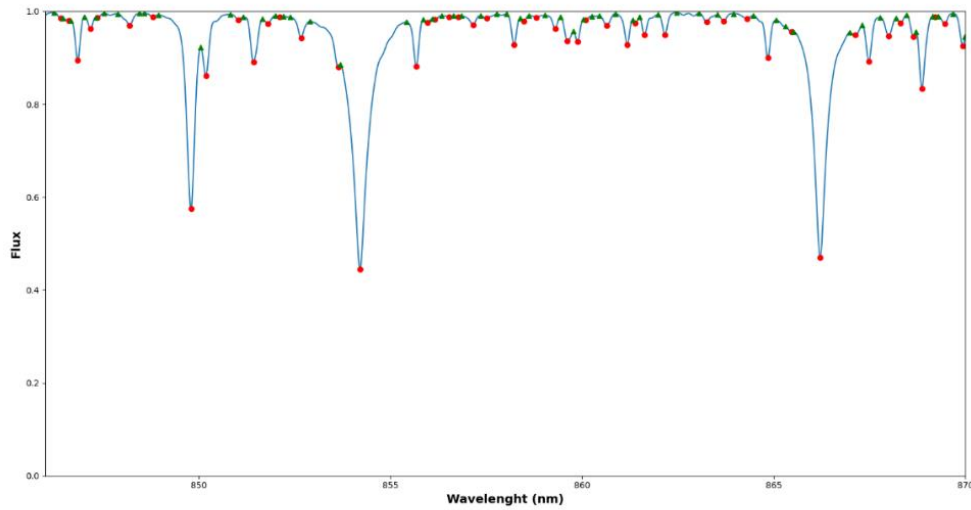
Este desplazamiento no es común ya que la mayoría de líneas de absorción coinciden con una frecuencia de la lista de VALD, sin embargo, ahora se ha de tener en cuenta que es una posibilidad. La razón se desconoce, ya que este espectro es sintético es posible que fuera un error de computación, pero a la hora de trabajar con los reales de RVS se ha de considerar que esta situación puede ser aún más frecuente debido a ruido, errores instrumentales o incluso algo de redshift.

Por tanto, como solución se añade una capa extra de complejidad al algoritmo de limpieza mediante límites desarrollado: antes de descartar una frecuencia, se probará a computar los límites con dos frecuencias a la derecha y otras dos a la izquierda (es decir, antes de descartar la 851.41nm, se probaría a calcular los límites con la 851.39, 851.4, 851.42 y 851.43), y si alguna de las cuatro muestra una anchura suficiente y no descartable, esa frecuencia se mantendrá, ya que el algoritmo de ajuste de Gauss no tendrá problema en descifrar que el pico se encuentra ligeramente desplazado respecto al propuesto.

La solución desemboca en otro problema, y es que las múltiples frecuencias que se encuentren concentradas alrededor del mismo pico de absorción van a resurgir. Para solucionarlo esta nueva capa se ejecutará tras el algoritmo inicial, y si una de las cuatro frecuencias alrededor de la descartada tiene límites adecuados, no se añadirá si esa frecuencia ya está presente en la lista.

Se ejecuta por tanto de nuevo la herramienta con la corrección sobre la lista de 168 líneas para el espectro sintético, y ahora se descartan 118: se han rescatado 26 de las descartadas en la limpieza anterior y se mantienen 50 líneas atómicas, que se pueden apreciar en la Figura 10 junto a sus límites.

Figura 10. Espectro normalizado con líneas de absorción y límites tras tercera selección

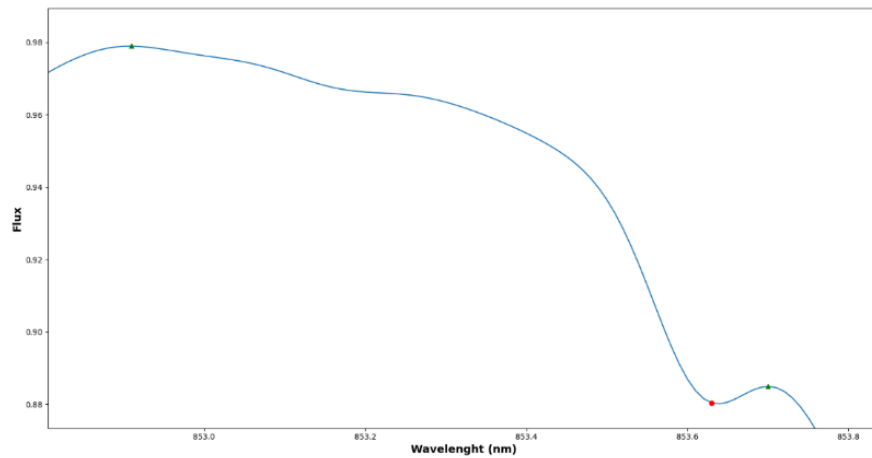


Fuente: Elaboración propia

3.2.3.4. Cuarta selección de líneas atómicas

En la Figura 10 es visible como la selección de frecuencias de absorción y sus límites generados aún no es correcta, habiendo varias dónde es claro que no se van a poder ajustar a distribuciones de Gauss. Un ejemplo claro es una de línea de absorción del Si 1 en 853.6 nm, no descartada ya que tanto sus límites izquierdos como derechos superaban una anchura de 0.02nm pero que como se puede observar en la Figura 11 sus límites están tan diferenciados en cantidad de flujo y distancia al pico máximo de absorción que será imposible ajustarse a una distribución de Gauss.

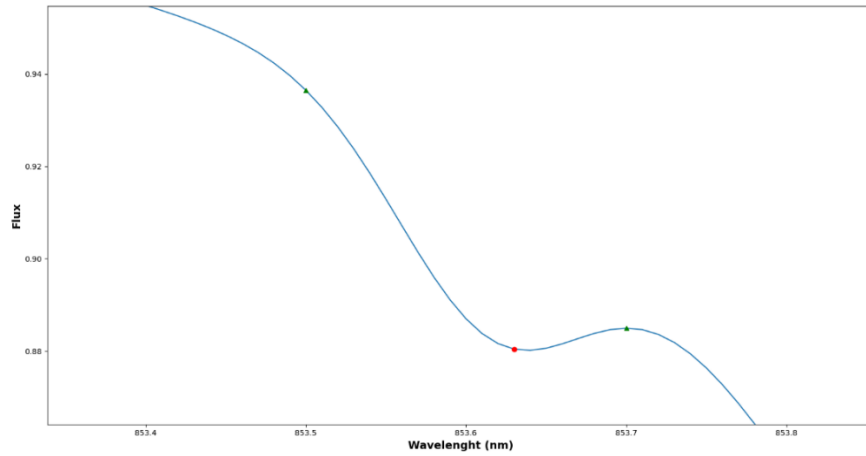
Figura 11. Límites erróneos de otra línea de absorción del Si 1 en espectro sintético



Fuente: Elaboración propia

Para solucionar el nuevo error se ha de hacer una cuarta iteración de la herramienta desarrollada para la selección de líneas atómicas y cálculo de límites. Siguiendo la filosofía de las distribuciones de Gauss y que los límites a ambos lados de las mismas deberían estar más o menos en un mismo plano (para espectros, con un mismo flujo) para las líneas de absorción mantenidas cuyos límites calculados tengan una diferencia de flujo mayor al 0.05, se modificará el límite de mayor flujo, desplazándolo las suficientes longitudes de onda a la derecha o a la izquierda (si se ha de modificar el límite derecho, se irá recorriendo hacia la izquierda, y viceversa) hasta encontrar un flujo que se encuentre a 0.05 o menor cantidad del límite con menor flujo, cuya frecuencia será establecida como el nuevo límite. Así, la línea de absorción del Si 1 en 853.6 nm tendría ahora los límites observados en la Figura 12, permitiendo un ajuste de Gauss tanto en la misma como en otras que sufrieron del mismo problema de cálculo.

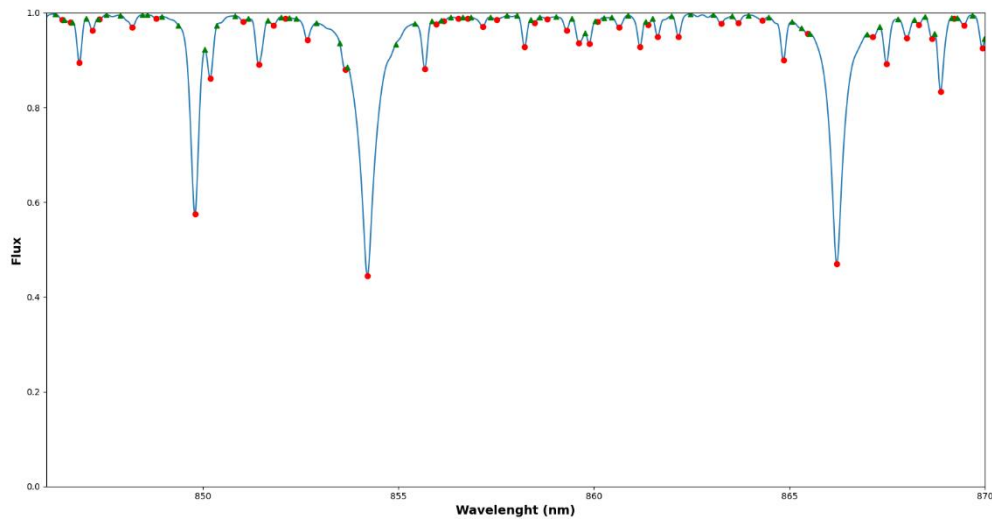
Figura 12. Límites corregidos de la otra línea de absorción del Si 1 en espectro sintético



Fuente: Elaboración propia

Con este último refinamiento se vuelve a ejecutar una última vez la herramienta sobre el set de líneas atómicas, seleccionando de nuevo las 50 anteriores, pero con límites reajustados, observables en la Figura 13.

Figura 13. Espectro normalizado con líneas de absorción y límites tras cuarta selección



Fuente: Elaboración propia

Este será el proceso automático para descartar líneas no aplicables de la lista de VALD en cada espectro que se use para entrenar el algoritmo de regresión, y para calcular los límites de las líneas de absorción seleccionadas, con el objetivo de seguidamente ajustar una distribución de Gauss a cada una.

3.2.3.5. Ajuste gaussiano a las líneas atómicas de absorción

Disecionadas las líneas atómicas ajustables en cada espectro, se han de calcular las distribuciones gaussianas que se ajusten a cada línea, con objeto de hallar los dos parámetros que luego la regresión simbólica ha de hallar (en función de los parámetros estelares del espectro) para poder computar esa línea de absorción y por consecuencia el espectro completo.

La fórmula de una distribución gaussiana de absorción (Ecuación 1, que se vuelve a añadir por comodidad del lector) es la siguiente:

$$f(x) = 1 - D \exp\left(\frac{-(x - x_o)^2}{2\sigma^2}\right) \quad (3)$$

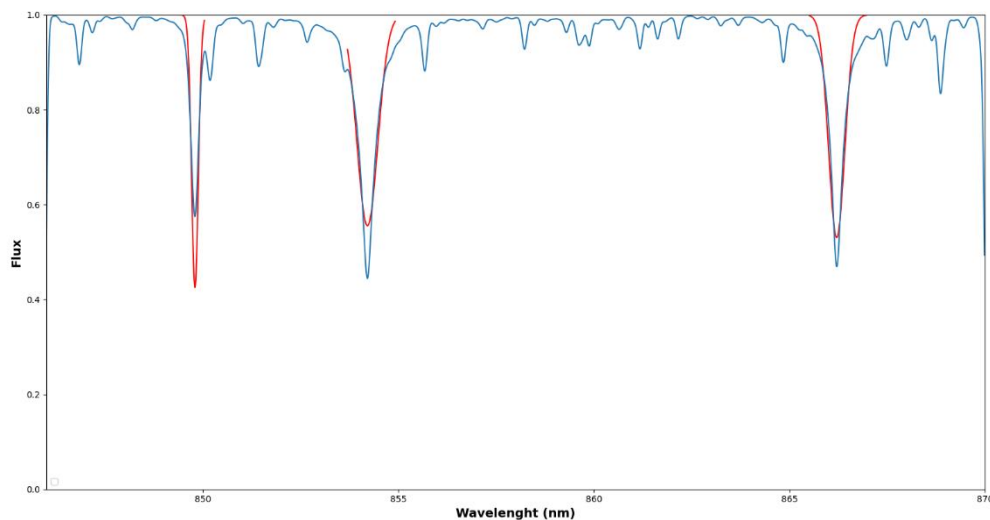
La profundidad (D) es el primer parámetro que se le debe dar al mecanismo de regresión simbólica y aquí no se ha de calcular, ya que ya es conocido. La media también se conoce, es la frecuencia exacta en la cual se encuentra el punto de mayor profundidad de la línea atómica. La desviación, el segundo parámetro que se le debe dar al mecanismo de regresión, es por tanto el único que se desconoce y que se ha de calcular en esta parte para cada línea atómica, intentando hallar un valor que maximice el ajuste de la distribución gaussiana al flujo presente en la línea.

Para hallar el valor, se decide usar la función `curve_fit` de la librería SciPy (Virtanen, 2020) de Python. La función mencionada utiliza el método de regresión de mínimos cuadrados no lineales para ajustar una función indicada a una serie de datos, hallando en el proceso todas las variables incógnitas presentes en la función. Para el problema presente, el algoritmo ha de intentar ajustar la distribución a las frecuencias contenidas entre los dos límites calculados

con anterioridad para cada línea atómica. La función gaussiana proporcionada como entrada al algoritmo tiene tres incógnitas, sin embargo, dos se conocen (la profundidad y la media), por lo que se usa el parámetro *bounds* opcional de la función para establecer valores límite a las incógnitas que ha de hallar: con esto se restringe la profundidad y la media, indicando los valores exactos que han de tener, y se indica un rango abierto de búsqueda para la desviación, permitiendo a la función centrarse únicamente en esta incógnita y en hallar el valor más óptimo.

En una primera ejecución del algoritmo sobre las tres líneas de absorción más profundas del espectro normalizado de iSpec (las tres producidas por el Ca II en 849.8nm, 854.21nm y 866.21nm) se produce el ajuste visible en la Figura 14.

Figura 14. Primer ajuste de gaussianas en espectro de iSpec



Fuente: Elaboración propia

Se eligen estas líneas en concreto como prueba de fuego inicial por la dificultad inherente que se encontrará el algoritmo al intentar ajustarlas, siendo las más profundas y distantes de campanas de Gauss ideales en el espectro presente. Como se puede apreciar, los ajustes no son del todo óptimos, presentando unas raíces del error cuadrático medio (*RMSE*) de 0.07,

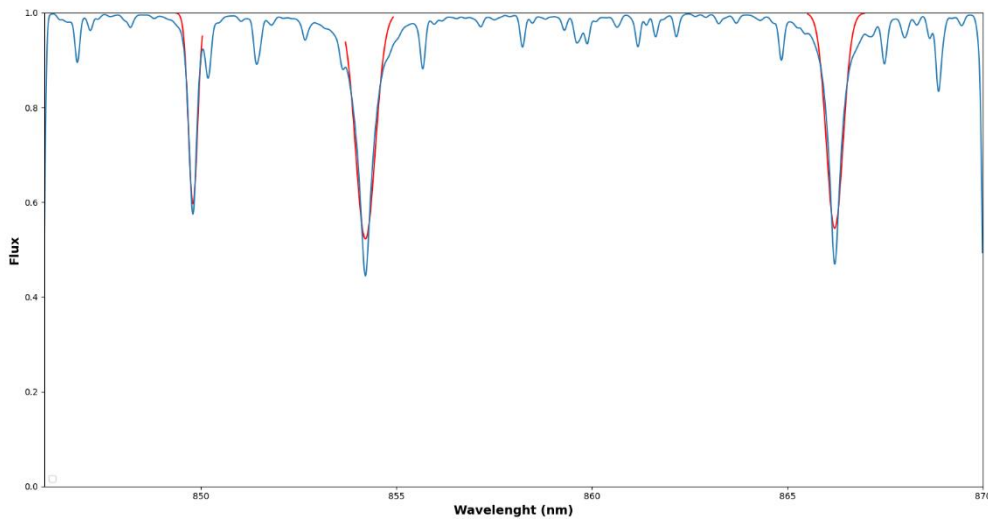
0.044 y 0.042 respectivamente, siendo el error más destacado el de la primera línea, la más estrecha de todas y la que más difiere de una campana de Gauss.

Tras una búsqueda referente a técnicas para mejorar el ajuste se encuentra que el fijar una profundidad concreta, aunque sea la real, puede dificultar el hallar una desviación que ajuste los datos de la manera más eficiente posible (principalmente en distribuciones muy profundas y estrechas, como la del Ca II en 849.8nm), por lo que una posible solución es añadir a esta profundidad un *offset* (Jaeyoung Chung, 2024), es decir, un valor de compensación, permitiendo por tanto a la función *curve_fit* de SciPy variar esta profundidad y modificar su valor si eso conlleva un mejor ajuste. La fórmula por tanto provista a la función sería:

$$f(x) = 1 - (offset + D) \exp\left(\frac{-(x - x_0)^2}{2\sigma^2}\right) \quad (4)$$

Una segunda ejecución del algoritmo con la fórmula modificada sobre las tres mismas líneas atómicas que se ajustaron anteriormente presenta el ajuste observable en la Figura 15.

Figura 15. Segundo ajuste de gaussianas con offset en espectro de iSpec



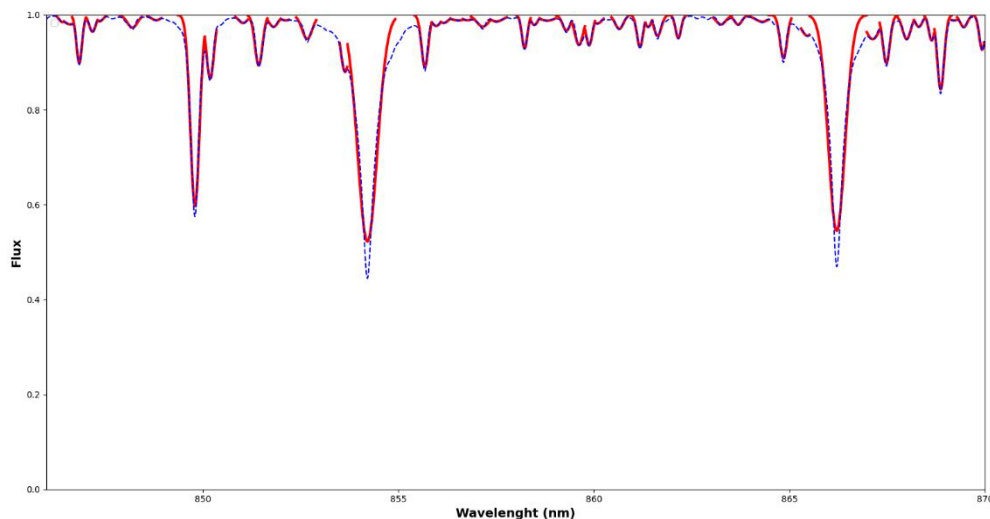
Fuente: Elaboración propia

Ya inicialmente se aprecia visualmente un ajuste mucho más fiable que el primero realizado con la función gaussiana con profundidad fija, corroborado por los *RMSE* que se han reducido a 0.02, 0.041 y 0.042 respectivamente. En las dos segundas líneas la mejora está presente, pero es prácticamente inapreciable, sin embargo, la primera línea, que más dificultad de ajuste había tenido, mejora significativamente reduciéndose el error casi un cuarto.

Con esta modificación, para el mecanismo de regresión simbólica, el valor de la variable de la profundidad que se le indique ya no será la profundidad de la línea en sí, sino esta más el offset calculado por el algoritmo de ajuste de SciPy.

Una vez afianzado el método de ajuste, se ejecuta el algoritmo en las 50 líneas atómicas seleccionadas, cuyo resultado se observa en la Figura 16.

Figura 16. Ajuste de gaussianas completo en espectro de *iSpec*



Fuente: Elaboración propia

El conjunto de las distribuciones gaussianas de todas las líneas atómicas presenta una versión bastante representativa del espectro original normalizado, dentro del alcance y las limitaciones del presente trabajo. Cómo se unen los espacios vacíos entre los ajustes de las líneas y el proceso para decidir a qué línea de absorción pertenecerá cada frecuencia (ya que,

a la hora de computar un espectro, los límites de cada línea de absorción serán desconocidos) se aborda preliminarmente con un set de líneas reducido en el siguiente apartado.

3.2.3.6. Síntesis de espectro preliminar

Antes de llegar a la regresión simbólica, se prueba la síntesis de los espectros en base a las gaussianas, ya que se ha de asegurar que el generado tiene un error mínimo con respecto al real. Tras la aplicación de la regresión, se hallará una desviación y una profundidad en función de los parámetros atmosféricos del espectro que se quiere sintetizar para poder ajustar una gaussiana (o intentarlo) a cada línea de absorción de la lista de VALD. Naturalmente habrá un error entre las desviaciones/profundidades reales para las líneas de espectro que se calcularon con el algoritmo recién explicado con respecto a las halladas por la regresión, y cuanto mayor sea ese error más diferirá del real el espectro generado ya que sus líneas de absorción serán distintas.

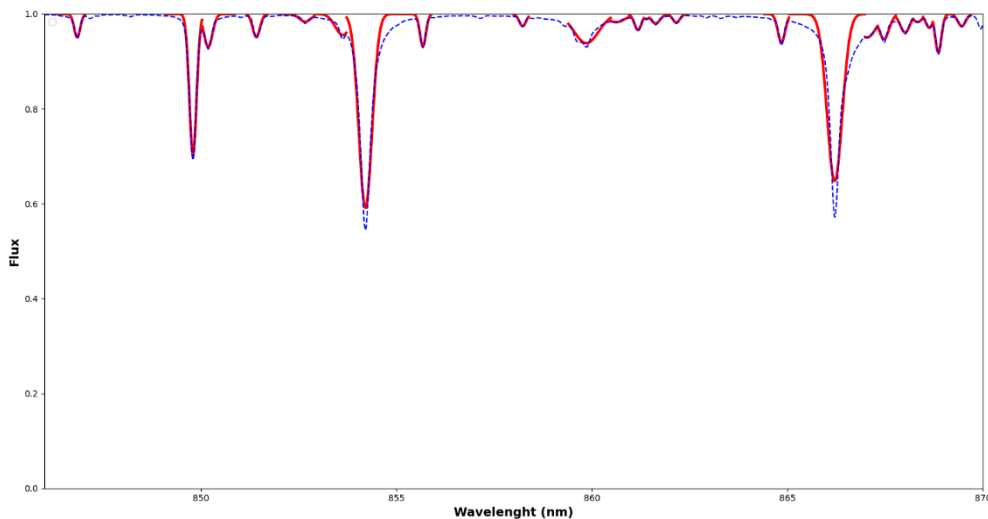
Lo que está claro es que para que el resultado tras la regresión sea mínimamente correcto, el espectro sintético generado al usar directamente los valores reales de las desviaciones y profundidades halladas por el algoritmo desarrollado ha de tener un error mínimo respecto al real. Esto es lo que se prueba preliminarmente en este apartado, computando el primer espectro usado en este capítulo con únicamente las 50 líneas que pudieron ser ajustadas y comparándolo con el sintético generado por iSpec.

El proceso de síntesis es el siguiente: con una precisión de 0.01nm, se ha de generar un flujo para cada frecuencia entre los 846 nm y los 870 nm. El dilema aquí es decidir a qué distribución pertenece cada frecuencia, o, mejor dicho, a qué línea de absorción a la cual se puede ajustar una distribución de Gauss pertenece la frecuencia. Para ello primero se establece un límite de comprobación, un rango de nm entre los cuales se computarán las distribuciones gaussianas de todas las líneas de absorción ubicadas en ese rango para la frecuencia cuyo flujo se está intentando hallar. Para decidir este límite se hace uso del mecanismo desarrollado para encontrar los límites de cada línea de absorción y se halla el de la línea de absorción con los límites más anchos encontrados, es decir, se encuentra la distancia máxima hallada entre una

frecuencia y el centro de su línea de absorción; el rango que se establezca no tendría sentido que fuera mayor que este límite por lo que se podría establecer como rango. En el caso del espectro de iSpec generado anteriormente este límite máximo es de 0.77 nm.

Hallar el límite máximo de un único espectro ajustado es naturalmente un método limitado para asegurar el rango que se ha de usar, por lo que se computa al menos un espectro adicional en iSpec con los parámetros de una estrella de tipo A7 (6780K, 4.29 log(g), -0.2 [M/H]) tomando como referencia Altair (Malagnini & Morossi, 1990) y se ajusta con el algoritmo desarrollado, observable en la Figura 17.

Figura 17. Ajuste de gaussianas en espectro sintético de estrella A7



Fuente: Elaboración propia

Este espectro tiene un límite máximo de 1.25. Para generalizar y asegurarse de que ninguna frecuencia no se pueda asociar a su frecuencia de absorción se establece el rango máximo de comprobación en 1.4 nm.

Para aclarar el algoritmo de síntesis, si por ejemplo se está computando el flujo para la frecuencia 865.55 nm, el proceso sería el siguiente:

1. Se recorren todas las líneas atómicas de la lista de VALD. Si la frecuencia de absorción de la línea está a 1.4 nm o menos de 865.55 nm (es decir, si está entre 864.15 nm y

875.95 nm) se calcula la fórmula de la distribución de Gauss para esa frecuencia con respecto a la línea de absorción en cuestión, usando la desviación típica y la profundidad hallada para esa línea. Si hay varias líneas en ese rango, se añaden todos los flujos calculados a una lista.

2. El flujo se decide en cuestión del número de flujos que haya en la lista:
 - a. Si ningún flujo es válido, es decir:
 - i. No hay líneas de absorción cercanas y no se ha calculado ninguno.
 - ii. Las líneas de absorción cercanas no se han podido ajustar a una distribución gaussiana con la desviación hallada (esto puede ocurrir si la desviación es 0 o casi 0).
 - iii. Se han hallado flujos superiores a 1 (valor del continuo).

Se establece 1 como el flujo para esa frecuencia.

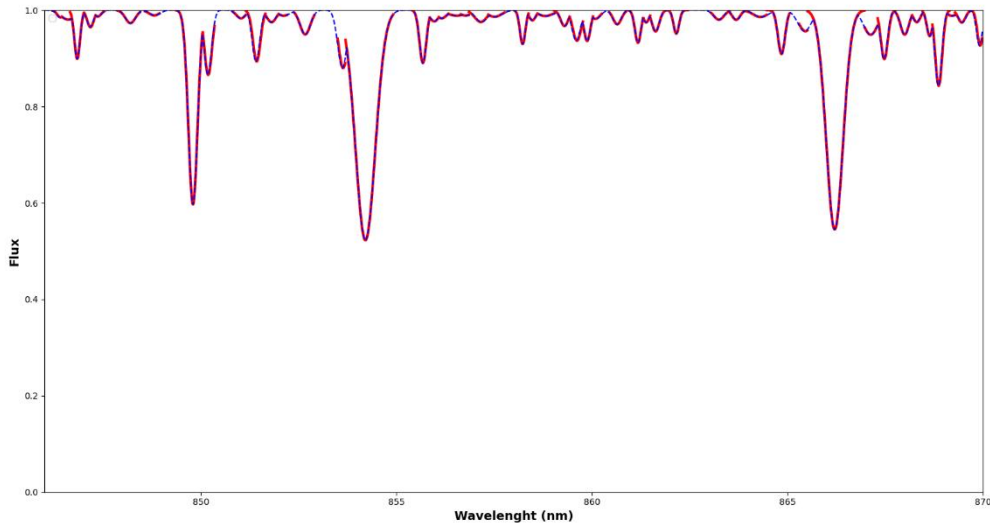
- b. Si hay un único flujo válido, se establece dicho flujo para esa frecuencia.
- c. Si hay varios flujos válidos en la lista, un análisis preliminar indica que el flujo más realista por lo general suele ser el más bajo, es decir, el más profundo³. Esto es porque al ser una gaussiana invertida cuánto mayor es el flujo, por lo general, más lejos se encuentra esa frecuencia del pico de absorción de la gaussiana y por tanto hay menos posibilidades de que pertenezca a la misma. A mayores, se quieren mantener los flujos que más contribuyan a la absorción del espectro ya que son los que definen el mismo. Por tanto, de los flujos hallados se selecciona el menor (más profundo).

Aplicando este algoritmo con las desviaciones y profundidades halladas para el espectro inicial de iSpec (Figura 5) se obtiene el observable en la Figura 18 (flujo generado correspondiente

³ El resultado del mencionado análisis se aprecia en la Figura 18, donde se ven todas las distribuciones gaussianas computadas en cada línea de absorción (línea roja) contra los flujos seleccionados (línea azul discontinua). Se observa cómo el método de selección del flujo más profundo encuentra correctamente la forma más similar posible al espectro real.

con la línea azul discontinua), con un RMSE de 0.016 con respecto al original normalizado, lo que se considera cómo un ajuste válido y relativamente realista.

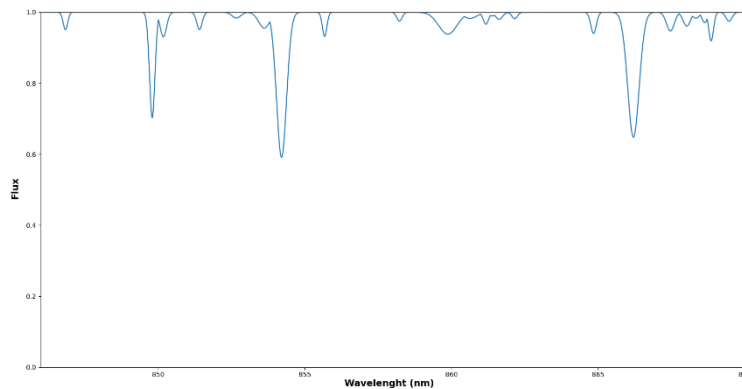
Figura 18. Espectro inicial sintetizado sin regresión contra gaussianas halladas



Fuente: Elaboración propia

Se ejecuta el algoritmo otra vez sobre los parámetros hallados para el segundo espectro sintético de iSpec de una estrella de tipo A7 (Figura 17) y se produce el visible en la Figura 19, con un RMSE de 0.015 con respecto al de iSpec.

Figura 19. Segundo espectro inicial sintetizado sin regresión



Fuente: Elaboración propia

Como era de prever, sin aplicar la regresión y usando los parámetros gaussianos reales calculados la síntesis es decente, con un error bastante pequeño y un resultado similar al espectro real. El siguiente paso es sintetizar estos espectros, pero usando las desviaciones y profundidades calculadas por las funciones que se han de construir mediante regresión simbólica, y reevaluar el error en este caso, que ha de permanecer pequeño ya que cuando se ejecute la regresión con miles de espectros el error en las funciones va a ser naturalmente mayor y por tanto el relativo a sintetizar espectros.

3.2.4. Regresión simbólica

3.2.4.1. Datos de entrada y salida

Se prueba inicialmente el mecanismo de regresión simbólica con dos espectros de entrenamiento ajustado, el generado por iSpec en este subapartado más uno adicional, con objeto de familiarizarse con la herramienta de regresión elegida, PySR (Sección 2.3), y de sentar las bases para el proceso que luego se ha de realizar de manera automática con cientos de espectros sintéticos de entrenamiento.

Lo primero que se ha de concretar, antes de poner el algoritmo en marcha, son los datos: por cada función se debe dar una lista con parámetros de entrada y valores de salida para esos parámetros, con el objetivo de que el algoritmo vaya ajustando las fórmulas para conseguir una mejor relación global.

El algoritmo de regresión simbólica ha de generar dos funciones por cada línea de absorción, una que represente la profundidad de la distribución gaussiana en la que se encuentra la frecuencia de la línea y otra representando la desviación típica de la distribución, permitiendo así obtener una posible distribución gaussiana por cada línea. Estos valores, que permitirán generar las distribuciones que forman el espectro completo, han de depender de los tres parámetros elegidos con anterioridad: temperatura, logaritmo de la gravedad y metalicidad.

Por tanto, teniendo en cuenta que la lista de líneas atómicas elegida y filtrada se compone de 176 líneas atómicas, y que se han de hallar dos funciones por línea, el trabajo del algoritmo de regresión simbólica será computar 352 funciones.

El proceso para hallar cada función es el siguiente:

- Se va recorriendo cada línea de las 176 de la lista.
- Se genera la primera función de la línea, que representa la profundidad de la distribución gaussiana:
 - Al algoritmo se le han de proporcionar dos listas, cada índice de cada lista representando los datos de un espectro de entrenamiento distinto.
 - En la primera lista (entrada), cada índice contendrá otra lista con los tres parámetros atmosféricos del espectro correspondiente al índice, es decir: los parámetros de entrada de la función para ese espectro, cuyo resultado se quiere que se ajuste al valor en el mismo índice, pero de la segunda lista.
 - La segunda lista (salida) contendrá el valor de las profundidades previamente halladas (si es que se ha podido ajustar una distribución gaussiana en esa línea de absorción para ese espectro), cada índice representando la profundidad hallada para cada espectro de entrenamiento cuyos parámetros se encuentran en la misma posición de esta lista, pero en la primera.
 - Así, se entrega al algoritmo una cantidad de ejemplos de entrenamiento que indican una serie de valores que se deberían hallar con unos respectivos parámetros de entrada, permitiéndole computar una función que mejor ajuste estas relaciones.
- Se genera la segunda función de la línea, que representa la desviación típica de la distribución gaussiana:
 - Mismo procedimiento que para la primera función. Se generan dos listas, cada índice representando un espectro de entrenamiento, la primera lista con los parámetros atmosféricos del espectro y la segunda con el valor de la desviación típica de la distribución de Gauss en la línea atómica en cuestión, si es que se ha podido ajustar.

El principal problema en esta parte de asignación de datos es el que se puede intuir leyendo el procedimiento, y es que el valor de las profundidades y desviaciones típicas de cada línea solo se podrá establecer correctamente si se ha podido ajustar una distribución de Gauss en

la misma en el espectro de entrenamiento, lo que ya se ha apreciado en el ajuste hecho en el subapartado previo que no es especialmente común: en el ajuste del espectro inicial sintético de iSpec, solo se pudieron ajustar 50 líneas de las 176 de la lista.

La idea preliminar fue no incluir los datos de espectros para los cuales no se hubiera podido ajustar una distribución Gaussiana para las funciones de la línea en cuestión con la que se esté trabajando. Sin embargo, es lógico que este enfoque no representaría con fidelidad la relación entre los parámetros de los espectros y esas líneas (el objetivo del presente trabajo), ya que si no se ha podido ajustar es porque esa frecuencia de absorción o no representa absorción ninguna en el espectro o forma parte de una más grande (y por tanto no se tiene en cuenta, ya que como se explicó en la Sección 3.2.3.2, solo se tienen en cuenta frecuencias de absorción que forman distribuciones independientes de Gauss), por lo que esta situación se ha de tener en cuenta: que para esos parámetros la línea no se debería poder ajustar al computar un espectro sintético, o al menos no debería tener relevancia ni contribuir al flujo. A su vez, toda función de cada línea debe tener ser entrenada con el mismo número de espectros para conseguir luego un resultado homogéneo y realista al computar todas las distribuciones de Gauss en cada línea para generar el espectro: cada función generada debe tener la misma contribución de cada espectro.

En conclusión, el único enfoque válido es dar un valor a ambos resultados incluso cuando no se ha podido ajustar una distribución gaussiana en esa línea. Para abordar la decisión de qué valores dar se revisa la ecuación de la distribución de Gauss aplicada (Ecuación 1) y se decide en base a los efectos sobre la curva establecer los siguientes valores:

Para la variable de la profundidad (D) es simple: un valor 0 indica que no hay profundidad en esa distribución, y el resultado de aplicar la distribución en cualquier frecuencia será siempre 1, es decir, el continuo. La desviación típica (σ), aunque no tan simple, decide ajustarse a 0 también, ya que la mayoría de las desviaciones halladas rondan ese valor por lo que no ensuciará en demasía el trabajo del algoritmo desviando los otros valores, y a su vez, una desviación 0 indica que esa distribución es incalculable, no se puede hallar, por lo que representa lo que se quiere que represente, que esa línea para esos parámetros no tiene

ajuste gaussiano. Resumiendo: ambos parámetros se igualan a 0 en las líneas sin ajuste hallado, pero independientemente se entregan al algoritmo para que los tenga en cuenta.

Aclarados los datos que se le entregan al algoritmo, lo siguiente antes de ponerlo en ejecución es concretar cómo se configura.

3.2.4.2. PySR

PySR proporciona multitud de opciones y parámetros para personalizar el funcionamiento del algoritmo: iteraciones, ciclos por iteración, procesadores, poblaciones, peso de los datos, restricciones, complejidad, operadores... Es una herramienta compleja que requiere de conocimientos avanzados y práctica para sacar el máximo provecho a la misma. Teniendo en cuenta el alcance del trabajo, que ya se encuentra en un punto avanzado, que llevaría mucho tiempo familiarizarse y entender el funcionamiento al completo de la herramienta, y que la falta de recursos no permitiría aprovechar la potencia completa de la misma, se decide usar una configuración por defecto propuesta por PySR que se comprueba que no consume una cantidad exorbitada de recursos y tiempo y que muestra resultados competentes.

Con falta de recursos se hace referencia al hecho de que la regresión simbólica es un algoritmo que encuentra soluciones a base de fuerza bruta, intento tras intento, es decir, que consume una mayor cantidad de tiempo y de recursos del sistema que otras técnicas de aprendizaje automático. Para hallar las funciones objetivo del presente trabajo, se habrá de entrenar el algoritmo con una cantidad elevada de espectros, y se ha de tener en cuenta que el algoritmo ha de descifrar dos funciones por cada línea de absorción de cada espectro. El equipo con el que se entrenará el algoritmo no es profesional, tiene importantes limitaciones, por lo que se ha de facilitar el proceso todo lo posible para obtener resultados.

La configuración elegida de PySR para la creación de funciones cuenta con la siguiente parametrización:

- 16 poblaciones, ya que para aprovechar el máximo de la potencia de la herramienta se deben añadir dos por *core* del procesador del equipo, teniendo así siempre una en ejecución.

- 500 ciclos por iteración, cada ciclo considera 10 muestras de ecuaciones, mutando cada una una vez.
- 40 iteraciones.
- Condición para parar la ejecución: si el error entre los valores reales y producidos se reduce lo suficiente o la complejidad asciende demasiado (10 niveles), el algoritmo se parará y guardará los resultados, sin necesidad de completar las iteraciones.
- Se asigna también un tiempo máximo de 5 minutos por función, que, en total, si se llegara a dicho tiempo límite, tardaría el algoritmo 29 horas de ejecución en generar las 352 funciones.
- Un tamaño máximo de ecuación de 50, contando operadores, constantes y variables. Se usa para reducir el tiempo de ejecución y evitar funciones demasiado grandes.
- Operadores binarios: se eligen los operados básicos para que puedan ser usados en las ecuaciones (suma, resta, multiplicación y división).
- Operadores unitarios: tomando como referencia la recomendación de PySR, se eligen cuadrados, cubos, cosenos, exponentes, senos e inversos respecto del producto ($1/x$).

Configurada la operativa, se ejecuta el algoritmo sobre las 352 funciones con dos espectros de entrenamiento, el ajustado en el subapartado anterior (5837K, 4.06 log(g), 0.24 [M/H]) más uno adicional generado y ajustado con los parámetros de una estrella de tipo A7 (6780K, 4.29 log(g), -0.2 [M/H]) tomando como referencia Altair (Malagnini & Morossi, 1990), cuyo ajuste se puede observar en la Figura 17, ya que el algoritmo de regresión carece de sentido si se hace con solo un ejemplo por función. Cada función por tanto tiene únicamente dos fuentes de entrenamiento para construirse, tres parámetros de entrada y un valor de salida por dos, por lo que la ejecución es bastante rápida y en unos pocos minutos se computan y guardan las funciones.

Como ejemplo se muestra la función hallada para la línea de absorción del Ca II, en la longitud de onda de 854.21 nm, que es la de mayor profundidad en ambos espectros. Los dos ajustes previos para cada espectro son visibles en la Figura 16 y la Figura 17 respectivamente. Dichos ajustes hallaron unas profundidades de 0.477 y 0.408 y unas desviaciones típicas de 0.252 y

0.169 respectivamente. Las funciones halladas para la profundidad y desviación, en orden, son las siguientes:

$$D = \cos(\cos(T_{eff} + 1.7227046)) - 0.16172126 \quad (5)$$

$$\sigma = (-0.6204483 / (\text{square}(\exp(\log g)) - T_{eff})) - 0.00150323 \quad (6)$$

Se puede apreciar cómo debido a tener únicamente dos ejemplos de entrenamiento y principalmente con ajustes muy similares en esta línea de absorción, el algoritmo es capaz de crear la fórmula para la profundidad con únicamente la temperatura como variable de entrada y para la desviación con la temperatura y el logaritmo de la gravedad, pudiendo prescindir en ambos casos de la metalicidad. El algoritmo realiza este descarte por el límite de parada basado en el error establecido, y es que aplicando estas ecuaciones a los parámetros de entrada de ambos espectros se hayan las respectivas profundidades y desviaciones prácticamente idénticas, con *RMSE* para los cuatro valores todos inferiores 10^{-5} .

Como era de esperar, el algoritmo funciona bien con únicamente dos espectros de entrenamiento, la incógnita será ver su eficiencia con cientos de ellos.

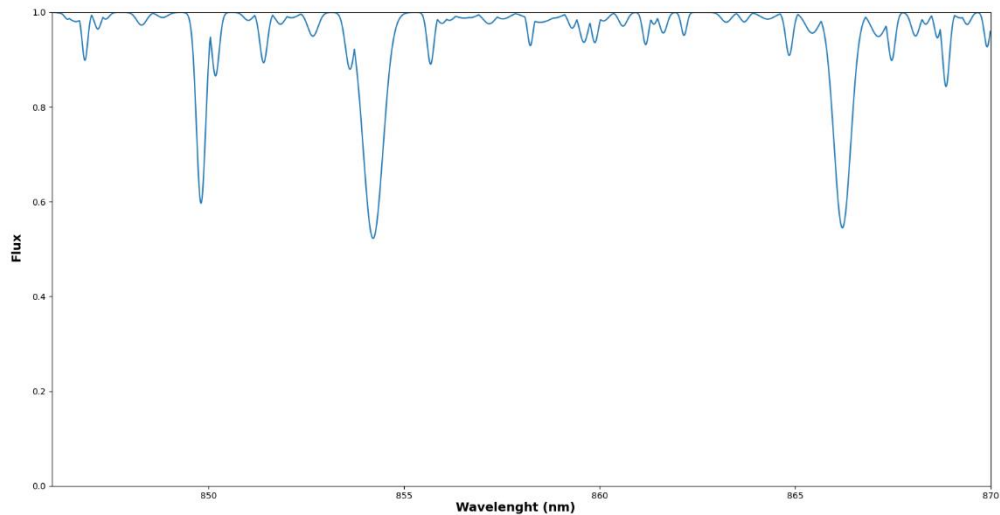
3.2.5. Síntesis del espectro

Una vez halladas las ecuaciones para cada línea de absorción en función de los parámetros, el siguiente paso es volver a seguir el procedimiento desarrollado en la Sección 3.2.3.6 para generar espectros sintéticos.

En la Sección 3.2.3.6 no se usaron ecuaciones de la regresión sino directamente los valores gaussianos calculados por la herramienta de ajuste desarrollada, para probar simplemente el mecanismo de síntesis. Ahora se ha de ejecutar de forma auténtica con las ecuaciones añadidas para probar su efectividad real, proporcionando a las ecuaciones los mismos parámetros que crearon el primer espectro sintético con el que se ha trabajado principalmente en este capítulo (5837K, 4.06 log(g), 0.24 [M/H]), para poder comparar el sintético generado con el previo generado por *iSpec* y valorar así la eficiencia del algoritmo.

Las ecuaciones producidas van a permitir hallar tanto una profundidad como una desviación típica para cada línea de absorción de la lista de VALD, en función de los tres parámetros atmosféricos establecidos. Por lo que, el proceso de síntesis es el mismo que el explicado en la Sección 3.2.3.6 pero previamente usando las funciones halladas por la regresión para computar las desviaciones y profundidades. Dicha síntesis produce el espectro observable en la Figura 20, con un RMSE con respecto al original generado por iSpec de 0.017. El RMSE al ejecutar la síntesis sin uso de las funciones (Figura 18), con los valores gaussianos calculados, fue de 0.016, por lo que con las funciones de la regresión el error únicamente ha aumentado 0.001.

Figura 20. *Espectro inicial sintetizado con regresión*



Fuente: Elaboración propia

3.3. Entrenamiento

El apartado de aproximación previo presentó un ciclo de entrenamiento el cual se iniciaba ajustando perfiles gaussianos a espectros, hallando una desviación típica y una profundidad por cada línea de absorción ajustable en dicho espectro, y posteriormente aplicando regresión simbólica sobre los parámetros de cada línea de absorción entre los parámetros atmosféricos del espectro, hallando funciones que relacionan ambos sets de datos. Dicho ciclo se validó sintetizando un espectro con las funciones halladas y se comprobó que la eficiencia era correcta, produciendo un espectro con un error aceptable con respecto al original.

Sin embargo, el ciclo se aproximó usando únicamente dos espectros de entrenamiento, lo que conlleva que la regresión tuvo únicamente dos fuentes por cada función, facilitando, naturalmente, en gran medida el proceso para hallar funciones adecuadas. En este apartado, el entrenamiento se debe amplificar enormemente, debido a que la regresión ha de tener en cuenta las relaciones entre los parámetros gaussianos para unos parámetros atmosféricos que abarquen la totalidad del rango de estrellas capturas por el espectrómetro RVS de Gaia, cuyos espectros se han de usar para validar el entrenamiento.

El apartado se divide en los tres entrenamientos completos que se realizaron con regresión simbólica y se describe el proceso previo seguido en cada uno mejorando y perfilando el anterior, hasta terminar con las funciones finales del trabajo presente.

3.3.1. Primer entrenamiento

El diagrama de Kiel del espectrómetro RVS de Gaia (Figura 2) muestra, como ya se mencionó, que las estrellas abarcan un rango aproximado de temperatura entre 3800K y 8000K, con logaritmos de gravedad variando entre 0.3 $\log(g)$ y 4.8 $\log(g)$, y metalicidades entre -1.1 [M/H] y 0.25 [M/H]. Como ya se indicó también, por varias razones, el algoritmo de regresión se entrena con espectros sintéticos de iSpec, permitiendo así que espectros computar exactamente y sus parámetros asociados.

Para intentar abarcar el máximo rango mostrado en el diagrama de Kiel sin producir un número exorbitante de espectros, cuyos datos producidos llevarían mucho tiempo y recursos

para el algoritmo de regresión, de los que no se disponen, se determina generar los espectros sintéticos con los siguientes parámetros:

- De 3800K a 8000K, en saltos de 300, por lo que se producen espectros para 15 temperaturas distintas.
- De 0.3 log(g) a 4.8 log(g), en saltos de 0.3, por lo que se producen espectros para 16 gravedades distintas.
- De -1.1 [M/H] a 0.25 [M/H], en saltos de 0.15, por lo que se producen espectros para 9 metalicidades distintas.

El total de combinaciones entre todos los parámetros generarían un total de 2400 espectros sintéticos, sin embargo, no todas las combinaciones son válidas. Previo a la generación del espectro, iSpec tiene que generar una atmosfera interpolada en base a los tres parámetros escogidos (el modelo atmosférico seleccionado, así como el resto de parámetros de generación, son los mismos que se decidieron en la Sección 3.2.1), y si no es capaz de interpolarla en el caso de que los valores no sean compatibles automáticamente los modifica para generar el espectro. En el caso de que se produzca el error de interpolación, se decide no dejar que iSpec compute el espectro y descartar esa combinación de parámetros, evitando así datos alterados para el entrenamiento.

Este error de interpolación se produce en 270 combinaciones, por lo que finalmente se tienen 2130 espectros sintéticos para entrenar el algoritmo de regresión simbólica.

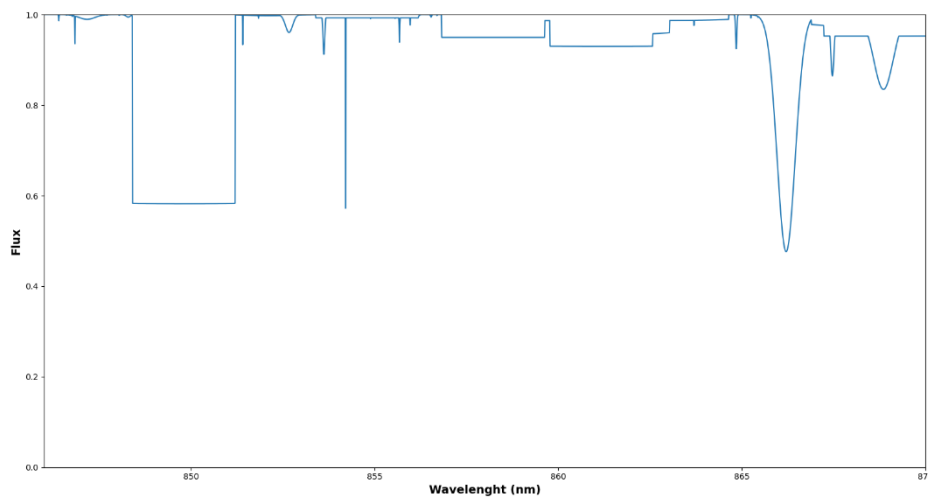
Aunque ya explicado en detalle en la Sección 3.2, se resumen aquí los pasos realizados por el proceso automatizado desarrollado para hallar las funciones, una vez se tienen los datos crudos de los espectros sintéticos.

1. Se tratan los datos crudos, extrayendo únicamente el rango de frecuencias entre los 846 nm y los 870 nm y redondeando valores tanto de frecuencias como de flujo.
2. Con el algoritmo de ajuste gaussiano desarrollado, por cada espectro se seleccionan las líneas atómicas de la lista de VALD que se pueden ajustar y se ajustan, almacenando la desviación y la profundidad hallada para cada línea, así como los parámetros del espectro cuyo ajuste a producidos esos datos.

3. Las listas con todas las desviaciones y profundidades, representando valores de salida, se importan al mecanismo de regresión simbólica junto con las listas de parámetros, representando los valores de entrada. Este mecanismo computa 352 funciones, dos para cada una de las 176 líneas de absorción de la lista de VALD.

Una ejecución preliminar de 3 horas del algoritmo de la regresión simbólica con los 2130 espectros sintéticos produjo unos resultados muy pobres, con errores muy grandes en las funciones y una síntesis posterior muy deficiente, en la cual no se podían apreciar las líneas de absorción. Un ejemplo se puede observar en la Figura 21, para la cual se generó el espectro sintético correspondiente a una estrella similar al Sol, de 5900K, $3.9 \log(g)$ y 0.1 [M/H]. El espectro formado no parece si quiera un espectro electromagnético estelar, con un RMSE enorme respecto al original de 0.146, y una síntesis claramente no válida y deficiente.

Figura 21. *Síntesis con primera generación de funciones por regresión*



Fuente: Elaboración propia

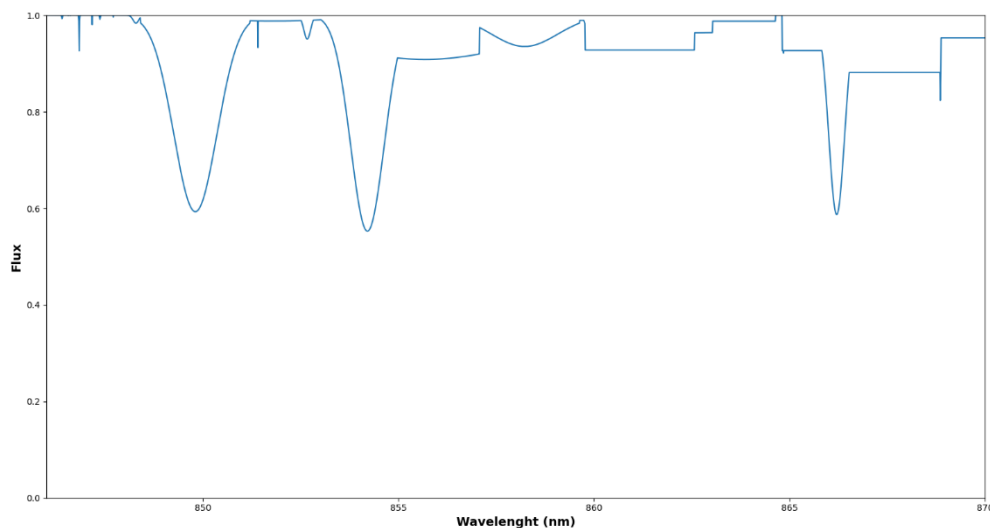
3.3.2. Segundo entrenamiento

En una segunda iteración se modifican algunos parámetros del algoritmo con el objetivo de conseguir que los resultados del entrenamiento mejoren, intentando obtener al menos espectros válidos. En particular, se aumenta el número de iteraciones a 100 y a su vez el límite

de parada por complejidad de las funciones de 10 niveles a 20, alcanzando el límite computacional de este trabajo, limitado por recursos y tiempo. La duración del algoritmo con 100 iteraciones para hallar las 352 funciones fue superior a 8 horas, consiguiéndose al menos disminuir el error de dichas funciones.

Se resintetiza el espectro que se computó con las funciones iniciales para comprobar así la mejoría alcanzada al aumentar el tiempo de entrenamiento de la regresión. El espectro sintetizado se observa en la Figura 22, con un RMSE calculado respecto al flujo de iSpec original de 0.078, consiguiendo reducir a la mitad el error con respecto a la síntesis anterior.

Figura 22. *Síntesis con segunda generación de funciones por regresión*



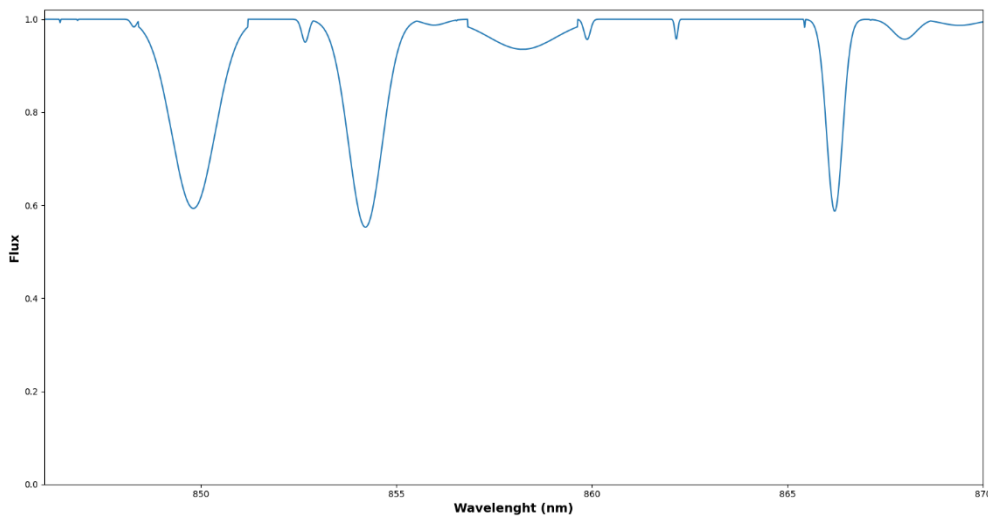
Fuente: Elaboración propia

Sin embargo, es claro que la síntesis aún dista mucho de ser válida. Gran parte del espectro está formado por líneas horizontales y verticales, lejos de ser distribuciones de Gauss que es, teóricamente, de lo único de lo que debería estar formado el espectro. Para ayudar a localizar el origen de esta fallida síntesis, incluso con el amplio tiempo de entrenamiento proporcionado a la regresión simbólica, se computan todas las distribuciones de Gauss de todas las líneas atómicas para el espectro recién sintetizado (de todas con un ajuste computable, ya que algunas tienen una desviación muy cercana a 0 por lo que no son

verticales) y a aquellas que superen una distancia a 0 de 1 (es decir, mayor que 1 o menor que -1, ya que produce ajustes prácticamente horizontales).

Con esta nueva iteración del algoritmo, de las 176 líneas atómicas de la lista de VALD para el espectro con parámetros 5900K, 3.9 log(g) y 0.1 [M/H] se descartarían 152, dejándose únicamente 24 para formar el espectro, cuya nueva síntesis se puede apreciar en la Figura 24.

Figura 24. *Síntesis filtrada con segunda generación de funciones por regresión*

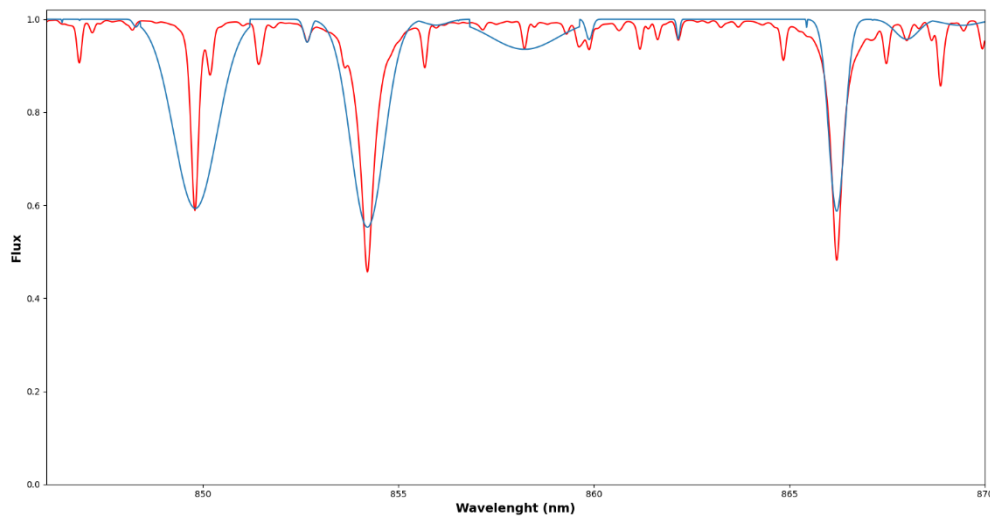


Fuente: Elaboración propia

Esta mejora del algoritmo computa un espectro con un RMSE de 0.071 con respecto al original, una ligera ventaja en error numérico en comparación con la anterior solución sin descarte de ajuste erróneos, sin embargo, la principal ventaja, claramente visible, de esta solución es que el espectro parece un espectro real de absorción, al contrario que los dos primeros generados con regresión.

Aun así, es evidente que el trabajo de la regresión simbólica no es eficiente, comparando la última síntesis con el espectro original (Figura 25) se observa como las funciones no están proporcionando valores correctos ni de desviaciones ni de profundidades, alterando las anchuras de las líneas más representativas y directamente siendo incapaz de hallar las menores.

Figura 25. Comparación de síntesis para segunda generación de funciones



Fuente: Elaboración propia

3.3.3. Tercer entrenamiento

Con el objetivo de intentar perfilar algo más el algoritmo y los datos de entrada y por probar a ejecutar un último entrenamiento se realiza un análisis final de los factores que pueden estar provocando la poca eficiencia de la regresión simbólica para encontrar funciones adecuadas que relacionen los parámetros con las desviaciones y profundidades.

Un factor sin duda de importancia, que desvía en gran medida las funciones, es que, para aquellos espectros con líneas atómicas no ajustables, se proporciona una desviación y profundidad de 0 (Sección 3.2.4.1), valores que la regresión está teniendo en cuenta para ajustar sus funciones pero que sin embargo no debería tener en cuenta, ya que no se ha podido ajustar una gaussiana en esa línea para ese espectro. Sin embargo, cómo se mencionó cuando se describió el porqué de la decisión, el mecanismo de regresión ha de tener en cuenta para generar la función de alguna manera que para esos parámetros no se puede ajustar esa línea, y los valores propuestos son los únicos que tienen sentido para representar dicha condición.

Un mayor tiempo de ejecución de PySR, es decir, aumentar las poblaciones, ciclos e iteraciones en el entrenamiento, proporcionarían muy probablemente unas mejores funciones y un menor error, como ya se demostró cuando se aumentó el número de iteraciones y por tanto se incrementó el tiempo de entrenamiento de 3 horas a 8 horas. Sin embargo, esta posibilidad no se contempla debido al ámbito de investigación amateur de máster del trabajo, en el que se cuenta con un sistema de recursos limitados para la ejecución de los algoritmos que además no puede estar disponible en su totalidad para el desarrollo presente.

Un factor clave, probablemente incluso más que las configuraciones de la propia regresión son los datos de entrada del mismo. Y el análisis presente encuentra que al observar el diagrama de Kiel del espectrómetro RVS de Gaia (Figura 2), cuyos valores fueron tomados de referencia para computar los espectros sintéticos de entrenamiento, muchas combinaciones de valores no tienen estrellas asociadas y sin embargo para el entrenamiento se computaron espectros para prácticamente la totalidad de combinaciones paramétricas posibles (excepto 270 descartados debido a que iSpec no pudo interpolar sus atmósferas). Por ejemplo, se aprecia que prácticamente no hay espectros de estrellas con baja gravedad y alta temperatura. Esto quiere decir que se están por tanto usando espectros de más para el entrenamiento, con combinaciones paramétricas no existentes en la realidad, o al menos no observadas por el RVS.

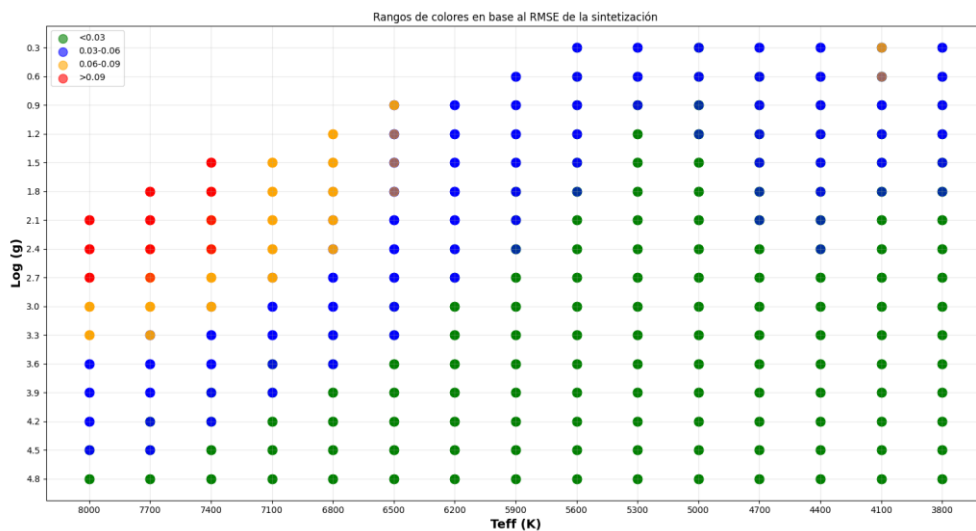
Por ello, como último intento, se deciden filtrar los espectros sintéticos generados por iSpec y usar un bloque más estricto, quedándose con los espectros con mejores ajustes, y volver a entrenar la regresión simbólica con éstos esperando que al eliminar el desecho provocado por los ajustes incorrectos las distribuciones gaussianas se desvíen menos y se mantengan más fieles a cómo deberían ser.

Para ello, se decide sintetizar cada espectro sintético de entrenamiento de iSpec directamente con los valores de desviaciones y profundidades hallados por el algoritmo de ajuste gaussiano desarrollado en el presente trabajo (Sección 3.2.3.5), tal y como se hizo en la Sección 3.2.3.6. Con la síntesis computada, se compara con el espectro original hallando el RMSE para cada

síntesis. Este proceso de síntesis y comparación dura unos 18 segundos en el sistema de trabajo para cada espectro, y se ha de realizar para los 2130 de entrenamiento originales, por lo que el proceso completo tiene una duración total de más de 10 horas y media.

Habiendo calculado los RMSE de todas los espectros recién generados en comparación con los sintéticos de entrenamiento de iSpec, se consigue una puntuación directa de la eficiencia del ajuste gaussiano generado para cada espectro en cuestión, y se usan solo aquellos con las mejores sintetizaciones para volver a entrenar el algoritmo de regresión una última vez. Los RMSE hallados se pueden apreciar en la Figura 26.

Figura 26. Errores de síntesis para todos los espectros de entrenamiento



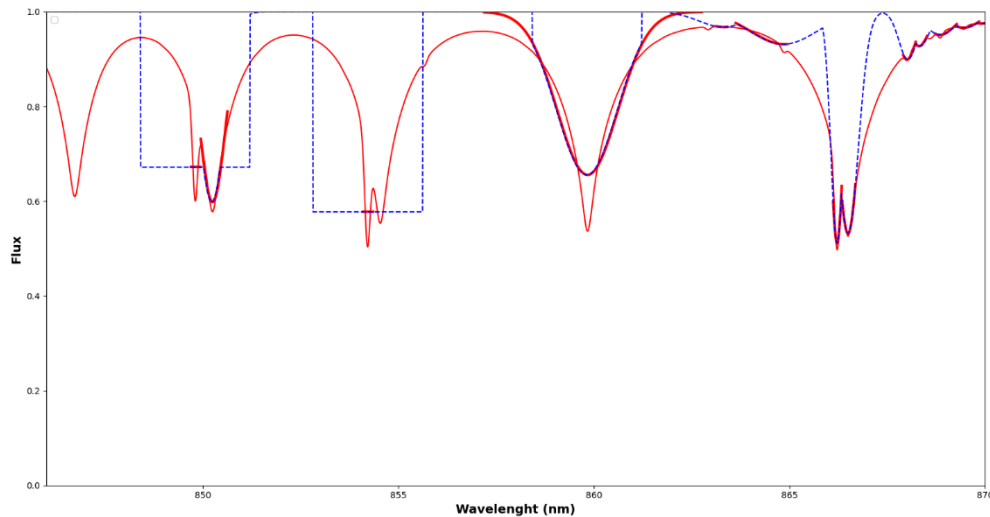
Fuente: Elaboración propia

Se aprecia ya de antemano que los 270 espectros que fueron descartados por los errores de interpolarización atmosférica son aquellos con baja gravedad y alta temperatura, que no eran tampoco observables en el diagrama de Kiel del espectrómetro RVS de Gaia (Figura 2). Los sintetizadores de espectros, como iSpec, usan cientos de datos espectrales reales para sus computaciones, por lo que es lógico que, para este tipo de estrellas poco comunes, con pocos datos reales de entrenamiento, las sintetizaciones sean pobres y por tanto el algoritmo

desarrollado de ajuste gaussiano tenga problemas para ajustar las absorciones a distribuciones.

Dicho problema se observa perfectamente en, por ejemplo, el espectro sintético de iSpec para una estrella de 8000K y 2.1 log(g), en la Figura 27, representado en color azul.

Figura 27. Espectro sintético erróneo de iSpec y ajuste



Fuente: Elaboración propia

El ajuste realizado por el algoritmo desarrollado se observa en rojo, que naturalmente no ajusta bien las distribuciones gaussianas ya que el propio espectro sintético original no está compuesto por absorciones en forma de distribuciones gaussianas. Se esperaba que iSpec informara de errores de interpolaciones para estas combinaciones tan extremas con sintetizaciones anormales, pero no es así, el programa produjo numerosos espectros erróneos que se usaron para entrenar al algoritmo de regresión simbólica, con lo que se descifra lo que fue el principal factor desencadenante de los dos primeros entrenamientos fallidos y de las funciones que se obtuvieron.

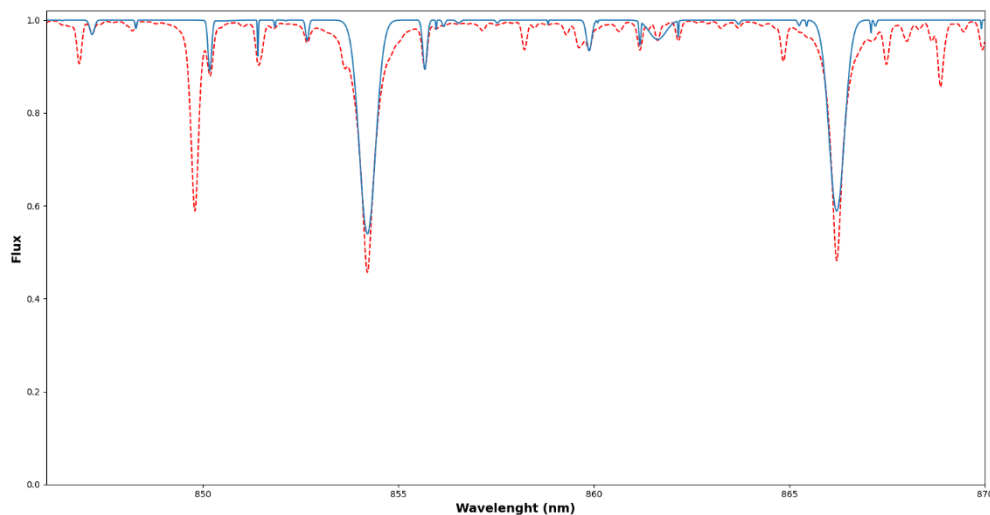
Se decide por tanto quedarse únicamente con aquellos espectros sintéticos cuyo RMSE calculado es menor a 0.03, es decir, los puntos verdes en la Figura 26. Con esto permanecen

para el entrenamiento 1412 espectros sintéticos, habiendo limpiado 718 de iSpec con sintetizaciones consideradas no adecuadas.

Con este set refinado se vuelve a ejecutar el algoritmo de regresión simbólica con 100 iteraciones. El tiempo de entrenamiento para computar las 352 funciones es de aproximadamente 7 horas.

Se vuelve a validar de forma preliminar la efectividad del algoritmo con los mismos parámetros que se usaron para validar en los dos primeros entrenamientos, 5900K, 3.9 log(g) y 0.1 [M/H], pudiendo así comparar fácilmente la mejora obtenida. El ajuste es visible en la Figura 28, en color azul, comparado con el espectro original en rojo.

Figura 28. Comparación de síntesis para tercera generación de funciones



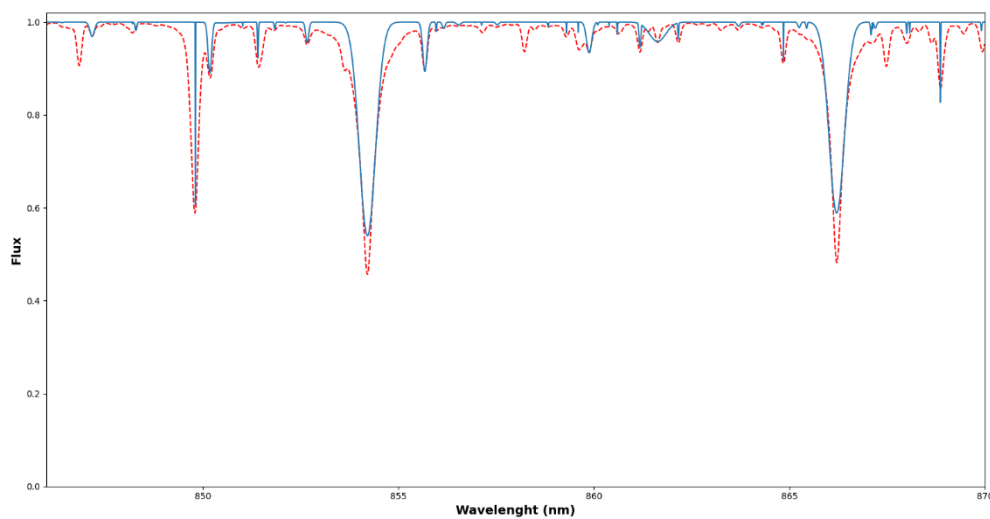
Fuente: Elaboración propia

El RMSE obtenido es ahora de 0.047, de nuevo se consigue una mejora sustancial respecto al ajuste obtenido por el entrenamiento previo, sin limpieza de espectros de entrenamiento, que obtuvo un RMSE para estos parámetros de 0.078. Incluso con la visible mejora, se aprecia que la generación a partir del ajuste y por tanto la eficiencia de las funciones de la regresión sigue sin poder ser considerada del todo correcta, ya que, aunque las líneas atómicas que son

visibles en la síntesis son bastante fieles a la reales (profundidades y anchuras similares, alguna consiguiendo un ajuste prácticamente idéntico) se pierden numerosas líneas de absorción.

Si se disminuye el límite de inferior establecido para la desviación típica en la síntesis (establecido en que si la distancia a 0 era menor de 0.01 se descartaban para evitar líneas verticales) a 0.001, es decir, solo descartar aquellas líneas cuyas desviaciones estén entre 0.001 y -0.001, se obtiene el ajuste de la Figura 29.

Figura 29. Comparación de síntesis con líneas verticales en tercera generación de funciones



Fuente: Elaboración propia

Con un RMSE prácticamente igual (ligeramente inferior) de 0.046, lo que se puede apreciar gracias a disminuir el límite es que muchas de las líneas de absorción perdidas están ahí, con una profundidad bastante similar a la real, pero con una desviación calculada por las funciones de la regresión muy desviada, que provoca que las líneas acaben siendo prácticamente verticales y por tanto irrealistas. Esta desviación está sin duda provocada por aquellas líneas no existentes en ciertos espectros a las que se les da un valor 0 de desviación para que se tengan en cuenta en el entrenamiento, lo que conlleva que para aquellas líneas solo existentes en ciertos espectros el algoritmo vaya a tender a desplazar el valor de la desviación hacia 0. Sin embargo, cómo se ha mencionado, no se encuentra otro método lógico de tratar estas

líneas no existentes para que se tengan en cuenta en la regresión, por lo que se considera que es un problema inherente del uso de la regresión simbólica en este desarrollo.

Se decide mantener estas líneas para la validación, ya que en su mayoría representan absorciones existentes en el espectro original, aún con una anchura incorrecta, y por tanto se considera que aportan valor a la síntesis. Para ello el límite inferior para la desviación típica se establece en 0.001.

Con esta última iteración se da por concluido el proceso de desarrollo de obtención de espectros sintéticos a partir de parámetros estelares por regresión simbólica, habiendo alcanzado unos resultados preliminares no óptimos pero que se consideran más o menos aceptables dada la naturaleza del trabajo y los límites de tiempo y de recursos. Las funciones, aun teniendo una amplia capacidad de mejora, parece que son capaces de obtener las líneas de absorción más significativas de los espectros estelares con una precisa profundidad.

En el siguiente capítulo se muestran cómo ejemplo alguna de las funciones finales halladas y se valida su eficacia computando espectros sintéticos y comparándolos con aquellos reales del espectrómetro RVS de Gaia.

4. Resultados

En este capítulo se muestran y describen algunas de las funciones finales encontradas, que representan el resultado del trabajo, y se valida la eficiencia de las mismas con datos de espectros reales de Gaia.

4.1. Funciones

Con una función de desviación y profundidad para ajustes gaussianos por cada línea de absorción de las 176 de la lista de VALD, se ha de probar la efectividad de estas mismas en un caso de uso real. Antes de la validación final se muestran algunas de las funciones más relevantes generadas por la regresión en la Tabla 1⁴, relativas a las frecuencias de absorción más predominantes y referenciadas a lo largo de este trabajo. Una lista completa con las 352 funciones halladas para las 176 líneas de absorción se encuentra en el Anexo A.

Tabla 1. Funciones finales para las líneas más relevantes

Elemento	Frecuencia (nm)	Resultado	Función
Ca II	849.8	Profundidad	$\sin(\sin(\sin(\text{square}(\text{square}(0.122663625454206 - 1.0346429 \cdot \cos(4.63919501030032e-5 \cdot \text{Teff} + 0.05084715 \cdot \text{logg} - 0.1016943 \cdot [\text{M}/\text{H}] - 0.05084715 \cdot \cos(\text{logg}) + 0.05084715 \cdot \cos(2.0537593 \cdot \text{Teff} - [\text{M}/\text{H}])))))))))$
Ca II	849.8	Desviación	$-1.2093169e-5 \cdot \text{Teff} \cdot \text{logg} \cdot [\text{M}/\text{H}] + 0.078072034 \cdot \text{logg} \cdot [\text{M}/\text{H}]$
Ca II	854.21	Profundidad	$-\sin(\sin(\sin(\sin(\sin(0.23192962 \cdot \text{logg} - 0.23192962 \cdot \sin([\text{M}/\text{H}])) + 5.269177 - 464.38162 \cdot \text{logg}/\text{Teff}))))))$
Ca II	854.21	Desviación	$\sin(\sin(\cos(\text{square}(\text{cube}(\cos(0.41173074 \cdot \text{logg}/\cos(34.77863 \cdot \text{Teff})))) + 0.2629335/\text{logg}))) - 0.25145125$

⁴ Tanto las fórmulas mostradas en esta tabla como las presentes en el Anexo A son factorizadas y simplificadas mediante la librería SymPy (Aaron Meurer, 2017) para facilitar ligeramente la visibilidad y el entendimiento, necesario debido a la complejidad de las producidas originalmente por PySR.

Si I	855.68	Profundidad	$\text{square}(\sin(\sin(0.16042033 * \sin(\text{square}(\sin(0.9102878 * \text{Teff} - 0.9102878 * \sin(\text{square}(0.10028628 * \text{logg} - 1.0593976)))))) * \exp([M/H]) + 0.18281925 * \text{square}(\sin(0.9102878 * \text{Teff} - 0.9102878 * \sin(\text{square}(0.10028628 * \text{logg} - 1.0593976)))))) + 0.18281925))$
Si I	855.68	Desviación	$\text{cube}(\text{square}(-0.000104252914233838 * \text{Teff}))$
Ca II	866.21	Profundidad	$\text{square}(\sin(\sin(\sin(2.1366665 / \text{Teff}))))$
Ca II	866.21	Desviación	$0.04349839 * [M/H] - 0.04349839 * \sin(0.0011901015 * \text{Teff} - 1.3542058) + 0.15644167 + 0.04349839 / \text{logg}$
H I	866.5	Profundidad	$\text{square}(-0.72500717213345 * [M/H] * \text{cube}(0.000159109156 * \text{Teff} - 0.51124054) - 1.4500143442669 * \text{cube}(0.000159109156 * \text{Teff} - 0.51124054) * \cos(\text{logg}))$
H I	866.5	Desviación	$-\sin(1 / (\text{Teff} - \text{cube}(\text{square}(-\text{logg} - 0.24412274))))$

Fuente: Elaboración propia

Estas funciones permiten inferir, incluso a pesar de su inherente complejidad, la dependencia de la absorción de los elementos químicos con los parámetros atmosféricos de la estrella, que era el objetivo del presente trabajo.

Alguna relación se puede derivar con facilidad, por ejemplo, la anchura de la línea de absorción del Ca II en 849.8 nm tiene una dependencia lineal incremental con la metalicidad y la gravedad, pero la temperatura es un modificador de esta dependencia invirtiéndola (la anchura decrecerá cuanto más temperatura). La profundidad del H I tiene asociada una función muy compleja, de la que sin embargo se puede inferir que es incrementada por la temperatura exponencialmente (ya que está elevada al cubo y luego al cuadrado) al mismo tiempo que la metalicidad, también con dependencia linealmente de la profundidad. Si se estudia la profundidad del Ca II en 866.21 se encuentra en una fórmula menos compleja el efecto contrario, a más temperatura tenga la estrella la profundidad de la línea de absorción disminuirá ligeramente.

Teniendo 352 funciones computadas no es posible con el tiempo disponible para el desarrollo de este trabajo estudiar y describir la dependencia de todas ellas con los parámetros atmosféricos elegidos, sin embargo, es algo que como se ha puesto de ejemplo se puede hacer. Finalmente se ha de demostrar que estas dependencias son fiables validando las funciones obtenidas.

4.2. Validación

Para la validación, cómo se ha explicado en la Sección 3.1, se van a usar espectros de alta resolución del espectrómetro RVS de Gaia, abarcando de 846 a 870 nm y cuyas especificaciones son la explicación detrás de los espectros de entrenamiento sintetizados para el entrenamiento del algoritmo de regresión simbólica llevado a cabo (así como de la elección de las lista de líneas atómicas y otras decisiones).

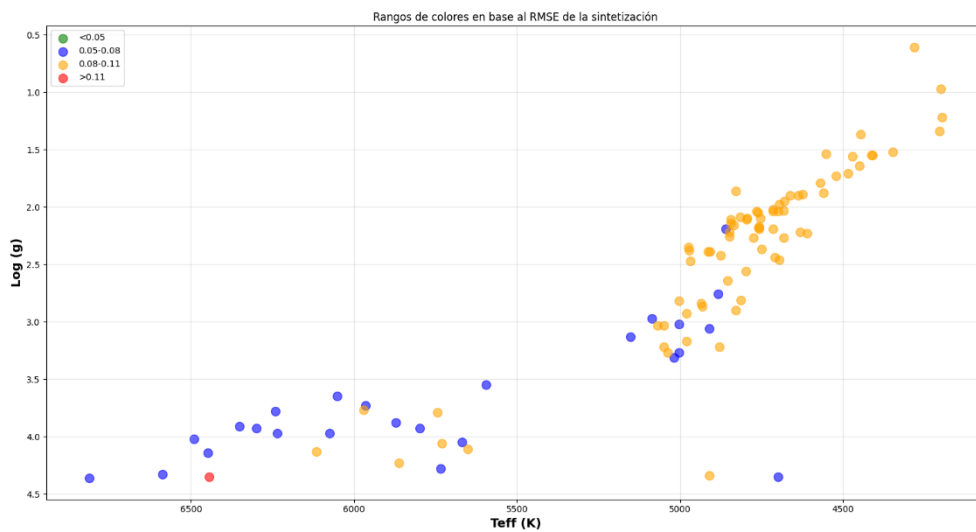
Se decide validar la síntesis con 100 espectros. Para seleccionarlos, primero, de la tabla *gaia_source* que contiene todas las estrellas capturadas por la misión Gaia presentes en su tercer catálogo, se seleccionan aquellas con un espectro RVS asociado y se ordenan de mayor a menor ratio señal a ruido (*rvs_spec_sig_to_noise*). Se seleccionan las 500 con una mejor relación, disminuyendo así el error obtenido al comparar que este causado por el ruido del espectro y no por una diferencia real en las líneas de absorción.

Seguidamente, de la tabla *astrophysical_parameters*, se obtienen los tres parámetros calculados por GSP-Spec que se han de usar para sintetizar los espectros y después comparar: temperatura, gravedad y metalicidad (*teff_gspspec*, *logg_gspspec* y *mh_gspspec*). A mayores, como segundo filtro de selección, se extrae de la misma tabla el logaritmo del nivel de ajuste de GSP-Spec (*logchisq_gspspec*), que muestra el logaritmo del error chi-cuadrado entre el espectro original y uno sintetizado por GSP-Spec. Se extraen de los 500 iniciales los 100 con un mejor ajuste, ya que es un medidor directo de la calidad de los parámetros producidos por el parametrizador.

Se descarga el flujo de los 100 espectros elegidos para su posterior comparación con lo sintetizados y se normalizan todos mediante convolución, usando un filtro de Gauss con desviación estándar de 5 (tal y como se hizo en la Sección 3.2.3.1 con los espectros de iSpec).

Para validar finalmente el resultado del trabajo, se sintetizan con el algoritmo desarrollado y las funciones obtenidas 100 espectros estelares correspondientes a los parámetros extraídos de los 100 seleccionados de Gaia, y se compara el resultado con estos originales, computando el error y validando así la eficacia del desarrollo. El resultado, es decir, los RMSE de la síntesis obtenidos al comparar con el espectro original, así como las gravedades y temperaturas asociadas a los espectros obtenidos de Gaia se puede observar en la Figura 30.

Figura 30. Errores de síntesis para los espectros de validación de Gaia



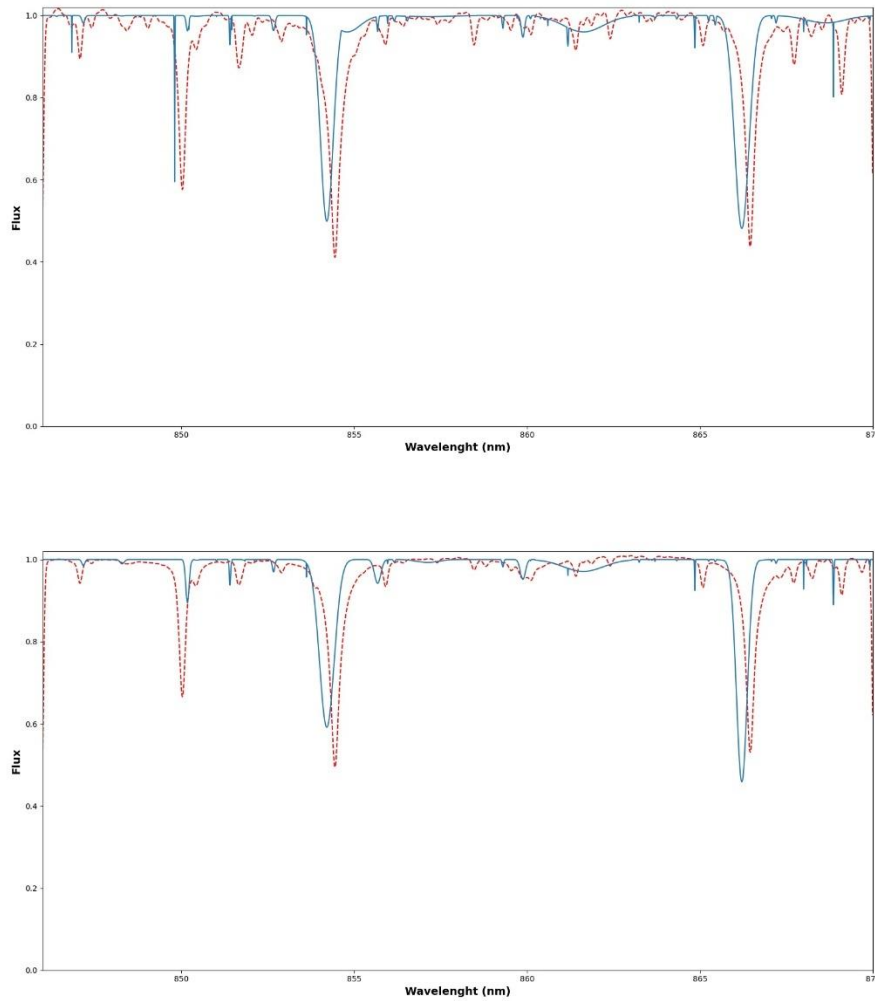
Fuente: Elaboración propia

Con un RMSE de media para los 100 espectros de 0.086, un mínimo obtenido de 0.067 y un máximo de 0.123, ninguna síntesis se ha acercado al umbral de 0.03 de RMSE aplicado para filtrar espectros de entrenamiento en el capítulo anterior, umbral a partir del cual se considera que el ajuste es adecuado.

Dos ejemplos se pueden apreciar en la Figura 31: la síntesis comparada con el espectro original para una estrella de 5050K, 3.03 log(g) y -0.12 [M/H] (arriba), con un RMSE obtenido de 0.084 y la síntesis para una estrella de 6448K, 4.14 log(g) y -0.21 [M/H] (abajo), con un RMSE obtenido de 0.074.

A parte de que el ajuste es visiblemente pobre, con muchas líneas verticales (mala desviación obtenida) y otras cuantas obviadas, se aprecia que hay una desviación de aproximadamente 0.23 nanómetros al rojo en todas las líneas de absorción presentes en los espectros RVS de Gaia con respecto a las frecuencias de las líneas de absorción presentes en la lista de VALD, usadas para computar los espectros sintéticos de entrenamiento de iSpec.

Figura 31. Sintetizaciones de validación para 2 espectros de Gaia

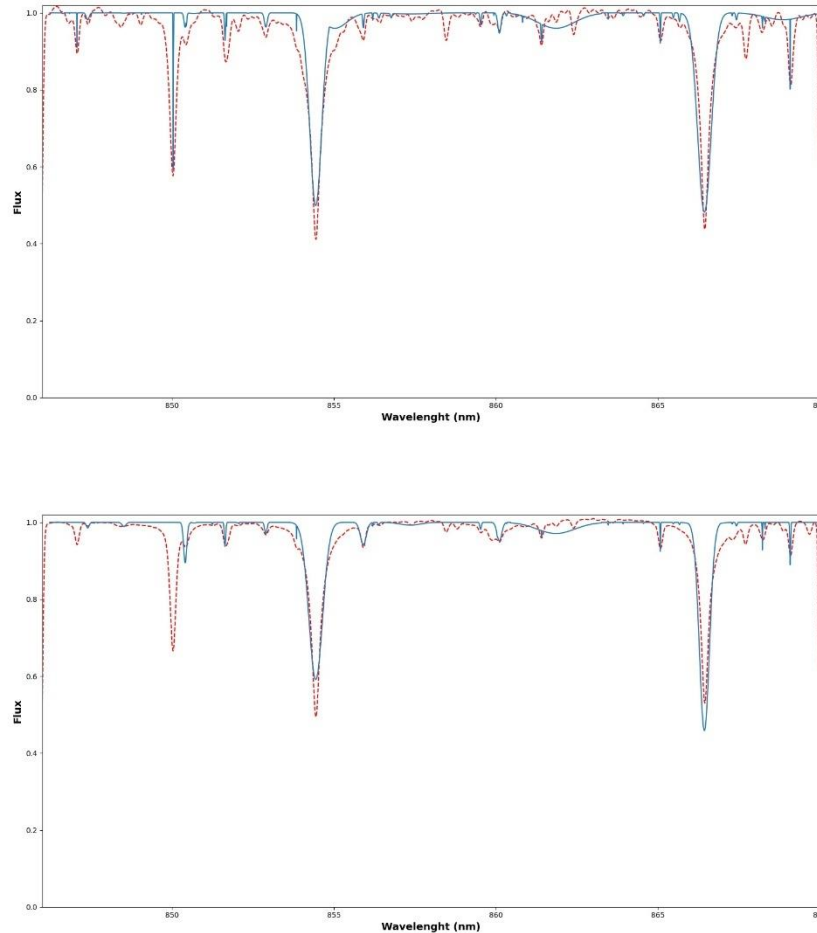


Fuente: Elaboración propia

Si se corrige esta desviación de 0.23 nm observada desplazando el espectro sintético hacia el rojo se consiguen dos ajustes ligeramente más certeros, visibles en la Figura 32, con errores

bastante disminuidos, ahora de 0.05 para el espectro superior y de 0.04 para el espectro inferior.

Figura 32. Sintetizaciones de validación con desviación corregida para 2 espectros de Gaia

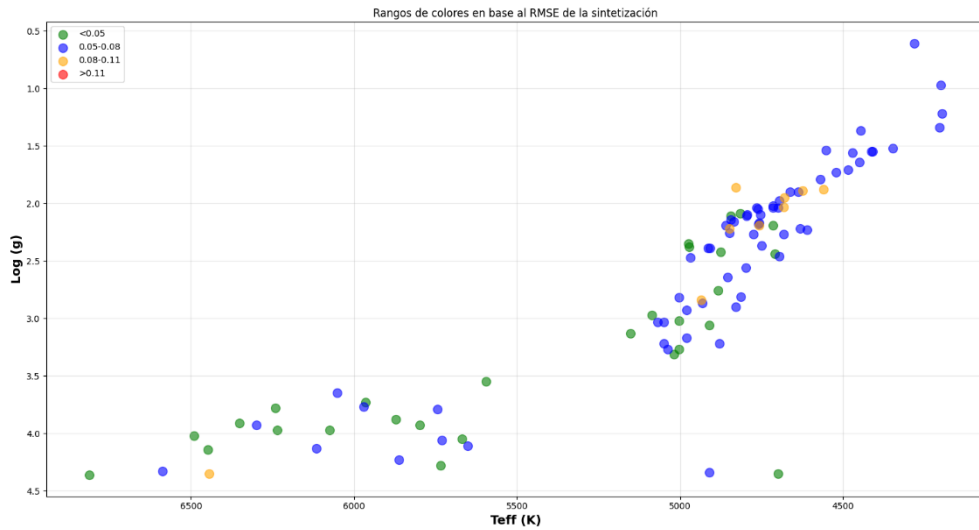


Fuente: Elaboración propia

Se vuelven por tanto a sintetizar los 100 espectros de validación con el desplazamiento de 0.23 nm para recalcular los errores, y una vez finalizado el proceso se muestran de nuevo en otro diagrama (Figura 33) pudiendo comparar el resultado fácilmente con el anterior.

Con un RMSE de media para los 100 espectros de 0.057, un mínimo de 0.025 y un máximo de 0.108, la calidad de la síntesis ha mejorado sustancialmente, siendo el problema de desviación al rojo un factor contribuyente de gran importancia al error previo.

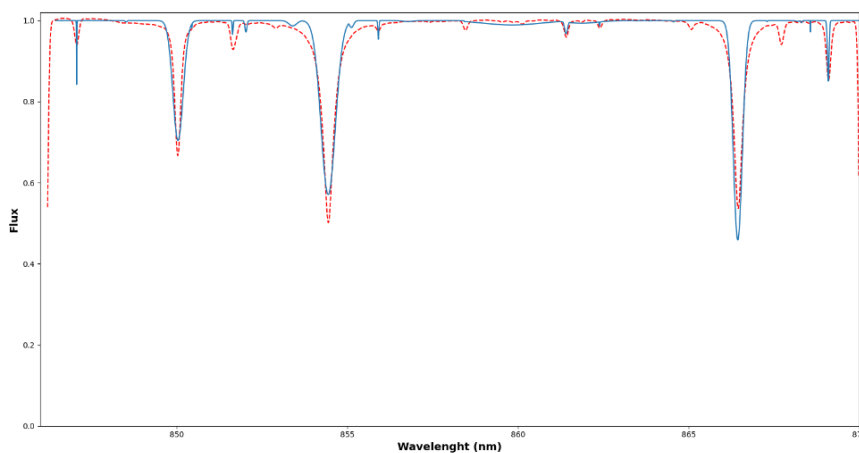
Figura 33. Errores de síntesis corregida para los espectros de validación de Gaia



Fuente: Elaboración propia

Incluso con los errores reducidos, alguno ahora incluso por debajo del umbral aceptable de 0.03 RMSE, el ajuste sigue sin ser del todo adecuado, como se puede apreciar en el mejor ajuste obtenido (Figura 34), con un RMSE de 0.025 para una estrella de 4700K, 4.35 log(g) y -1.64 [M/H].

Figura 34. Mejor síntesis de validación obtenida para un espectro de Gaia



Fuente: Elaboración propia

Casi todas las líneas de absorción aparecen, pero muchas tienen una anchura muy alejada de la realidad, en concreto por ser demasiado estrechas (provocado por una desviación típica cercana a 0), y otras, principalmente las que forman parte de una frecuencia de absorción más profunda, directamente son ignoradas y no están presentes. De todas formas, el ajuste superior en concreto se considera correcto para el ámbito del trabajo, representando un espectro bastante similar al original que permite descifrar al menos la mayoría de elementos químicos que contribuyen a la absorción estelar.

5. Conclusiones

En este capítulo final se compilan las conclusiones del trabajo y se mencionan posibles líneas de investigación futuras en base a las limitaciones y problemas encontrados.

5.1. Conclusiones

Tras un largo periodo de desarrollo, correcciones constantes e iteraciones se consigue crear un denso algoritmo automatizado para la extracción de parámetros gaussianos a partir de espectros de entrenamiento, un programa que ejecuta un algoritmo de regresión simbólica para que cree funciones enlazando estos parámetros gaussianos con los atmosféricos de la estrella cuyo espectro computó esas gaussianas, y finalmente un sintetizador que en base a unos parámetros atmosféricos, haciendo uso de esas funciones de la regresión, es capaz de producir un espectro sintético.

La calidad del resultado del trabajo se debe medir en base a la similitud del espectro sintético generado con las funciones con el espectro real cuyos parámetros estelares se han usado para generar el artificial. Y como se ha visto, tras validar el desarrollo con espectros reales del espectrómetro RVS de Gaia, los espectros generados permiten atisbar las líneas atómicas de absorción más predominantes del espectro real, no todas, pero las generadas con una profundidad y una anchura relativamente fieles a las originales. Sin embargo, el resultado, considerado al menos decente para el ámbito del trabajo limitado por recursos y tiempo, no es adecuado para ningún uso real en el campo de la astronomía, todavía no al menos. Los espectros sintetizados a partir del trabajo de la regresión siguen teniendo un error demasiado grande con respecto al original, con multitud de líneas de absorción perdidas y otras cuantas infieles al producto original, principalmente debido a una desviación de Gauss demasiado pequeña que tiende a desviarlas a ser líneas de absorción verticales y con una anchura prácticamente despreciable.

El desarrollo presente se puede interpretar como una primera aproximación a la generación de espectros sintéticos mediante el uso de regresión simbólica, un enfoque no adoptado hasta la fecha y que este trabajo amateur con recursos computacionales limitados, así como tiempo

y conocimientos, ha demostrado que con un desarrollo más exhaustivo y profundo puede producir resultados mucho más satisfactorios y usables en el campo de la astronomía. A su vez, con esta aproximación ya es posible mediante el algoritmo entrenado determinar con sencillez las tendencias de las líneas de absorción en función de los parámetros estelares, pudiéndose hacer por tanto una idea de los elementos químicos presentes en la estrella, un proceso mucho más costoso con otros métodos.

5.2. Dificultades encontradas y posibles líneas futuras de investigación

Las mejoras que se pueden aplicar con tiempo y recursos a cada iteración del desarrollo son numerosas.

El algoritmo desarrollado de ajuste de distribuciones de Gauss ha de conseguir un ajuste prácticamente perfecto si se quiere que luego el algoritmo de regresión simbólica produzca funciones que representen con fidelidad las líneas de absorción. Para ello, lo principal sería tener en cuenta las frecuencias de absorción contribuyentes a una más grande, que forman distribuciones de Gauss superpuestas y que en este desarrollo directamente se ignoraron, perjudicando naturalmente el ajuste en un grado bastante alto. A mayores, no todas las líneas de absorción se van a poder ajustar correctamente a una distribución gaussiana, por lo que se debería considerar usar alternativas como distribuciones de Voigt o de Cauchy para ciertas líneas, y luego ver cómo aplicar los valores obtenidos para estas en la regresión simbólica.

También, uno de los principales factores que perjudicaron el trabajo de la regresión fue el asignar valores de desviación de 0 a aquellas líneas de absorción a las que no se podían ajustar distribuciones gaussianas, ya que conllevó el desplazamiento del valor producido por las funciones hacia 0, disminuyendo la anchura en todas las líneas atómicas generadas. Se debería buscar una alternativa para que se tuviera en cuenta que para ciertos parámetros hay líneas no ajustables sin añadir valores a las desviaciones gaussianas que luego generen error en las funciones producidas.

Sobre la regresión simbólica en sí, el tiempo limitado disponible en este desarrollo, así como los recursos computacionales, no permitieron una investigación en profundidad de las distintas configuraciones y alternativas, así como límites, que pueden ofrecer los algoritmos.

PySR, la herramienta usada, permite multitud de personalizaciones en la ejecución, con altas posibilidades de que muchas de ellas beneficien sustancialmente la eficacia de la regresión para encontrar enlaces entre parámetros de entrada y salida. Que cada ejecución del algoritmo tardara en torno a las 8 horas no permitió explorar más valores para los parámetros ni crear entrenamientos variados para comparar resultados, pero es un proceso que debería de hacerse para maximizar la eficacia al computar las funciones. El tiempo de ejecución (número de iteraciones, ciclos, poblaciones...) debería de aumentarse también, ya que, en el único aumento realizado en este trabajo, que conllevó una subida de 3 a 8 horas, se apreció una sustancial mejora en las funciones obtenidas. Finalmente, también se debería probar con otros programas de regresión simbólica alternativos, ya que siempre es posible que para este problema en concreto otorguen mejores resultados.

Para terminar, el mecanismo de síntesis adoptado en este trabajo podría ser completamente distinto para obtener resultados más fieles. Por ejemplo, se podrían computar funciones adicionales a parte de las 352 usadas que permitieran obtener otros datos que ayuden a la síntesis, como que líneas no son ajustables para una serie de parámetros (es decir, que no tienen absorción), que líneas pueden necesitar de una distribución distinta la de Gauss (Cauchy o Voigt, por ejemplo), etc.

Una síntesis para un uso real debería tener en cuenta parámetros adicionales, como es el ruido de los espectros, y se debería probar con abundancias de metales concretas, e incluso con la edad de la estrella o tamaño, que podrían mejorar la relación entre parámetros de entrada y de salida.

Referencias bibliográficas

- Aaron Meurer, C. P. (2017). SymPy: symbolic computing in Python. *PeerJ Computer Science*.
- Abdurro'uf, K. A. (2022). The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar and APOGEE-2 Data. *The Astrophysical Journal Supplement Series*, 259.
- Alvarez, R., & Plez, B. (1998). Near-infrared narrow-band photometry of M-giant and Mira stars: models meet observations. *Astronomy and Astrophysics*, v.330, 1109-1119.
- Astropy Collaboration. (2022). The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. *The Astrophysical Journal*, 935.
- Blanco-Cuaresma, S. (2019). *Modern stellar spectroscopy caveats*.
- Blanco-Cuaresma, S., Soubiran, C., Heiter, U., & Jofré, P. (2014). Determining stellar atmospheric parameters and chemical abundances of FGK stars with iSpec. *A&A Volume 569*.
- C., M. P. (2009). Fitting models for Spectrum lines - Graduate Seminar I. *The Solar Spectrum 2008: Gaussianity of Absorption Lines*.
- Cranmer, M. (2023). *Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl*.
- Cropper, M., Katz, D., Sartoretti, P., & al., e. (2018). Gaia Data Release 2. *A&A Volume 616*.
- Dubcáková, R. (2011). Eureka: software review. *Genetic Programming and Evolvable Machines, Volume 12*, 173-178.
- Feynman, R., Leighton, A., & Sands, M. (1964). *The Feynman Lectures on Physics: The New Millennium Edition: Mainly Mechanics, Radiation, and Heat, v. 1*.
- Gaia Collaboration. (2023a). Gaia Data Release 3: Summary of the content and survey properties. *A&A Volume 674*.

- Gaia Collaboration. (2023b). Gaia Data Release 3: A golden sample of astrophysical parameters. *A&A Volume 674*.
- Gerber, J. M., Magg, E., Plez, B., & al., e. (2023). Non-LTE radiative transfer with Turbospectrum. *A&A v.669*.
- Gilmore, G., Randich, S., Asplund, M., & al., e. (2012). The Gaia-ESO Public Spectroscopic Survey. *The Messenger, vol. 147*, p. 25-31.
- Gray, R. O., & Corbally, C. J. (1994). The Calibration of MK Spectral Classes Using Spectral Synthesis. I. The Effective Temperature Calibration of Dwarf Stars. *Astronomical Journal v.107*.
- Grevesse, N., Asplund, M., & Sauval, A. J. (2007). The Solar Chemical Composition. *Space Science Reviews Volume 130*.
- Gustafsson, B., Edvardsson, B., Eriksson, K., & al, e. (2008). A grid of MARCS model atmospheres for late-type stars. *A&A Volume 486*, 951 - 970.
- Heiter, U., Jofré, P., Gustafsson, B., & al., e. (2015). Gaia FGK benchmark stars: Effective temperatures and surface gravities. *A&A v.582*.
- Heiter, U., Lind, K., Bergemann, M., & al., e. (2021). Atomic data for the Gaia-ESO Survey. *A&A v.645*.
- Herbert Goldstein, C. P. (2002). *Classical Mechanics*.
- Jackson, J. D. (1999). *Classical electrodynamics*.
- Jaeyoung Chung, J. O. (2024). Depth-Regularized Optimization for 3D Gaussian Splatting in Few-Shot Images.
- Kordopatis, G., Schultheis, M., McMillan, P. J., & al., e. (2023). Stellar ages, masses, extinctions, and orbital parameters based on spectroscopic parameters of Gaia DR3. *A&A Volume 669*.
- Kupka, F., & Dubernet, M. (2011). VAMDC as a Resource for Atomic and Molecular Data and the New Release of VALD. *Open Astronomy*.

- Kurucz, R. L. (1992). The Stellar Populations of Galaxies. *IAU Symposium*.
- Kurucz, R. L. (2005). ATLAS12, SYNTHE, ATLAS9, WIDTH9, etcetera. *Memorie della Società Astronomica Italiana Supplement*.
- Laverny, P. d., Recio-Blanco, A., Worley, C. C., & Plez, B. (2012). The AMBRE project: A new synthetic grid of high-resolution FGKM stellar spectra. *A&A* 544.
- Li, Z., Zhao, G., Chen, Y., Liang, X., & Zhao, J. (2022). The stellar parameters and elemental abundances from low-resolution spectra - I. 1.2 million giants from LAMOST DR8. *Monthly Notices of the Royal Astronomical Society*, 517.
- Malagnini, M. L., & Morossi, C. (1990). Accurate absolute luminosities, effective temperatures, radii, masses and surface gravities for a selected sample of field stars. *Astronomy and Astrophysics*, 85.
- Michael Affenzeller, B. B. (2024). *Symbolic Regression*. Chapman and Hall/CRC.
- Recio-Blanco, A., Bijaoui, A., & Laverny, P. d. (2006). Automated derivation of stellar atmospheric parameters and chemical abundances: the MATISSE algorithm. *Monthly Notices of the Royal Astronomical Society, Volume 370*, 141–150.
- Recio-Blanco, A., Laverny, P. d., Palicio, P. A., & Kordopatis, G. (2023). Gaia Data Release 3: Analysis of RVS spectra using the General Stellar Parametrising from spectroscopy. *A&A* 674.
- Snedden, C., Bean, J., Ivans, I., Lucatello, S., & Sobek, J. (2012). MOOG: LTE line analysis and spectrum synthesis. *Astrophysics Source Code Library*.
- Udrescu, S.-M., & Tegmark, M. (2020). AI Feynman: A physics-inspired method for symbolic regression. *Science advances, Vol.6*.
- Van Rossum, G. a. (2009). *Python 3 Reference Manual*. CreateSpace.
- Virtanen, P. G. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*, 17.

Anexo A. Lista completa de funciones generadas

Elem.	Freq (nm)	Resultado	Función
CN I	846.35	Profundidad	0.0000244
CN I	846.35	Desviación	$\sin(3.7435321 * \text{square}(\text{cube}(-\sin(9.7221460 / (\text{Teff} * [M/H] - \text{Teff} * \sin(0.2746740 * \text{Teff} / \text{logg})))) / \text{logg})) / \text{logg}$
Fe I	846.4	Profundidad	$1.6607376 * \text{square}(-2 * \text{logg} * \exp([M/H]) + 5.8610716 * \exp([M/H]) * \sin(0.5257731 * \text{Teff} - 0.5257731 * \cos(\text{logg} + 2.4176967)) + 9.9084950 * \exp([M/H])) / \text{Teff}$
Fe I	846.4	Desviación	$0.0238246 * \text{cube}(\cos([M/H] + \cos(\exp(-0.8486673 / \text{logg}) * \exp(-[M/H]) * \exp(-\sin(\text{logg} - [M/H] + \text{square}(0.3752419)) / \text{logg}) / \cos(\sin(1.1637069 * \text{Teff}))))$
Fe I	846.51	Profundidad	$\exp(\text{cube}(1.2334405 - \text{Teff}))$
Fe I	846.51	Desviación	$\text{square}(\text{square}(\text{cube}(1 / \text{Teff})))$
Fe I	846.52	Profundidad	$0.3209861 * \exp(-\text{Teff})$
Fe I	846.52	Desviación	0
CN I	846.56	Profundidad	$\text{cube}(0.4254033 * \exp([M/H]) * \cos(0.0019638 * \text{Teff}) * \cos(1.0245011 * \text{Teff} + 0.5905095 * \text{logg} - 0.3684824)) * \cos(\text{logg} * [M/H])$
CN I	846.56	Desviación	$\text{square}(\cos(\text{Teff})) / (\text{Teff} * [M/H] * \cos(\text{logg} - 1 / \cos(\text{Teff})))$
Fe I	846.61	Profundidad	$1 * (17.2135560 * [M/H] * \text{cube}(\sin(\text{logg}) + \cos(0.5253212 * \text{Teff}) + 0.7877619) * \text{cube}(\text{cube}(\sin(\text{logg} * [M/H] + 0.7537196 * \text{logg} + \cos(\text{Teff})))) + 13.5601853 * \text{cube}(\sin(\text{logg}) + \cos(0.5253212 * \text{Teff}) + 0.7877619) * \text{cube}(\text{cube}(\sin(\text{logg} * [M/H] + 0.7537196 * \text{logg} + \cos(\text{Teff})))))) / \text{Teff}$
Fe I	846.61	Desviación	$-\text{logg} * * 2 * \sin(2.1513867 / \text{cube}(\text{cube}(\sin(18.5588100 - \text{square}(\text{Teff}) / [M/H]) - 1.9696400))) / (\text{Teff} * [M/H])$
Fe I	846.65	Profundidad	0
Fe I	846.65	Desviación	0
Ti I	846.71	Profundidad	$\text{cube}(\text{cube}(1 / \text{Teff}))$

Ti I	846.71	Desviación	$\text{cube}(\text{cube}(1/\text{Teff}))$
Ti II	846.77	Profundidad	$\text{square}(\sin(\sin(\sin(\text{square}(\sin(\sin(\sin(2.0155617*\sin(\text{square}(2.1006897*\text{square}(\sin(\sin(10.3924542*\text{cube}(\cos([M/H])*\cos(0.0203655*\text{Teff} + 0.0203655*\sin(4.3820099*\exp(\log)))\text{))*exp(-\log))))))))))))))$
Ti II	846.77	Desviación	$1*(-0.1007204*\log*[M/H]*\exp(1.0672538*\text{cube}(\text{cube}([M/H] - 0.2766949)))*\exp(1.0672538*\sin(\text{Teff} + \log))*\exp(1.0672538*\cos(2.8667440*\text{Teff})) + 0.0509114*\log*[M/H] - 0.1005273*\log*\exp(1.0672538*\text{cube}(\text{cube}([M/H] - 0.2766949)))*\exp(1.0672538*\sin(\text{Teff} + \log))*\exp(1.0672538*\cos(2.8667440*\text{Teff})) + 0.0508138*\log)*\exp(-\log**2)$
Fe I	846.84	Profundidad	$-0.0000100*\text{Teff}*\log + 0.0000100*\text{Teff}*\sin(\text{square}(\text{Teff})) + 0.0770361*\log - 0.0770361*\sin(\text{square}(\text{Teff}))$
Fe I	846.84	Desviación	$\text{square}(0.3525184/(\text{Teff}*\text{cube}(\text{square}(\cos(\cos(0.3048896*\text{Teff})))))*\text{square}(\text{square}(0.0000833*\text{Teff}))*\cos(\cos(\text{square}(-3.6677356/[M/H])))$
Ti I	846.85	Profundidad	$0.0142834*\exp(\sin(\text{square}(\text{Teff}))*\exp(\cos(\text{Teff} + 0.9966036))*\exp(\cos(\log - 1.2295916))$
Ti I	846.85	Desviación	$\text{square}(\sin([M/H]*\text{cube}(0.7687448*\sin(\log - 0.5950723*\sin(\text{square}(\text{Teff}))))*\exp([M/H])*exp(\sin(-0.5273182*\text{Teff}))/\log - 0.2533983*\text{cube}(0.7687448*\sin(\log - 0.5950723*\sin(\text{square}(\text{Teff}))))*\exp([M/H])*exp(\sin(-0.5273182*\text{Teff}))/\log))$
Fe I	846.86	Profundidad	0
Fe I	846.86	Desviación	$\text{cube}(1/\text{cube}(\text{Teff} + 0.4694899))$
Mg I	846.88	Profundidad	$\text{cube}(\text{square}(\text{square}(-0.0656387)/\text{Teff}))$
Mg I	846.88	Desviación	$\text{cube}(-0.2841506/\text{cube}(\text{Teff}))$
CN I	846.99	Profundidad	$\text{square}(\exp(-2.2513271*\text{Teff}))$
CN I	846.99	Desviación	$\text{square}(1/\text{cube}(\text{square}(\text{Teff})))$
Fe I	847.04	Profundidad	$5.7605553*\exp(-\text{Teff})$
Fe I	847.04	Desviación	0
Fe I	847.17	Profundidad	$1*\sin(-0.0106224) + 0.0656471*\exp([M/H])*exp(\cos(-\cos(\sin([M/H])) + 1/\cos(\cos(1.1528023*\text{Teff} - 0.1528023*\log + 0.1528023*\sin(\log))) + 1/(\log + 0.0500062)))/\log$

Fe I	847.17	Desviación	$\text{square}(1.3837757*\text{square}(0.4843894 - 0.0883464*\exp(-\sin(\text{logg})*\cos(\text{Teff}))*\exp(\sin(\text{logg})*\cos(\text{Teff}*\text{logg}))*\exp(-\sin(\text{cube}([\text{M}/\text{H}])*\text{square}(\text{logg}))*\cos(\text{Teff}))*\exp(\sin(\text{cube}([\text{M}/\text{H}])*\text{square}(\text{logg}))*\cos(\text{Teff}*\text{logg}))*\exp(0.8622622*\sin(\text{logg}))*\exp(0.8622622*\sin(\text{cube}([\text{M}/\text{H}])*\text{square}(\text{logg}))))))$
CN I	847.24	Profundidad	$1/\text{square}(\text{square}(\text{square}(\text{Teff})))$
CN I	847.24	Desviación	$\text{cube}(-0.0641905/\text{square}(\text{Teff}))$
Mg I	847.37	Profundidad	$2.0538098*\text{square}(\text{cube}(-\cos(\text{logg} - [\text{M}/\text{H}])/(\cos(\exp([\text{M}/\text{H}])) + \cos(\text{cube}(\text{cube}(\text{Teff} - 0.2339992)))))/(\cos(\cos(\exp([\text{M}/\text{H}]))))/\text{Teff}$
Mg I	847.37	Desviación	$0.0009983*\text{cube}(\text{cube}(-0.1986267*\exp([\text{M}/\text{H}])*\sin(1/\text{cube}(0.9469467*[\text{M}/\text{H}]*\sin(\text{cube}(\text{Teff})))))/\sin(\text{logg}))$
Fe I	847.64	Profundidad	0.0000309
Fe I	847.64	Desviación	$\text{square}(\sin(\text{square}(\text{cube}(-0.0559418*\text{E}*\exp(3.4868574*[\text{M}/\text{H}])*\exp(\cos(0.9406059*\text{Teff}))*\cos(\sin(\text{logg}))))/\cos(\text{square}(-0.1111217/\text{cube}(\exp(0.2833560*[\text{M}/\text{H}]) - \cos(0.8703783*\text{Teff}) - 0.9406059))*\exp(\sin(\text{logg}))))))$
Fe I	847.71	Profundidad	$7.8028263*\text{square}(\exp(\cos(0.3116546*\text{logg} - 2*[\text{M}/\text{H}]))*\cos(0.2942009*\text{Teff} - 0.2942009*\text{logg}))*\text{square}(\exp(\cos(0.2942186*\text{Teff}))/\text{Teff}$
Fe I	847.71	Desviación	$1*(-0.0211211*[\text{M}/\text{H}] - 0.0211809)/\text{square}(\text{logg})$
CN I	848.04	Profundidad	$\text{square}(0.0332054*\text{logg}*\sin(\text{square}(\text{Teff}))/([\text{M}/\text{H}]*\exp(\text{logg}) - 0.0391713*\exp(\text{logg})))$
CN I	848.04	Desviación	$\sin(4.4132873*[\text{M}/\text{H}]*\text{cube}([\text{M}/\text{H}]/\sin([\text{M}/\text{H}]) + \sin(\text{logg})/\sin([\text{M}/\text{H}]))/(\text{logg}*\text{square}(2.5171941*\text{Teff}*\sin(0.1281195*\text{Teff}*\text{logg}) - 0.5424406*\text{Teff}*\sin(0.1281195*\text{Teff}*\text{logg}) + 0.6256824*[\text{M}/\text{H}]*\sin(0.1281195*\text{Teff}*\text{logg}) - 0.1348309*\sin(0.1281195*\text{Teff}*\text{logg})))$
Fe I	848.06	Profundidad	$\text{cube}(\text{square}(-\text{square}(\text{logg} + [\text{M}/\text{H}])*\sin(\text{square}(\text{Teff}))/(\text{Teff}*\sin(\text{logg}) - 0.4242120*\text{Teff})))$
Fe I	848.06	Desviación	$\text{logg}*\text{square}(\exp([\text{M}/\text{H}]*\text{cube}(\cos(2.1832404*\text{logg}) + 1.2033801))*\exp(\cos(\text{logg} - 1.3178773))/(\text{Teff}*\sin(\text{cube}(\cos(\exp(\sin(\text{Teff})))))))$
CN I	848.16	Profundidad	$\text{cube}(\text{cube}(1/\text{square}(\text{Teff})))$
CN I	848.16	Desviación	0

Fe I	848.2	Profundidad	$\text{square}(0.0930119*[M/H]*\cos(1.6657008*\text{Teff}) + 0.0465060*\sin(\text{logg} - [M/H]))*\cos(1.6657008*\text{Teff}) - 0.0465060*\sin(\text{square}(\text{Teff}))*\cos(1.6657008*\text{Teff}) + 0.1818797*\cos(1.6657008*\text{Teff})$
Fe I	848.2	Desviación	$-0.0641742*\cos(\text{logg}*\text{cube}(\cos(0.8269670*\text{Teff} + [M/H] + 0.3521183)))$
Fe I	848.29	Profundidad	$-\sin(\text{cube}(-\text{square}(\exp(\sin(\text{cube}(-\text{Teff} - 2.1594906)))))) + \cos(\text{logg} - \sin(\text{Teff}))*\exp([M/H])/(Teff*\text{logg} + 0.5437859*\text{Teff})$
Fe I	848.29	Desviación	$-0.0018945/(0.0002817*\text{Teff} - 1)$
CN I	848.83	Profundidad	$\sin(0.0054717/(\text{logg}*\text{square}(\text{cube}(0.8505489/\cos(0.0011920*\text{Teff} - [M/H])))))$
CN I	848.83	Desviación	$\text{square}(4.6345751/(\text{cube}(\text{logg}) - 118023.5261016*\text{cube}([M/H])/Teff + 118023.5261016*\exp(\cos(\text{Teff} + \text{logg} + [M/H] - 0.8527659))/Teff))$
Si I	849.21	Profundidad	$\text{cube}(1/\text{cube}(\text{Teff}))$
Si I	849.21	Desviación	$\exp(-1.0174187*\text{Teff})$
Fe I	849.38	Profundidad	0
Fe I	849.38	Desviación	0
Ti I	849.44	Profundidad	$0.1316736*\sin(\text{square}(\text{square}(\sin(0.3989008*\exp(\sin([M/H] - \cos(1.1734405*\exp(-\cos(9.7614405*\text{Teff})))))))*\sin(\text{logg}) + 0.4854004*\exp(\sin([M/H] - \cos(1.1734405*\exp(-\cos(9.7614405*\text{Teff})))))))))$
Ti I	849.44	Desviación	$\text{cube}(3.1544774/(0.7866366*\text{Teff}*[M/H]*\sin(\text{logg}) - 248.6815026*\sin(\text{logg}) + 0.7866366*\sin(\text{logg})/\cos(\text{logg} - 0.8278186)))$
Fe I	849.57	Profundidad	$\text{cube}(\text{cube}(1/\text{Teff}))$
Fe I	849.57	Desviación	0
Ti I	849.6	Profundidad	0
Ti I	849.6	Desviación	0
Fe I	849.65	Profundidad	$\exp(-0.9325192*\text{Teff}/\text{logg})$
Fe I	849.65	Desviación	$\text{cube}(\text{cube}(1/\text{Teff}))$
Fe I	849.7	Profundidad	$1/\text{cube}(\text{cube}(\text{Teff}))$

Fe I	849.7	Desviación	$\text{cube}(0.0373874/\text{cube}(\text{Teff}))$
CN I	849.79	Profundidad	$\text{exp}(-0.2895413*\text{Teff})$
CN I	849.79	Desviación	$\text{exp}(\text{cube}(1/(-0.2385325)))$
Ca II	849.8	Profundidad	$\text{sin}(\text{sin}(\text{sin}(\text{square}(\text{square}(0.1226636 - 1.0346429*\text{cos}(0.0000464*\text{Teff} + 0.0508472*\text{logg} - 0.1016943*[M/H] - 0.0508472*\text{cos}(\text{logg}) + 0.0508472*\text{cos}(2.0537593*\text{Teff} - [M/H])))$
Ca II	849.8	Desviación	$-0.0000121*\text{Teff}*\text{logg}*[M/H] + 0.0780720*\text{logg}*[M/H]$
Ca II	849.81	Profundidad	$\text{cube}(0.6417904/\text{cube}(\text{Teff}))$
Ca II	849.81	Desviación	0
Fe I	850.03	Profundidad	$\text{exp}(-\text{Teff})*\text{exp}([M/H])$
Fe I	850.03	Desviación	$\text{square}(1/\text{cube}(\text{Teff}))/\text{Teff}$
Si I	850.15	Profundidad	$\text{square}(\text{square}(0.4626432/\text{square}(\text{Teff})))$
Si I	850.15	Desviación	$\text{cube}(\text{sin}(\text{square}(\text{cube}(-0.0491759))))$
Ni I	850.18	Profundidad	$\text{square}(-0.3322654*\text{cube}(\text{cos}(\text{square}(\text{cos}(0.2179293*\text{square}(\text{logg}*\text{cos}(0.1717625*[M/H]*\text{square}(\text{logg}))/\text{logg} + \text{cos}(0.0819025*\text{Teff} + [M/H] - \text{sin}(\text{logg}) - 0.1534477))/\text{logg}))))$
Ni I	850.18	Desviación	$0.0117677*\text{logg}$
Si I	850.22	Profundidad	$0.1069223/(1*\text{logg} - 1*[M/H])$
Si I	850.22	Desviación	$\text{square}(\text{cube}(0.0000938*\text{Teff}))$
Ni I	850.23	Profundidad	$0.0395852 - 0.0085657*\text{logg}$
Ni I	850.23	Desviación	$[M/H]*\text{sin}(0.0179998*\text{sin}(2.7873032*\text{logg}))/\text{sin}(4.1885304*\text{Teff}) - 0.0119887/\text{sin}(4.1885304*\text{Teff})$
CN I	850.45	Profundidad	$1*\text{square}(6.9815363/(\text{Teff}*\text{logg}*\text{sin}(2*\text{Teff} + [M/H] - 0.1471017))) + 0.0018725$
CN I	850.45	Desviación	$-1*\text{logg}*\text{exp}(\text{cube}(1.6633225*[M/H] - 1.6633225*\text{cos}(\text{Teff} - \text{exp}(1/[M/H]))*\text{exp}(\text{cos}([M/H] - 0.7973747/(\text{Teff}*[M/H]))))) + 0.7426583*\text{logg} + (-1*\text{logg}*\text{exp}(\text{cube}(1.6633225*[M/H] -$

			$1.6633225 \cdot \cos(\text{Teff} - \exp(1/[M/H]) \cdot \exp(\cos([M/H] - 0.7973747 / (\text{Teff} \cdot [M/H]))) + 0.7426583 \cdot \logg) / (1 \cdot [M/H] + 1 \cdot \cos(0.7962679 \cdot \logg))$
Fe I	850.46	Profundidad	$\text{square}(\text{cube}(\text{cube}(3372.0928000 / (\text{Teff} + \exp(\logg) \cdot \exp(-[M/H]) \cdot \exp(\text{square}(\exp(30.2514820 \cdot \text{Teff} / \text{square}(-0.0978101 \cdot \text{Teff})))))) \cdot \exp(\logg))$
Fe I	850.46	Desviación	$5.1790466 \cdot \text{Teff} \cdot \text{square}(\logg - 2.7940564) / (0.4618200 \cdot \text{Teff} \cdot [M/H] \cdot \text{square}(\logg - 2.7940564) - 0.0787457 \cdot \text{Teff} \cdot \text{square}(\logg - 2.7940564) + 0.4618200 \cdot \logg \cdot \text{square}(\logg - 2.7940564) + 0.4618200 \cdot \logg + 1 \cdot \text{square}(\logg - 2.7940564))$
Si I	851.02	Profundidad	$-\sin(0.1502693 / (\logg \cdot \text{cube}([M/H] - \sin(2.8889826 \cdot \text{Teff}) - 2.6367180)))$
Si I	851.02	Desviación	$-0.0172885 \cdot \exp(\sin(0.5420605 \cdot \text{Teff})) / \logg$
Fe I	851.17	Profundidad	$\sin(\sin(1 / (\text{Teff} / \text{cube}(\exp(-\sin(\sin(\logg) + 0.5697396)) \cdot \exp(\sin(0.6542735 \cdot \logg \cdot [M/H] - 0.0793153 \cdot [M/H] + \sin(0.7772820 \cdot \text{Teff})))) + \exp(\logg)))$
Fe I	851.17	Desviación	$\sin(\sin(291.0650975 \cdot \sin([M/H]) \cdot \sin(\text{cube}(-\text{Teff} / \text{square}(-1.1526774) + 0.7401777 / \text{square}(-1.1526774))) - [M/H]) / (\text{Teff} \cdot \logg \cdot [M/H]))$
Fe I	851.21	Profundidad	$\exp(-0.0992378 \cdot \text{Teff})$
Fe I	851.21	Desviación	$\text{cube}(\text{cube}(1 / \text{Teff}))$
Fe I	851.41	Profundidad	$\text{square}(\sin(1258.7847440 \cdot \logg \cdot [M/H]) / (\text{Teff} \cdot \logg / (0.0467866 \cdot \logg^2 - 0.5407972) - 0.7039320 \cdot \text{Teff}) + 1208.3997447 \cdot [M/H] / (\text{Teff} \cdot \logg / (0.0467866 \cdot \logg^2 - 0.5407972) - 0.7039320 \cdot \text{Teff}) + 336.9069897 \cdot \logg / \text{Teff} + 323.4217147 / \text{Teff}))$
Fe I	851.41	Desviación	$\sin(1 / (0.0057036 \cdot \text{Teff} \cdot \cos(-1.0071413 \cdot \text{Teff} + [M/H] + \cos(\logg)) - 0.0117301 \cdot \text{Teff} + 140.7018100))$
CN I	851.46	Profundidad	$5.2045877 \cdot \text{square}(\text{square}([M/H] + \sin(\logg - [M/H]) + \sin(0.0400764 \cdot \text{Teff} + [M/H] + 1.3557659) + 1.9926128)) / \text{Teff}$
CN I	851.46	Desviación	$-0.0061463 \cdot [M/H] / (1 \cdot \logg \cdot [M/H] - 0.1658981)$
Fe I	851.51	Profundidad	0
Fe I	851.51	Desviación	$\exp(-2.4912977 \cdot \text{Teff})$
Fe I	851.71	Profundidad	0
Fe I	851.71	Desviación	$\text{square}(1 / \text{cube}(-1.6188677 \cdot \text{Teff}))$

Ti I	851.8	Profundidad	$\sin(1.4673775 \cdot \text{cubo}(\exp([M/H])) \cdot \sin(0.7945738 \cdot \text{Teff}) / (\text{cubo}(\text{logg}) + 27.9704820))$
Ti I	851.8	Desviación	$-45.8147130 / (1 \cdot \text{Teff} - 1 \cdot \text{square}(-1.2492699 \cdot \text{logg} / \sin(\text{Teff})))$
Ti I	851.84	Profundidad	$0.5287785 \cdot \text{cubo}(0.0383926 \cdot [M/H] \cdot \text{square}(-1.1854665 \cdot \cos(0.7754909 \cdot \text{Teff})) / \cos(\cos(0.2064304 \cdot \text{logg}) - 0.3824252) + 0.4262998 \cdot \text{square}(-1.1854665 \cdot \cos(0.7754909 \cdot \text{Teff})) / \cos(\cos(0.2064304 \cdot \text{logg}) - 0.3824252))$
Ti I	851.84	Desviación	$1 \cdot \text{logg} \cdot \text{square}(\text{cubo}(\text{square}(\cos(\text{square}(\text{cubo}(\exp(\text{square}(\text{square}(\text{square}(\text{cubo}(\text{square}(\text{cubo}(\cos(\text{Teff} + [M/H])))))))))))) - \cos(\exp(\text{square}(\text{square}(\sin(\cos(\cos(\text{square}(1.2621312 - \text{square}(\exp(\sin(\text{cubo}(\text{cubo}(\cos(\cos(\text{cubo}(\text{Teff} - \text{square}(\text{logg})))))))))) - 0.7119476)))))) - 0.4014105) - 0.1475632 \cdot \text{logg} - 1 \cdot [M/H] \cdot \text{square}(\text{cubo}(\text{square}(\cos(\text{square}(\text{cubo}(\exp(\text{square}(\text{square}(\text{square}(\text{cubo}(\text{square}(\text{cubo}(\cos(\text{Teff} + [M/H])))))))))))) - \cos(\exp(\text{square}(\text{square}(\sin(\cos(\cos(\text{square}(1.2621312 - \text{square}(\exp(\sin(\text{cubo}(\text{cubo}(\cos(\cos(\text{cubo}(\text{Teff} - \text{square}(\text{logg})))))))))) - 0.7119476)))))) - 0.4014105) + 0.1475632 \cdot [M/H]$
CN I	852.12	Profundidad	$0.0396813 \cdot \exp(-\text{square}(-0.3456333 \cdot \text{logg} + 1.3456333 \cdot [M/H] + 0.6912666 \cdot \sin(0.5671950 \cdot \text{Teff}) - 0.1381819))$
CN I	852.12	Desviación	$0.0028883 \cdot \cos(\text{logg}) / [M/H]$
Zr II	852.51	Profundidad	$\sin(\text{cubo}(\text{logg} / \text{cubo}(\text{Teff})))$
Zr II	852.51	Desviación	0
Fe I	852.67	Profundidad	$-0.2004801 \cdot \sin(\cos([M/H]) / (\text{logg} + 1.1776744)) / (0.0525054 \cdot \text{logg} \cdot [M/H] \cdot \cos(0.7107948 \cdot \text{Teff} - 1.1772573) - 0.0525054 \cdot \text{logg} \cdot \cos(0.7107948 \cdot \text{Teff} - 1.1772573) + 1 \cdot [M/H] - 1)$
Fe I	852.67	Desviación	$-0.1421611 \cdot [M/H] \cdot \sin(1.2379969 \cdot \text{Teff} - 1.2379969 \cdot [M/H]) + 0.1421611 \cdot \cos([M/H]) - 0.1006525$
Fe I	852.79	Profundidad	0
Fe I	852.79	Desviación	0
CN I	853.17	Profundidad	$-0.0445858 + 1 \cdot \text{square}(\exp([M/H]) + 6.0825167) / (0.1526662 \cdot \text{Teff} + 1 \cdot \text{square}(\exp([M/H]) + 6.0825167) \cdot \sin(\text{square}(\text{Teff})))$
CN I	853.17	Desviación	$1 / (\text{Teff} \cdot \cos(0.1894986 \cdot \text{Teff}))$
Si I	853.62	Profundidad	$\sin(\sin(23.9623583 \cdot \text{logg} \cdot \text{cubo}(0.1068756 / (\cos(\cos(\text{logg}))) + \exp(-\cos(8.9210990 \cdot \text{Teff})))) \cdot \exp([M/H]))$

Si I	853.62	Desviación	$1.6114655/(1*\text{Teff}*\cos(\text{Teff} - 3.1311224) - 1*\exp(\log\text{g}))$
Fe I	853.8	Profundidad	$\text{cube}(\text{cube}(\text{square}(\text{square}(\sin(0.0340486))))))$
Fe I	853.8	Desviación	$1/\text{cube}(\text{cube}(\text{Teff}))$
CN I	853.82	Profundidad	$\text{cube}(\text{square}(\text{square}(0.0000005*\sin(\text{Teff}))*\sin(\text{Teff})))$
CN I	853.82	Desviación	0
Ce II	853.91	Profundidad	$\text{cube}(\text{cube}(\cos(\log\text{g})/\text{Teff}))$
Ce II	853.91	Desviación	$\text{square}(\text{cube}(-0.2578995/\text{Teff}))$
Ti I	853.93	Profundidad	$\text{square}([[\text{M}/\text{H}]/\text{cube}(\text{Teff} + \text{cube}(\text{Teff}))])$
Ti I	853.93	Desviación	$\text{square}(0.5006905/([\text{M}/\text{H}] - \text{square}(\text{square}(\text{Teff}))))$
Ca II	854.21	Profundidad	$-\sin(\sin(\sin(\sin(\sin(0.2319296*\log\text{g} - 0.2319296*\sin([\text{M}/\text{H}])) + 5.2691770 - 464.3816200*\log\text{g}/\text{Teff}))))))$
Ca II	854.21	Desviación	$1*\sin(\sin(\cos(\text{square}(\text{cube}(\cos(0.4117307*\log\text{g}/\cos(34.7786300*\text{Teff})))))) + 0.2629335/\log\text{g})) - 0.2514513$
Ca II	854.22	Profundidad	0
Ca II	854.22	Desviación	$\text{cube}(\text{cube}(1/\text{Teff}))$
Ca II	854.23	Profundidad	$3.9369240*\exp(-\text{Teff})$
Ca II	854.23	Desviación	0
Cr I	854.37	Profundidad	0
Cr I	854.37	Desviación	0
Ni I	854.77	Profundidad	$\text{cube}(\text{cube}(1/\text{Teff}))$
Ni I	854.77	Desviación	$\text{cube}(\text{square}(1/\text{cube}(\text{Teff})))$
Fe I	854.8	Profundidad	$0.0583193 - 0.0000088*\text{Teff}$
Fe I	854.8	Desviación	$0.0240884*\log\text{g}*\text{[M/H]} - 154.8535046*\log\text{g}*\text{[M/H]}/\text{Teff}$

Ti I	854.81	Profundidad	$\sin(\text{square}(-0.0001805 * \text{Teff} * \sin(\cos([M/H])) + 1.1507545 * \sin(\cos([M/H])))$
Ti I	854.81	Desviación	$-1 * \exp(0.3569828 * [M/H]) + (-1 * \text{square}(-\text{logg} * \sin(\text{Teff}) + 0.0711007 * \text{logg}) * \exp(0.3569828 * [M/H]) * \sin(\text{logg} + [M/H]) + 0.1284640 * \exp(0.3569828 * [M/H]) * \sin(\text{logg} + [M/H]) - 1 * \exp(0.3569828 * [M/H])) / (1 * [M/H] * \sin(\text{logg} + [M/H]) - 1 * \text{square}(-\text{logg} * \sin(\text{Teff}) + 0.0711007 * \text{logg}) * \sin(\text{logg} + [M/H]))$
Cr I	854.89	Profundidad	$\text{cube}(0.4593918 / \text{Teff} ** 2) / \text{logg}$
Cr I	854.89	Desviación	0
CN I	854.92	Profundidad	$\text{cube}(\text{square}(0.5884600 / \text{Teff}))$
CN I	854.92	Desviación	0
Fe I	854.99	Profundidad	$0.2303497 / \text{square}(33.5576505 * \text{square}(-18.1243528 * \text{Teff}))$
Fe I	854.99	Desviación	$\text{cube}(-0.0430661 / \text{cube}(\text{Teff}))$
Si I	855.04	Profundidad	$\sin(10.5658767 * \text{cube}(3170988.2563258 * \text{logg} ** 2 * \exp(-\text{logg}) / \text{Teff} ** 2 + 3170988.2563258 * \text{logg} * [M/H] * \exp(-\text{logg}) / \text{Teff} ** 2 + 5345807.6980373 * \text{logg} * \exp(-\text{logg}) / \text{Teff} ** 2))$
Si I	855.04	Desviación	$0.1258958 * [M/H] ** 2 * \exp(\sin(\text{logg})) * \sin(3.2007406 * \text{cube}(\sin(\text{Teff}))) + 0.1258958 * [M/H] * \exp(\sin(\text{logg})) * \sin(\text{Teff}) * \sin(3.2007406 * \text{cube}(\sin(\text{Teff})))$
Cr I	855.55	Profundidad	$\text{cube}(\text{logg} + \text{logg} / \cos(5.8106510 * \text{Teff}) + \cos(\text{square}(2 * [M/H])) / \cos(5.8106510 * \text{Teff})) / \text{cube}(\text{Teff} / \text{square}(\text{logg}))$
Cr I	855.55	Desviación	$\sin(\text{cube}(\sin(\text{square}(-6.4363463 * \text{logg} * \sin(\text{Teff}) / [M/H] - 1.7890726) / (\text{Teff} * \sin(\exp(\text{logg} / \sin(\text{Teff}))))))$
CN I	855.56	Profundidad	$\sin(\exp(0.1727893 * [M/H] / \sin(0.3033987 * \text{Teff})) * \exp(0.2842810 / \sin(0.3033987 * \text{Teff})) / (\text{Teff} * \text{logg}))$
CN I	855.56	Desviación	$0.2273600 - 0.5278895 * \text{cube}(2 * [M/H] * \text{square}(\cos(\text{Teff} / ([M/H] - 2.1640632) + \text{square}(\text{Teff}) - 0.8784949)) + \text{square}(\cos(\text{Teff} / ([M/H] - 2.1640632) + \text{square}(\text{Teff}) - 0.8784949)) * \cos(0.1974634 * \text{Teff} * [M/H] + 0.6867637) + 1.2169906 * \text{square}(\cos(\text{Teff} / ([M/H] - 2.1640632) + \text{square}(\text{Teff}) - 0.8784949))) / (0.5278895 * \text{logg} * \sin(\text{Teff} * \text{logg}) - 1)$
Si I	855.68	Profundidad	$\text{square}(\sin(\sin(0.1604203 * \sin(\text{square}(\sin(0.9102878 * \text{Teff} - 0.9102878 * \sin(\text{square}(0.1002863 * \text{logg} - 1.0593976)))))) * \exp([M/H]) +$

			$0.1828192 * \text{square}(\sin(0.9102878 * \text{Teff} - 0.9102878 * \sin(\text{square}(0.1002863 * \text{logg} - 1.0593976)))) + 0.1828192))$
Si I	855.68	Desviación	$\text{cube}(\text{square}(-0.0001043 * \text{Teff}))$
Fe I	855.97	Profundidad	$1 * (-17.4153451 * \text{logg} + 169.3880900 * \exp([M/H])) / \text{Teff}$
Fe I	855.97	Desviación	$\text{cube}(-\sin([M/H] * \exp(\cos(\text{Teff})) * \cos(\text{Teff})) / (-[M/H] + \text{square}(\text{logg}) + 0.8079100)) - 0.2076467)$
Fe I	856.16	Profundidad	$\sin(5.1142854 / (0.0339823 * \text{Teff} * \text{logg} * \text{square}([M/H] + \text{square}([M/H] - 0.8641807)) + \text{logg} * \text{cube}(3.4749910 / \text{logg})))$
Fe I	856.16	Desviación	$-0.0314677 * \exp([M/H])$
Fe I	856.21	Profundidad	$\text{cube}(1 / \text{cube}(\text{Teff}))$
Fe I	856.21	Desviación	$\text{cube}(\text{cube}(0.6215193 * \text{logg} / \text{Teff}))$
Ni I	856.54	Profundidad	$\text{square}(-\sin(\sin(\sin(0.1839720 - \text{square}(\text{square}(1.0305486 * \exp(1 / \text{Teff}) * \exp(-0.1734684 * \text{logg}) * \exp([M/H]) * \exp(-0.1734684 * \sin(\text{square}(\text{Teff})) + 5.6590470)) / \text{Teff}))))))$
Ni I	856.54	Desviación	$-45.9033280 * \text{logg} / (1 * \text{Teff} + 1 * \text{cube}(-12.3927010 * \exp(\sin(\text{Teff}))))$
CN I	856.7	Profundidad	$\cos(\text{logg}) / (\text{Teff} * \text{logg})$
CN I	856.7	Desviación	$\text{cube}(31.1178948 * [M/H] * \exp(\cos(\text{Teff} - 0.6094887))) / (\text{Teff} * \text{logg} * \text{square}(\text{logg}))$
Fe I	856.78	Profundidad	$\text{square}(-\text{square}(-0.2574091 * [M/H] - 0.2021941) * \sin(\text{square}(\text{Teff}) - 1.0845592) + 0.7183294 * \text{square}(-0.2574091 * [M/H] - 0.2021941))$
Fe I	856.78	Desviación	$\text{cube}(\text{square}(1.4808736 * \text{cube}(1.1906737 * \text{logg} ** 3 / (\text{Teff} * \sin(\text{Teff} - 1.4836079)) - 1.1906737 * \text{logg} ** 2 * \sin(\cos(\text{Teff} ** 3 / (\text{logg} * [M/H] + 0.5612900 * \text{logg} - 0.0081289 * [M/H] - 0.0045627) - \text{Teff} * [M/H] / (\text{logg} - 0.0081289))) / (\text{Teff} * \sin(\text{Teff} - 1.4836079))))))$
Co I	856.96	Profundidad	$\text{square}(\text{cube}(-\sin(2 * \text{Teff} + 0.2815290 * \text{logg} * \exp(-\text{logg}) + 0.1124621 * [M/H] + 0.1032475)))$
Co I	856.96	Desviación	$0.0000034 / (0.4483208 * \text{cube}([M/H]) * \cos(\text{Teff}) + 1 * \text{cube}([M/H]))$
Fe I	857.02	Profundidad	$\text{cube}(1 / \text{square}(\text{square}(\text{Teff})))$
Fe I	857.02	Desviación	$0.3655309 * \exp(-\text{Teff})$
Si I	857.08	Profundidad	0

Si I	857.08	Desviación	$\sin(\text{cubecube}(0.0059271/\text{Teff}))$
Si I	857.13	Profundidad	$\sin(0.0000023*\text{Teff}*[M/H] + 0.0000038*\text{Teff} - 0.0097975*[M/H] - 0.0162379)$
Si I	857.13	Desviación	$\text{cubecube}(\sin(0.5233903*\text{Teff}))$
Fe I	857.18	Profundidad	$-0.0038101*\text{logg} - 1*\text{cubecube}(\sin(1/([M/H] - 3.1764590)))$
Fe I	857.18	Desviación	$-0.0018561*[M/H]/\text{square}(\sin(1/\sin(0.1665242*\text{Teff})))$
CN I	857.33	Profundidad	$\text{square}(-43.0665200*\exp(2.5492590*[M/H])/(Teff*\sin(0.5853986*\text{square}(Teff))))$
CN I	857.33	Desviación	$\sin(\exp(\text{logg}*[M/H])/(1.4470879*\text{Teff}*\sin([M/H])*\cos(1.5534155*\text{Teff} - 0.3750306*\exp(\text{logg})) + \text{cubecube}([M/H])*\sin([M/H])*\cos(1.5534155*\text{Teff} - 0.3750306*\exp(\text{logg})))$
Co I	857.45	Profundidad	0
Co I	857.45	Desviación	$[M/H]/\text{square}(\text{square}(\text{square}(Teff)))$
Co I	857.53	Profundidad	$\sin(\text{logg}**3*\text{square}(\text{cubecube}(1.0540727*\cos(0.7710425*\text{logg}))*\text{cubecube}(\cos(Teff*\text{cubecube}(0.7092983))))*\exp([M/H])/Teff)$
Co I	857.53	Desviación	$\text{logg}*\text{square}(3.3206327*\exp([M/H])*exp(\sin(Teff + 0.6837800)))/Teff$
Fe I	857.65	Profundidad	1/Teff
Fe I	857.65	Desviación	$\sin(\text{square}(\sin(\text{square}(-0.0861606*\text{cubecube}([M/H])*\text{square}(\sin(\text{square}(Teff - 0.8412462))))*\sin(\text{square}(\text{logg} - 0.5547559) + 1.0261389*\text{square}(\sin(\text{square}(Teff - 0.8412462))))*\sin(\text{square}(\text{logg} - 0.5547559) + \sin(\text{square}(\text{logg} - 0.5547559)*\cos(Teff))))))$
Si I	857.91	Profundidad	$\sin(0.0061283*\text{square}(\exp([M/H]))*\exp(\cos(\text{square}(\text{square}(-0.3110441*\text{Teff}))))/\text{logg}$
Si I	857.91	Desviación	$0.0137512*\exp(-\text{logg})*\cos(Teff)/[M/H]$
Ni I	857.99	Profundidad	$\text{cubecube}(\text{cubecube}(-0.6984227/\text{Teff}))$
Ni I	857.99	Desviación	0
Fe I	858.23	Profundidad	$\sin(\sin(\sin(-151.0505886*[M/H]*\exp([M/H])/(Teff - 0.0066203*\text{Teff}*\text{cubecube}(-1.4407791*\text{logg}))/(\sin([M/H]/(\text{logg} - [M/H]) + \text{cubecube}(Teff)) + 3.0422480)) + 151.0505886*\exp([M/H])*sin(2.5536908*\text{Teff} + 0.0169062)/(Teff - 0.0066203*\text{Teff}*\text{cubecube}(-1.4407791*\text{logg}))/(\sin([M/H]/(\text{logg} - [M/H]) + \text{cubecube}(Teff)) + 3.0422480)) +$

			$459.5333510 \cdot \exp([M/H]) / (\text{Teff} - 0.0066203 \cdot \text{Teff}^3 - 1.4407791 \cdot \log) / (\sin([M/H] / (\log - [M/H]) + \text{cube}(\text{Teff})) + 3.0422480))$
Fe I	858.23	Desviación	$-2.6045954 / (1 \cdot [M/H] \cdot \cos(1.4941061 \cdot \text{Teff} \cdot \log) + 0.6997598 \cdot [M/H])$
Cr I	858.29	Profundidad	$\text{Teff} \cdot \log^2 \cdot \text{square}(\text{cube}(0 \cdot \text{square}(1.0605249 \cdot \text{Teff}) \cdot \sin(1046705500000000000 / (\log - [M/H]))) \cdot \sin(\sin(\text{Teff} \cdot \text{square}(\text{cube}(0 \cdot \text{square}(\text{Teff} \cdot \sin(1070869959763130048147619840 \cdot \text{Teff} \cdot \log - 1070869959763130048147619840 \cdot \text{Teff} \cdot [M/H])))$
Cr I	858.29	Desviación	$0.0787876 \cdot \text{cube}(\text{cube}(\text{square}(\cos(0.0003974 \cdot \text{Teff}))))$
Fe I	858.48	Profundidad	$1 \cdot (5.7683606 - 1 \cdot \log) / (\text{Teff} \cdot \text{cube}(0.2917461 \cdot \exp(-0.6120107 \cdot [M/H])))$
Fe I	858.48	Desviación	$(-0.5258367 \cdot \log^2 \cdot [M/H] \cdot \sin(\text{cube}(\text{Teff})) \cdot \cos(0.1510664) - 0.5258367 \cdot \log^2 \cdot [M/H] - 0.5258367 \cdot \log^2 \cdot \sin(\text{cube}(\text{Teff})) \cdot \cos(0.1510664) - 0.5258367 \cdot \log^2 \cdot \sin(\text{cube}(\text{Teff})) / (0.4729443 \cdot \log^2 \cdot [M/H]^3 \cdot \cos(0.1510664) - 0.2803229 \cdot \log^2 \cdot [M/H]^2 \cdot \cos(0.1510664) + 0.0254412 \cdot \log^2 \cdot [M/H] \cdot \cos(0.1510664) - 1 \cdot [M/H]^3 \cdot \cos(0.1510664) + 0.5927187 \cdot [M/H]^2 \cdot \cos(0.1510664) - 0.0537933 \cdot [M/H] \cdot \cos(0.1510664))$
CN I	858.52	Profundidad	$6.4180885 / (\text{Teff}^2 \cdot \text{square}(\cos(\text{Teff})) \cdot \cos(\text{cube}(1.1480450 - \text{Teff}) \cdot \text{square}(\cos([M/H])))$
CN I	858.52	Desviación	$\text{cube}(-\sin(1.3494020 \cdot \text{Teff}) / (\text{Teff} \cdot [M/H] \cdot \cos(\text{Teff}) - \text{Teff} \cdot \text{square}(\cos(\log)) \cdot \cos(\text{Teff}))$
Co I	858.67	Profundidad	$\text{square}(\text{square}(\text{square}(1/\text{Teff})))$
Co I	858.67	Desviación	0
Fe I	858.79	Profundidad	0
Fe I	858.79	Desviación	$\text{square}(\text{square}(0.9884979/\text{Teff}^2))$
Fe I	858.83	Profundidad	$0.0826357 \cdot \log \cdot \exp(-\log) \cdot \exp(\log \cdot [M/H]) \cdot \exp(\text{cube}([M/H]))$
Fe I	858.83	Desviación	$\sin(\text{cube}(-\sin(\sin(\sin(\sin(\sin(1.6534309 \cdot \text{Teff} \cdot \sin(4.6270661 \cdot \exp(1.2144313 \cdot [M/H]) \cdot \exp([M/H] \cdot \exp([M/H])) \cdot \exp(-\exp([M/H]) \cdot \sin(1.1722236 \cdot \text{Teff})) \cdot \exp(1.2614324 \cdot \exp([M/H])) \cdot \exp(-1.2144313 \cdot \sin(1.1722236 \cdot \text{Teff}) / \text{square}(-1.3059305 \cdot \text{Teff})) / (0.2333514 - 0.2504994 / (\log - 0.3632525)) + 2.8297570 \cdot \sin(4.6270661 \cdot \exp(1.2144313 \cdot [M/H]) \cdot \exp([M/H] \cdot \exp([M/H])) \cdot \exp(-\exp([M/H]) \cdot \sin(1.1722236 \cdot \text{Teff})) \cdot \exp(1.2614324 \cdot \exp([M/H])) \cdot \exp(-1.2144313 \cdot \sin(1.1722236 \cdot \text{Teff}) / \text{square}(-1.3059305 \cdot \text{Teff})) / (0.2333514 - 0.2504994 / (\log - 0.3632525))))))))))$

Co I	858.97	Profundidad	$\sin(\sqrt{-0.0655978 \cdot \exp([M/H]) \cdot \exp(\sin(\sin(\log g)))} \cdot \cos(0.4064565 \cdot T_{\text{eff}}))$
Co I	858.97	Desviación	$28.3579430 \cdot \cos(T_{\text{eff}}) / (T_{\text{eff}} \cdot \log g \cdot [M/H])$
CN I	859.15	Profundidad	$1 \cdot (-10.7588728 \cdot \log g \cdot \text{cub}(-\exp([M/H]) + \sin(\sqrt{T_{\text{eff}}})) \cdot \exp([M/H]) + 10.7588728 \cdot \log g \cdot \text{cub}(-\exp([M/H]) + \sin(\sqrt{T_{\text{eff}}})) \cdot \sin(\log g + 3.2192640) - 1.7644743 \cdot \text{cub}(-\exp([M/H]) + \sin(\sqrt{T_{\text{eff}}})) \cdot \exp([M/H]) + 1.7644743 \cdot \text{cub}(-\exp([M/H]) + \sin(\sqrt{T_{\text{eff}}})) \cdot \sin(\log g + 3.2192640)) / (T_{\text{eff}} \cdot \log g)$
CN I	859.15	Desviación	$0.0164814 \cdot [M/H] \cdot \text{cub}([M/H] + 1.2662773) \cdot \exp(\sin(\log g)) \cdot \sin(T_{\text{eff}}) + 0.0164814 \cdot \text{cub}([M/H] + 1.2662773) \cdot \exp(\sin(\log g)) \cdot \sin(T_{\text{eff}}) \cdot \sin(\cos(\log g) + 0.7753303)$
Fe I	859.3	Profundidad	$(0.0367378 \cdot [M/H] + 0.0424747) / (0.0442605 \cdot [M/H] \cdot \text{cub}(\log g - 1.3457156) + 0.0511721 \cdot \text{cub}(\log g - 1.3457156) + 1)$
Fe I	859.3	Desviación	$-0.0247535 \cdot \sin(\log g - 0.1547317 + 1.2166482 / [M/H])$
Si I	859.6	Profundidad	$\sqrt{-0.0387545 \cdot \log g \cdot \cos(0.7130163 \cdot T_{\text{eff}})}$
Si I	859.6	Desviación	$\sqrt{\text{cub}(-0.1798640 \cdot \log g \cdot [M/H] - 0.0722781 \cdot \log g)}$
Si I	859.7	Profundidad	$1 / (\log g \cdot \text{cub}(\log g - \text{cub}([M/H] - 2.3087597 / \log g) - \sin(2.4747058 \cdot T_{\text{eff}})))$
Si I	859.7	Desviación	$1.9469627 \cdot \text{cub}(\exp([M/H]) \cdot \exp(\cos(2.4512711 \cdot \log g))) / (T_{\text{eff}} \cdot \sin(1 / \text{cub}(\cos(T_{\text{eff}}))))$
Ti I	859.82	Profundidad	$0.0000188 \cdot T_{\text{eff}} - 0.0822211$
Ti I	859.82	Desviación	$\sin(\sin(\text{cub}(\sqrt{\sqrt{\sin(1.7867266 \cdot \sqrt{\cos(1.4135632 \cdot T_{\text{eff}} - 1.0000026)})} \cdot \sin(\log g + [M/H] - 0.5921045) + 4.3511899 \cdot \sqrt{\cos(1.4135632 \cdot T_{\text{eff}} - 1.0000026)}))))))$
Fe I	859.88	Profundidad	$0.0596865 \cdot \exp([M/H])$
Fe I	859.88	Desviación	$\sqrt{T_{\text{eff}} \cdot \sqrt{-0.0066362 \cdot \cos([M/H])}}$
Fe I	860.1	Profundidad	$(1 \cdot [M/H] + 1 \cdot \exp(\exp([M/H]))) / (1 \cdot T_{\text{eff}} \cdot \sqrt{\sin(\exp(448.1901337 / T_{\text{eff}}) \cdot \exp(T_{\text{eff}} / (T_{\text{eff}} + \text{cub}(\log g + 1.5787910))))} + 0.0049883 \cdot T_{\text{eff}})$
Fe I	860.1	Desviación	$-0.0126407 \cdot \exp(\sin(\log g - 3.7918040 \cdot [M/H] + 3.6930659)) + 1 \cdot (8.1603695 \cdot \log g \cdot \exp(\sin(\log g - 3.7918040 \cdot [M/H] + 3.6930659)) + 80.0789893 \cdot \exp(\sin(\log g - 3.7918040 \cdot [M/H] + 3.6930659))) / T_{\text{eff}}$
CN I	860.39	Profundidad	$\sqrt{508.2494500 \cdot \exp([M/H]) / T_{\text{eff}}}$
CN I	860.39	Desviación	$-0.0621415 \cdot \sqrt{\sin(\text{cub}(-0.2448521 / (\log g \cdot [M/H])))}$

Si I	860.6	Profundidad	$\exp([M/H]) \cdot \exp(3.9284966 \cdot \cos(0.8177213 \cdot \text{Teff} - [M/H])) / \text{Teff}$
Si I	860.6	Desviación	$0.0199678 \cdot \text{cube}(\sin(1.3621646 \cdot \text{Teff}) + 0.2099347)$
Ni I	860.64	Profundidad	$\sin(0.6304999 \cdot \text{square}(\text{cube}(\sin(\sin(\sin(\sin(\sin(\exp(-0.2266046 \cdot \text{logg})) \cdot \exp([M/H]) \cdot \exp(\sin(\cos(1.3598443 \cdot \text{Teff} + 1.3598443 \cdot [M/H]))) \cdot \text{Teff})))))))))$
Ni I	860.64	Desviación	$-0.0249331 / \text{logg}$
Fe I	860.71	Profundidad	$-\sin(4.6486370 \cdot \exp(\sin(\text{Teff} - 1.3340008))) / (-\text{Teff} / ([M/H] + 1.1052930) + 703.2118500 + 2471.2969000 / ([M/H] + 1.1052930)) + 4.6486370 \cdot \sin(\text{logg}) / (-\text{Teff} / ([M/H] + 1.1052930) + 703.2118500 + 2471.2969000 / ([M/H] + 1.1052930))$
Fe I	860.71	Desviación	$\sin(\text{cube}(0.0989413 \cdot \text{Teff} \cdot \cos(\text{logg}) - 0.0989413 \cdot \text{cube}(3.9799535 / [M/H]) \cdot \cos(\text{logg})))$
CN I	861.01	Profundidad	$\text{cube}(\text{cube}(1 / \text{Teff}))$
CN I	861.01	Desviación	0
Fe I	861.06	Profundidad	$\sin(\text{square}(\text{square}(-0.0372637 / \text{square}(\text{Teff}))))$
Fe I	861.06	Desviación	$\text{cube}(0.1433674 \cdot [M/H] / \text{square}(\text{Teff}))$
Fe I	861.18	Profundidad	$\sin(\sin(\text{square}(\sin(-81.0492192 \cdot \text{logg} / \text{Teff} + 347.8406746 \cdot [M/H] / \text{Teff} + 81.0492192 \cdot \sin(1.2959286 \cdot \text{Teff}) / \text{Teff} + 1715.2405850 / \text{Teff} + 32.5438113 / (\text{Teff} \cdot \text{logg}))))))$
Fe I	861.18	Desviación	$0.8281975 \cdot \text{square}([M/H] + 1.4951494 / \text{logg}) / (0.0032185 \cdot \text{Teff} \cdot \text{square}([M/H] + 1.4951494 / \text{logg}) + 1 \cdot \text{logg} - 1 \cdot [M/H] - 0.0611590)$
Fe I	861.39	Profundidad	$0.0420932 \cdot \text{square}(\exp([M/H]) \cdot \cos(0.2328420 \cdot \text{logg}))$
Fe I	861.39	Desviación	$\text{cube}(0.2567878 \cdot \cos([M/H] + 0.2996601))$
Fe I	861.63	Profundidad	$0.0843985 \cdot \exp(-0.2058512 \cdot \text{logg}) \cdot \exp([M/H])$
Fe I	861.63	Desviación	$0.6819228 \cdot \exp(1.8812199 \cdot \text{square}(1.8137726 \cdot \exp([M/H]) \cdot \sin(\text{square}(23.3056370 \cdot \text{logg} \cdot [M/H] / \sin(\text{Teff}) + 23.3056370 \cdot [M/H] \cdot \text{Teff} / \sin(\text{Teff}) - 3.7938296 \cdot [M/H]))) \cdot \exp(\text{cube}(\text{cube}(\sin(\text{logg} \cdot [M/H] / \sin(\text{Teff}) + [M/H] \cdot \text{Teff} / \sin(\text{Teff})))) \cdot \text{square}(\exp([M/H])) \cdot \sin([M/H])) - 1.1111541$
Fe I	861.81	Profundidad	$17.7043229 \cdot \exp([M/H]) \cdot \exp(-\sin(0.0196529 \cdot \text{Teff})) / \text{Teff}$
Fe I	861.81	Desviación	$-0.0129569 \cdot \sin(\text{logg} \cdot [M/H])$

Ni I	862.01	Profundidad	$[M/H]/\text{cube}(\text{logg} + \text{cube}(\text{Teff}))$
Ni I	862.01	Desviación	$\text{square}(\text{cube}(\text{square}(1/\text{Teff})))$
Fe I	862.16	Profundidad	$-0.0062504*\text{logg} + 0.0434564*\sin([M/H] - \sin(0.9623603*\text{Teff} - 1.5660009)) + 0.0762068$
Fe I	862.16	Desviación	$0*\text{Teff}*\cos(\text{cube}(\text{logg} + [M/H]))/(\text{cube}(\exp(\sin(\text{Teff})))\cos(\text{Teff})) + 0.0118841$
Fe I	863.24	Profundidad	$\text{square}(-\exp([M/H])/(0.0012337*\text{Teff} + \text{logg} - 2.3241506))$
Fe I	863.24	Desviación	$\sin(\sin(\sin(\text{square}(11.6978725*\text{logg}*\text{cube}(\exp(\sin(\sin(\text{logg}*[M/H] + 2.7550110))))/\text{Teff}))))$
S I	863.31	Profundidad	$\sin(\text{square}(0.0000002*\text{Teff}*\text{square}(0.0019572*\text{Teff}) - 0.0001287*\text{logg}*\text{square}(0.0019572*\text{Teff}) + 0.0001287*[M/H]*\text{square}(0.0019572*\text{Teff}))$
S I	863.31	Desviación	$\text{square}(\text{square}(\text{square}(0.8085244*\sin([M/H]*\text{square}(0.3434113*\text{logg}))*\sin(0.6192942*\text{Teff} - \text{logg}))))$
Cr I	863.62	Profundidad	$\exp([M/H])/\text{cube}(1.1118558*\text{Teff}*\text{square}(0.0000071*\text{Teff}))$
Cr I	863.62	Desviación	$18.2149101*\text{logg}*\cos(16.4153820*[M/H]/\text{logg} + 21.6041386/\text{logg} - 16.4153820*\text{cube}(-0.9997269*\text{logg})/(\text{Teff}*\text{logg}*\sin(\text{Teff} - 1.7177280)))/\text{Teff}$
Ni I	863.7	Profundidad	$0.0284718*\exp(2*[M/H])/\text{logg}$
Ni I	863.7	Desviación	$-\text{square}(\text{cube}(1.1298776/\text{square}(\cos(\sin(\sin(1.0459923*\text{Teff} + 1.0459923*[M/H]))))))/(\text{Teff} - [M/H])$
Cr I	864.3	Profundidad	$\sin(\sin(\sin(\text{square}(0.0804942*\text{Teff}/(1.8588628*\text{Teff} + 1/\text{cube}(-1/(1.0041662*[M/H] + 18.0133349 - 15.7281252*\text{cube}(\text{logg})/\text{Teff} + 7.2593096/\text{Teff})))) - 0.0033589))$
Cr I	864.3	Desviación	$-0.0381483*\cos(1.6552044*\text{Teff})*\cos(\text{logg} - [M/H])$
Fe I	864.33	Profundidad	$0.1310896*\sin(0.0836487*\exp(-\cos(724.7638000*\text{Teff}))/\text{logg}) - 0.0171845*\sin(0.0836487*\exp(-\cos(724.7638000*\text{Teff}))/\text{logg})/(1*\exp(5.6059650*\sin([M/H])) + 0.1310896)$
Fe I	864.33	Desviación	$\text{cube}(\sin(\sin(\sin(\sin(\text{cube}(-\sin(1.4245716*\text{Teff}))/\text{logg}))))$
Si I	864.64	Profundidad	$\text{square}(5.0110285*\text{logg}*\exp([M/H])/(\text{Teff}*\text{cube}(\exp(\sin(0.5216491*\text{Teff} + 0.2608246/(\sin(1.2893342*\text{Teff}/[M/H]) + [M/H]/\text{logg}))))*\cos(\exp([M/H])) + 0.1938072*\text{cube}(\exp(\sin(0.5216491*\text{Teff} + 0.2608246/(\sin(1.2893342*\text{Teff}/[M/H]) + [M/H]/\text{logg}))))*\exp(\text{logg}^2)*\cos(\exp([M/H])))$

Si I	864.64	Desviación	$-\sin(1/(0.2421618 * T_{\text{eff}} + 0.7256159 * \text{cube}(\log g + \sin(T_{\text{eff}}) - 1.1867301/[M/H])/ \text{square}(-0.3056470/\log g) - 0.5035166/\text{square}(-0.3056470/\log g)))$
Fe I	864.78	Profundidad	$\text{square}(\text{cube}(1.3442721/\text{cube}(T_{\text{eff}})))$
Fe I	864.78	Desviación	$\exp(-0.3780302 * T_{\text{eff}})$
Si I	864.85	Profundidad	$\sin(\sin(\text{square}(0.3177639 * \cos(\sin(1/(\exp(\sin([M/H] - \sin(0.5621256)))) * \sin(0.6510599 * T_{\text{eff}} + 0.5631281/\log g) + 2.7051554 * \exp(\sin([M/H] - \sin(0.5621256))))))))))$
Si I	864.85	Desviación	$\text{cube}(\text{square}(\sin(3.9056249 * T_{\text{eff}})))$
Fe I	865.25	Profundidad	$-0.0238025 + 1 * \text{cube}([M/H] + 5.8520646)/T_{\text{eff}}$
Fe I	865.25	Desviación	$\text{cube}(-0.2907217 * \exp([M/H]))$
Fe I	865.41	Profundidad	$\text{cube}(\text{square}(0.6404873/T_{\text{eff}}))$
Fe I	865.41	Desviación	$\text{cube}(\text{cube}(\log g/T_{\text{eff}}))$
Fe I	865.44	Profundidad	$2.5416436 * \text{cube}(1.7504772 * [M/H] + \sin(\log g) + 3.6768879)/T_{\text{eff}}$
Fe I	865.44	Desviación	$0.0790854/(1 * \log g + 0.6776123)$
Fe I	865.67	Profundidad	0
Fe I	865.67	Desviación	$\text{cube}(\text{square}(-0.2342865/T_{\text{eff}}))$
Co I	866.11	Profundidad	$\text{cube}(\text{cube}([M/H]/T_{\text{eff}}))$
Co I	866.11	Desviación	0
Si I	866.12	Profundidad	$\exp(-0.3812157 * T_{\text{eff}})$
Si I	866.12	Desviación	0
Fe I	866.19	Profundidad	$5.4529628 * \exp(-T_{\text{eff}})$
Fe I	866.19	Desviación	0
Ca II	866.21	Profundidad	$\text{square}(\sin(\sin(\sin(2.1366665/T_{\text{eff}}))))$
Ca II	866.21	Desviación	$0.0434984 * [M/H] - 0.0434984 * \sin(0.0011901 * T_{\text{eff}} - 1.3542058) + 0.1564417 + 0.0434984/\log g$

Ca II	866.22	Profundidad	$\sin(\sin(\text{cube}(\sin(\sin(\text{square}(-\sin(1/(0.1590806*\text{logg}*\sin(0.5782931) - 4.9691039*\sin(0.5782931) + 15.5830718*\exp(-\cos(\text{logg}))/(\sin(\text{logg}) - 2.5813682)))*\exp(\sin(0.9430461*\text{Teff}))*\sin(0.5782931))))))))))$
Ca II	866.22	Desviación	$\text{square}(\text{square}(\text{cube}(-\sin(6.2932413*\text{Teff}))))$
Ca II	866.23	Profundidad	$0.0001225/(0.0001225*\text{square}(\text{square}(-\text{Teff} - \text{logg} - \text{square}(\text{square}(\text{logg} - 0.7497073)) + 8164.1455000))*\text{square}(\text{square}(-0.5351206*\sin(0.7807372*\text{logg})) - 1*\text{square}(\text{square}(-0.5351206*\sin(0.7807372*\text{logg}))))$
Ca II	866.23	Desviación	$\sin(\text{cube}(1/(\text{logg}*\text{square}(-1.3484063*\text{logg} + 81.5894968*\sin(1.0592745*\text{Teff} - 1.0592745*\text{logg}/\cos(\text{Teff} - \text{logg})) + 1.0592211*\cos(0.3005115*[M/H]) + 1.5553443) - 8.3323440*\text{square}(-1.3484063*\text{logg} + 81.5894968*\sin(1.0592745*\text{Teff} - 1.0592745*\text{logg}/\cos(\text{Teff} - \text{logg})) + 1.0592211*\cos(0.3005115*[M/H]) + 1.5553443))))$
Ca II	866.24	Profundidad	$\text{square}(\text{logg}/\text{square}(\text{cube}(\text{Teff})))$
Ca II	866.24	Desviación	$\text{square}(-0.1512671/(\text{Teff}*\text{square}(\text{Teff})))$
Fe I	866.37	Profundidad	0
Fe I	866.37	Desviación	$0.0665993*\text{square}(\text{cube}(\sin([M/H]/\text{Teff})))$
H I	866.5	Profundidad	$\text{square}(-0.7250072*[M/H]*\text{cube}(0.0001591*\text{Teff} - 0.5112405) - 1.4500143*\text{cube}(0.0001591*\text{Teff} - 0.5112405)*\cos(\text{logg}))$
H I	866.5	Desviación	$-\sin(1/(\text{Teff} - \text{cube}(\text{square}(-\text{logg} - 0.2441227))))$
Si I	866.74	Profundidad	$\sin(\text{square}(\text{logg}*\text{square}(\text{logg}*\exp(\cos(0.1398683*\text{logg} + 0.7160383)/[M/H]))*\exp(1.3900956*\cos(0.1226996*\text{Teff} + 0.1226996*\text{logg}))/\text{Teff}))$
Si I	866.74	Desviación	$\text{cube}(-2.5083077/(-\text{Teff}*[M/H]/\text{square}([M/H] + \sin(\text{logg}) + 4.1067118) + 4.0837654*\text{logg} + 4.0837654*\sin(\text{Teff})))$
S I	866.84	Profundidad	$\text{cube}(1/(\text{Teff}*\text{square}(0.2589413 - \text{Teff})))$
S I	866.84	Desviación	$\text{cube}(1.1248729/\text{cube}(\text{Teff}))$
Fe I	866.89	Profundidad	$\text{cube}(\text{cube}(\text{cube}(\text{square}(0.1210507))))$
Fe I	866.89	Desviación	$\text{cube}(\text{cube}([M/H]/\text{Teff}))$
S I	867.02	Profundidad	$0.0000034*\text{Teff}*[M/H]/(-0.0255442*\text{Teff}*[M/H]*\text{cube}(\text{logg}) + 0.0255442*\text{Teff}*\cos(\text{Teff}) - 1*\text{logg}*[M/H])$

Si I	867.02	Desviación	$-1/(Teff*\sin(Teff - [M/H]/\text{square}(\text{logg} + 0.6595848) + \text{cube}(1.8852239)))$
Si I	867.06	Profundidad	$\text{square}(0.1742148*[M/H] + 0.1742148)/\text{logg}$
Si I	867.06	Desviación	$-0.0175695*\sin(\text{logg}*[M/H] - 0.7036429*\text{logg})$
Mn I	867.09	Profundidad	$\sin(6.2035925*\text{cube}(\exp([M/H])* \exp(\cos(38035.1766715/Teff))))/Teff$
Mn I	867.09	Desviación	$\sin(\exp(-2.6149895*\exp(-[M/H])* \exp(\exp(\cos(-0.7936176*Teff + 0.6367725 + 1/[M/H]))) * \exp(\exp(-[M/H])* \exp(\exp(\cos(-0.7936176*Teff + 0.6367725 + 1/[M/H]))) * \sin(\text{logg}))))$
Si I	867.13	Profundidad	$\text{logg}*\text{cube}(\exp([M/H])* \exp(\sin(0.3122310*Teff)))/Teff$
Si I	867.13	Desviación	$Teff**2*\text{logg}*\text{square}(\exp([M/H]))/(\text{cube}(Teff)*\sin(861.2204786*Teff*\sin([M/H]) - 130.2403750*\sin([M/H])))$
Fe I	867.19	Profundidad	$1*(15.1172470*\text{logg}*[M/H] + 54.0190301)/Teff$
Fe I	867.19	Desviación	$0.0121832*\cos(-1.1712877*Teff)$
Mn I	867.2	Profundidad	$\text{square}(-409.9383500*[M/H]*\sin(83592.0550000*Teff)/Teff - 748.2258305*\sin(83592.0550000*Teff)/Teff)$
Mn I	867.2	Desviación	$0.0264118*\exp([M/H])$
Mn I	867.39	Profundidad	$\text{cube}(1/\text{cube}(\text{square}(Teff)))$
Mn I	867.39	Desviación	$\text{cube}(\text{cube}([M/H]/Teff))$
Fe I	867.47	Profundidad	$\text{square}(\text{square}(0.5921013*\cos(0.9116209*Teff + 0.4308263*\cos(2.1159824*Teff + \cos(Teff))))))$
Fe I	867.47	Desviación	$-38.9781340*\text{logg}*\cos(Teff) - 39.5039412*\text{logg} + 1*(1.3441982*\text{logg}*\sin(51.2060200*\text{logg}*\cos(Teff) + 0.3681352)*\cos(Teff) + 1.3623312*\text{logg}*\sin(51.2060200*\text{logg}*\cos(Teff) + 0.3681352) - 1.3194132*\text{logg}*\cos(Teff) - 1.3372118*\text{logg} + 1.9157213*\sin(51.2060200*\text{logg}*\cos(Teff) + 0.3681352)*\cos(Teff) + 1.9415640*\sin(51.2060200*\text{logg}*\cos(Teff) + 0.3681352) - 1.8803983*\cos(Teff) - 1.9057645)/\text{square}(-\sin(\text{logg} + 0.9633075 + \exp(-\text{cube}([M/H]))) - 55.5507674*\cos(Teff) - 56.3001361$
Ti I	867.54	Profundidad	$11.5596833*\text{square}(\text{logg}*\text{cube}(\text{logg})*\text{cube}(1/(0.0590614*Teff*\exp(\sin(Teff)))/(\text{logg}*[M/H]) + \cos(1.2865735*\text{cube}(\text{logg}*\text{cube}([M/H]))))))$
Ti I	867.54	Desviación	$\text{square}(\text{logg}*\text{cube}(\text{square}(\text{logg}*[M/H] + [M/H]*\sin(Teff**2))))/\text{square}(0.2341633*Teff)$

Fe I	867.58	Profundidad	$\text{cube}(1/(\text{Teff}*\text{square}(\text{Teff} + 1.9680064)))$
Fe I	867.58	Desviación	$0.6761022*\exp(-\text{Teff})$
Fe I	867.71	Profundidad	$\text{cube}(\exp(-4170.4023000*\text{Teff}))$
Fe I	867.71	Desviación	$\text{cube}([M/H]*\text{square}(0.8030859)/\text{Teff}**2)$
Fe I	867.96	Profundidad	$260.5277709*\text{square}(\text{square}(\text{logg})*\sin(2.4005493*[M/H])/Teff)$
Fe I	867.96	Desviación	$-452.8622692*\text{square}(-\text{logg}**2 - \text{logg}*\text{square}(0.6465847 - \sin([M/H])) - 0.7574918*\text{logg} - 0.7574918*\text{square}(0.6465847 - \sin([M/H]))) * \sin(\text{square}([M/H] - 0.6465847)) / (0.4592939*\text{Teff}**2 + 1*\text{Teff})$
Fe I	868.0	Profundidad	$126.7684760*\exp([M/H])/Teff$
Fe I	868.0	Desviación	$\sin(8.7883548*\text{logg}*[M/H]/Teff - 8.7883548*\text{logg}*\text{square}(\text{square}(\sin([M/H]) + \cos(\text{logg} - 0.4997641*[M/H]) - 1.3067797)/\text{logg})/Teff)$
N I	868.03	Profundidad	$\sin(\exp(\text{cube}([M/H] - \sin(\sin([M/H]) + \sin(64.2563345*\text{Teff} - 34.6291034) + 2.3059508) + 0.0520169) - 1.0327716))/\text{logg})$
N I	868.03	Desviación	$0.0050360/(0.1546875*\text{logg} - 1*[M/H] + 0.0417731)$
S I	868.05	Profundidad	$\text{cube}(\text{square}(0))$
S I	868.05	Desviación	$-8.5038331/\text{square}(\text{square}(\text{square}(\text{Teff})))$
Fe I	868.08	Profundidad	$\exp(-0.3746296*\text{Teff})$
Fe I	868.08	Desviación	$-0.1350518/\text{cube}(-\text{Teff}**2 + \text{Teff}*[M/H])$
Ti I	868.3	Profundidad	$-0.0469279 + (0*\text{Teff}*\text{logg} + 0.6414852*\text{Teff}*[M/H] + 17906.0370769)/(0.0073532*\text{Teff}**2 + 0.7387750*\text{Teff}*\text{logg} - 1*\text{Teff})$
Ti I	868.3	Desviación	$\text{square}(\text{cube}(-[M/H]*\sin(0.2698990/(\exp(\text{logg}) - \exp(\cos(-0.3286756*\text{logg} + 0.3286756*\text{logg}/([M/H]*\sin([M/H])) + 2.1918025)))) - 0.1942173*[M/H] + 1.5438174*\exp(3.6842039*[M/H])* \exp(2.7712571*\sin(\text{Teff}*[M/H]))*\sin(0.2698990/(\exp(\text{logg}) - \exp(\cos(-0.3286756*\text{logg} + 0.3286756*\text{logg}/([M/H]*\sin([M/H])) + 2.1918025)))) - 0.0369399*\exp(3.6842039*[M/H])* \exp(2.7712571*\sin(\text{Teff}*[M/H])))$
N I	868.34	Profundidad	0

N I	868.34	Desviación	$-0.0004535*[M/H]*\text{cube}(-661.5927000/(\text{logg}*[M/H] + \exp(\cos(\text{logg})) - 18.9893950 + 1/(\text{logg}*[M/H]) + 1/\text{Teff}))/((0.0004535*\text{Teff}*[M/H] + 1*\text{Teff} - 0.0004535*\text{logg}*[M/H] - 1*\text{logg}))$
Si I	868.64	Profundidad	$1*(0.3085761*\text{Teff}*\exp(211.0597313/\text{Teff})*\exp([M/H]) - 1233.3585865*\exp(211.0597313/\text{Teff})*\exp([M/H]))*\exp(-1974.4679000*\text{logg}/\text{Teff})$
Si I	868.64	Desviación	$1*\text{logg}*\exp(0.0694998*\text{logg}*\cos([M/H]*\text{square}(\text{Teff})*\sin(\cos(\text{Teff}))**2))*\exp(\cos([M/H]*\text{square}(\text{Teff})*\sin(\cos(\text{Teff}))**2)/\sin(\text{logg}))*\exp(\cos(\text{Teff}))/((1*[M/H] - 0.2879220))$
Fe I	868.68	Profundidad	$\text{square}(\sin(0.0248924/([M/H] - 1/(\cos(\text{cube}(\text{Teff}))*\cos(\sin(\text{logg} + 0.9077179)) + 1.9090210*\cos(\sin(\text{logg} + 0.9077179))))))*\cos(-2.2442627/\text{square}(\sin(\text{cube}(\text{Teff}))))$
Fe I	868.68	Desviación	$-\sin(5.6322517/(\text{cube}(1008.5947000*[M/H])*\text{square}(\text{square}(1 - 5.1068210/\text{logg}) + \sin(\text{Teff}**2))))$
Cr I	868.75	Profundidad	0
Cr I	868.75	Desviación	0
Fe I	868.86	Profundidad	$-0.0000944*\text{Teff} - 0.0000396*\text{Teff}/(0.4399334*[M/H] - 1) + 0.4848214$
Fe I	868.86	Desviación	$1.5310598*\exp(\text{logg})*\sin(4.9800648*[M/H])/Teff$
Fe I	868.97	Profundidad	0
Fe I	868.97	Desviación	$\text{square}(\text{cube}(\text{square}(\text{square}(0.0734373*\cos(0.5141332))))))$
Fe I	868.99	Profundidad	$\text{square}(\text{cube}(\cos(\text{Teff})/\text{square}(\text{Teff})))$
Fe I	868.99	Desviación	0
Fe I	869.15	Profundidad	0
Fe I	869.15	Desviación	$\text{cube}(\text{square}(-\sin(0.6477838/\text{Teff})))$
Ti I	869.23	Profundidad	$\text{square}(\text{square}(\sin(\text{square}(1.3799812*\sin(1.5602129*\text{Teff})))))*\sin(\cos([M/H]))$
Ti I	869.23	Desviación	$44.4832519/\cos(0.0147643*\text{Teff}*\text{logg} + 0.0147643*\text{Teff}/[M/H])$
Si I	869.4	Profundidad	0
Si I	869.4	Desviación	0
Si I	869.47	Profundidad	$0.0000062*\text{Teff}*\exp([M/H]) - 1*\sin(18.2328524*\text{logg}/\text{Teff})$

S I	869.47	Desviación	$\text{square}(0.0943121 * \text{cube}(\text{cube}(-\text{square}(\text{square}(\text{cube}(\text{square}(\text{cube}(\text{square}(\text{cube}(\cos(\text{square}(\text{Teff} + \text{logg})))))))))) - 0.0990011)) / (\sin([M/H] - 0.2566668))$
Fe I	869.87	Profundidad	0
Fe I	869.87	Desviación	$0.4798296 * \exp(-\text{Teff})$
Mn I	869.91	Profundidad	$\text{square}(0.0724468 - 0.2099630 * \sin(1.4649953 * \text{Teff} + [M/H] - 0.7439278))$
Mn I	869.91	Desviación	$-0.0097780 * \exp([M/H])$
Fe I	869.95	Profundidad	$\text{square}(\text{cube}(\text{logg} / \text{square}(\text{Teff})))$
Fe I	869.95	Desviación	0

Índice de siglas y acrónimos

A

ADQL: Astronomical Data Query Language

ANITA: Astronomía no supervisada e interpretable con técnicas avanzadas

API: Application Programming Interface

F

FLAME: Final Luminosity Age Mass Estimator

G

GSP-Phot: General Stellar Parametrised-Photometry

GSP-Spec: General Stellar Parametrised-spectroscopy

M

MARCS: Model Atmospheres in Radiative and Convective Scheme

P

PySR: Python Symbolic Regression

R

RMSE: Root Mean Square Error

RVS: Radial Velocity Spectrometer

V

VALD: Vienna Atomic Line Database