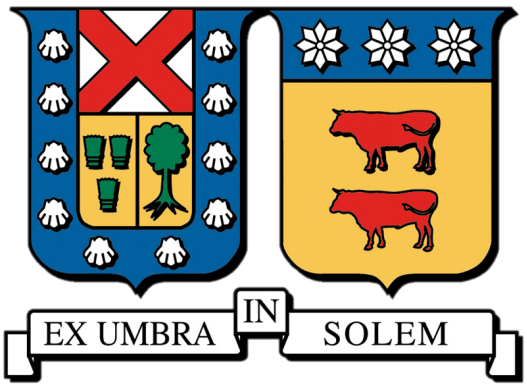
UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE FÍSICA



Testing mass-radius relations using *Gaia* distances as constraints in Machine Learning algorithms

Valentina Antonella Ortúzar Garzón

MEMORIA PARA OPTAR AL TÍTULO DE LA LICENCIATURA EN ASTROFÍSICA

PROFESORA GUIA: Odette Toloza

PROFESORA CO-GUIA: Pía Amigo

December 23, 2024

Abstract

White dwarfs (WDs) are the endpoint of evolution for over 90% of the stars in our Galaxy. They serve as essential instruments for studying various phenomena. For example, $25 - 50\%$ are **metal polluted**, however the timescales that elements heavier than He remain in their surface are extremely short (i.e., they are **self-cleaning**). The contamination can be explained by recent -or current- accretion of planetary remnants onto the surface of the star, which means we can study **the fate of planetary systems** by analysing the chemical abundances of WDs. Not only this, but because they no longer produce energy through nuclear fusion, they cool gradually in well known sequences, **which means we can determine their age, and with this, the age of the Milky way**. Both require an accurate determination of their main atmospheric parameters: effective temperature ($T_{\rm eff}$) and surface gravity ($\log(g)$). Mass-Radius (M-R) relations are necessary to determine the evolution and cooling of WDs, which in turn makes them necessary when trying to determine the atmospheric parameters mentioned above; they are determined by the degenerate nature of the star. The two main M-R relations are Montreal and La Plata. There have been efforts in constraining these M-R relations observationally, but the conclusions are not straightforward, because of the dispersion of the objects (they are small, therefore faint and difficult to observe) and the dependence of the M-R relations in multiple parameters.

Fortunately, ongoing and upcoming surveys are providing more observations of WDs, which should help with increasing the size of the sample. The problem now is how we analyse the data. Current techniques to fit a model atmosphere to a WD are time consuming and very computationally expensive, because they require numerically solving the differential equations of the stellar structure for each parameter combination, until we find the best fit by minimizing a metric like χ^2.

To confront this problem, we implement Machine Learning models to predict the atmospheric parameters of 262 hydrogen WDs, using ultraviolet spectroscopy from the Hubble Space Telescope. We use two decision tree-based ensemble algorithms: Random Forest (RF) and Extreme Gradient Boosting (XGBoost). We train the algorithms with synthetic atmosphere models. In order to compare the M-R relations mentioned above, we take two precautions: (1) In order to have model-independent evaluation metrics, we introduce the distance as a third variable to predict, since we

can compare this result to the *Gaia* distances, (2) We create two training sets, with synthetic fluxes created with the same $T_{\rm eff}$, $\log(g)$, and distance, but that represent WDs with radii according to each M-R relation.

We find that the best performing model on observational data is XGBoost for both M-R relations. With the model we find a distance Mean Absolute Error of $\sim$ 12.5 pc on average for the synthetic models and of $\sim$ 19.5 pc for the observational data. We find a mean difference ($\pm$ st. deviation) between Montreal and La Plata of $\sim 145 \pm 738$ K for $T_{\rm eff}$, $\sim -0.019 \pm 0.21$ [dex] for $\log(g)$, and $\sim -0.46 \pm 15.72$ pc for the distance. We also confirm that there is a problematic zone in the M-R relations, a fact that was already proposed in the literature, since our highest distance errors and the worst fits to the observational spectra are located at $T_{\rm eff} \lesssim 25000$ K and $\log(g) \lesssim 8.0$ [dex].

We compare our resulting parameters with previous works that use the traditional fitting techniques, and while we find that our errors are significantly higher, we obtain an accurate mass distribution with a mean mass of $0.60 \pm 0.18 M_{\odot}$ for Montreal and $0.62 \pm 0.22 M_{\odot}$ for La Plata.

We find that our results are affected by observational bias, and further analyses should be conducted on volume limited samples, with objects with distances < 50 pc. Given the error we obtain with our ML model, we believe it is still preferable to fit spectroscopic data with the conventional techniques in samples < 1000 objects. However, our algorithm still has room for improvement, and future work involves analysing the effects of feature sampling (that is, the flux points that serve as input), size of training set (to improve the resolution of the parameter space), and upgrade the computational capabilities, to test if it is possible to improve the performance of the model and obtain lower errors that are comparable to the results of using the current available techniques.

Resumen

Las enanas blancas (WD, por sus siglas en inglés) son el punto final de la evolución de más del 90% de las estrellas de nuestra galaxia. Sirven como instrumentos esenciales para estudiar diversos fenómenos. Por ejemplo, $25-50\%$ están **contaminadas por metales**, sin embargo el tiempo que los elementos más pesados que He permanecen en su superficie es extremadamente corto (es decir, son **autolimpiantes**). La contaminación puede explicarse por la reciente -o actual- acreción de restos planetarios en la superficie de la estrella, lo que significa que podemos estudiar **el destino de los sistemas planetarios** analizando las abundancias químicas de WD. No sólo esto, sino que debido a que ya no producen energía a través de la fusión nuclear, se enfrían gradualmente a tasas bien conocidas, **lo que significa que podemos determinar su edad y la edad de la Vía Láctea**. Ambos requieren una determinación precisa de sus principales parámetros atmosféricos: temperatura efectiva ($T_{\rm eff}$) y gravedad superficial ($\log(g)$).

Las relaciones Masa-Radio (M-R) son necesarias para determinar la evolución y enfriamiento de las WD, lo que a su vez las hace necesarias al intentar determinar los parámetros atmosféricos mencionados anteriormente; están determinados por la naturaleza degenerada de la estrella. Las dos principales relaciones M-R son Montreal y La Plata. Se han realizado esfuerzos para limitar estas relaciones M-R observacionalmente, pero las conclusiones no son sencillas debido a la dispersión de los objetos (son pequeños, por lo tanto débiles y difíciles de observar) y la dependencia de las relaciones M-R en múltiples parámetros.

Afortunadamente, los surveys en curso y próximos están proporcionando más observaciones de WDs, lo que debería ayudar a aumentar el tamaño de la muestra. El problema ahora es cómo analizamos los datos. Las técnicas actuales para ajustar una atmósfera modelo a una WD consumen mucho tiempo y son muy costosas desde el punto de vista computacional, porque requieren resolver numéricamente las ecuaciones diferenciales de la estructura estelar para cada combinación de parámetros, hasta encontrar el mejor ajuste al modelo minimizando una métrica como χ^2.

Para afrontar este problema, implementamos modelos de Machine Learning para predecir los parámetros atmosféricos de 262 WD de hidrógeno, utilizando espectroscopía ultravioleta del Telescopio Espacial Hubble. Utilizamos dos algoritmos de conjunto basados en árboles de decisión: Random Forest (RF) y Extreme Gradient Boosting (XGBoost). Entrenamos los algoritmos con

modelos sintéticos de atmósferas. Para comparar las relaciones M-R mencionadas anteriormente, tomamos dos precauciones: (1) Para tener métricas de evaluación independientes del modelo, introducimos la distancia como una tercera variable a predecir, ya que podemos comparar este resultado con las distancias de *Gaia*, (2) Creamos dos conjuntos de entrenamiento, con flujos sintéticos creados con los mismos $T_{\rm eff}$, $\log(g)$y distancia, pero que representan WD con radios de acuerdo con cada relación M-R.

Encontramos que el modelo con mejor rendimiento en datos de observación es XGBoost para ambas relaciones M-R. Con este modelo encontramos un error absoluto medio de distancia de ~ 12.5 pc en promedio para los modelos sintéticos y de ~ 19.5 pc para los datos observacionales. Encontramos una diferencia media ($\pm$ desviación estándar) entre Montreal y La Plata de $\sim$145$\pm$738 K para $T_{\rm eff}$, $\sim -0.019 \pm 0.21$ [dex] para $\log(g)$y $\sim -\ 0.46 \pm 15.72$ pc para la distancia. También confirmamos que existe una zona problemática en las relaciones M-R, hecho que ya a propuesto en la literatura, ya que nuestros errores de distancia más altos y los peores ajustes a los espectros observacionales se ubican en $T_{\rm eff} \lesssim 25000$ K y $\log(g) \lesssim 8.0$ [dex].

Comparamos nuestros parámetros resultantes con trabajos anteriores que utilizan las técnicas de ajuste tradicionales, y aunque encontramos que nuestros errores son significativamente mayores, obtenemos una distribución de masa adecuada con una masa media de $0.60 \pm 0.18 M_{\odot}$ para Montreal y $0.62 \pm 0.22 M_{\odot}$ para La Plata.

Descubrimos que nuestros resultados se ven afectados por un sesgo de observación, y se deben realizar análisis adicionales en muestras de volumen limitado, con objetos con distancias < 50 pc. Dado el error que obtenemos con nuestro modelo de ML, creemos que sigue siendo preferible ajustar los datos espectroscópicos con las técnicas convencionales en muestras de < 1000 objetos. Sin embargo, nuestro algoritmo todavía tiene margen de mejora y el trabajo futuro implica analizar los efectos del muestreo de características (es decir, los puntos de flujo que sirven como entrada para el entrenamiento), el tamaño del conjunto de entrenamiento (para mejorar la resolución del espacio de parámetros) y la mejora de las capacidades computacionales, para determinar si es posible mejorar el rendimiento del modelo y obtener errores más bajos que sean comparables a los resultados de las técnicas disponibles actualmente.

Agradecimientos

Gracias a mis amigos más cercanos, por caminar conmigo -y mantenerme cuerda- durante esta travesía de cuatro años. Espero que nuestros caminos se crucen muchas veces más en el futuro. A mis profesores y profesoras, especialmente con los y las que tuve el privilegio de trabajar este año: gracias por creer en mi y ayudarme a alcanzar mi máximo potencial. A mi familia, por su apoyo incondicional e impulsarme a alinear mis acciones con mis sueños y deseos, y no con mis miedos. Este logro es gracias a ustedes.

Contents

Chapter 1

Introduction

A white dwarf is, in brief terms, a dead star. It is the final remnant for low mass stars like the Sun, and it is formed when the core of the initial star is exposed after the expulsion of the outer layers. They correspond to the final outcome for over 90% of the stars in our Galaxy. They are degenerate objects, meaning they are supported by electron degeneracy pressure. Because they are degenerate, they have extremely high densities ($\sim 10^4 - 10^7$ times Earth's gravity), and since they correspond to the cores of stars, they have temperatures up to 10^5 K.

Because of their high gravities, we expect elements heavier than helium to sink to the core of the white dwarf in extremely short timescales. This means that if we observe a white dwarf, we should only see features that are produced by the transition of electrons between orbitals in hydrogen or helium, nothing else. We see however, that $\sim 25 - 50\%$ of white dwarfs are **metal polluted**, that is, they have other elements (Ca, Si, etc) in their surface (Zuckerman et al., 2003; Koester et al., 2014). Since the timescales during which these elements should remain visible in the atmosphere are extremely short, we can hypothesize that they are caused by recent -or even current- accretion of planetary materials onto the surface of the white dwarf. This means we can understand **the fate of planetary systems** by studying these white dwarfs (e.g. Zuckerman et al., 2007; Vanderburg et al., 2015; Xu et al., 2014, 2016; Farihi et al., 2022; Vanderbosch et al., 2021, and references therein). Not only this, but since they are dead, and do not produce energy through nuclear fusion, they cool gradually. If we were able to obtain accurate cooling sequences, we could determine the age of the white dwarf based on its atmospheric parameters. This would turn these stars into **cosmic clocks, with which we could even determine the age of the Milky Way**.

In order to study white dwarfs, we use **spectroscopy**, which is a technique where we observe the flux (radiation) of the star at each wavelength. This allows us to build a spectrum, which is a *fingerprint* of the star, where we can identify each element from their effects in the radiation. Metals will usually produce absorption features, since they absorb light as the electrons jump

between levels. In order to analyse these metals, not only do we need to model the absorption feature, but we also need to model the *standard* white dwarf atmosphere.

As expected, these atmosphere models are extremely complex, since we need to account for the full structure of the star, because radiation travels outwards from the core: this means we need to consider the chemical composition, temperatures, and many other parameters. We also need to scale the model flux by the surface area of the star: this means we also need to find a way to relate the mass of the star (which somewhat determines its composition, since it depends on the previous evolution) and the radius (which is necessary to measure the surface area of the star). Because these objects are degenerate, we know that the mass-radius relation is of the form $M \sim R^{-3}$ (Chandrasekhar, 1931).

Obtaining an accurate, empirical mass-radius relation is maybe even more complex than modelling the atmosphere of the white dwarf. In this case we need to consider the core composition of the star to obtain an equation of state, the mean molecular weight of the particles, the Coulomb interaction, relativistic cases, and much more. The shape of this relation changes drastically if we vary any of its defining parameters. There have been attempts at constraining the mass-radius relation observationally, for example, using eclipsing binaries (Parsons et al., 2017). In their work, Parsons et al. 2017 use a model-independent method to measure the radius and mass of stars in binary systems, and they find good agreement with models that favour thick outermost hydrogen layers, and are able to separate different core compositions from the agreement of the mass-radius relation of the different objects. While they have very high precision measurements (they average errors of 2.4% in mass, 2.7% in radius), their data is very scattered across the parameter space (temperature vs. gravity), so the assessment of the agreement with the different mass-radius relations is not straightforward (see their Fig. 9) because of how scattered the objects are.

It is clear that in order to constraint the mass-radius relation we need to analyse spectroscopic observations of white dwarfs. Up to today, observing white dwarfs has been a difficult task, because they are tiny objects (their radii are similar to Earth!) and they are not bright in most wavelength ranges: because of their high temperature their maximum emission is usually in the ultraviolet. And thus, missions in the ultraviolet as the *Hubble Space Telescope (HST)* are very important to study white dwarfs. Fortunately, there are many upcoming large scale surveys, for example the Ultraviolet Explorer (UVEX; Kulkarni et al., 2021), which will perform photometric observations of the full sky to fill in the ultraviolet range in the family of new wide-field surveys like the Vera Rubin (optical), Euclid (near-infrared) and Roman (infrared) observatories. There have been proposals of spectroscopic surveys, like the Ultraviolet Spectroscopic Characterization of Planets and their Environments (UV-SCOPE) mission, which also plans to observe multiple objects in the ultraviolet range (Ardila et al., 2022).

While all these surveys will provide us with an unimaginable amount of new knowledge, they

also carry an absurd amount of data we need to process and analyse in an efficient way. Current techniques to model the spectral energy distributions, not only of white dwarfs, but of many objects, are tedious and require high computational resources. Because of this, we have seen an uprise in the use of machine learning algorithms in astronomy. We have seen machine learning applied to the determination of members of globular clusters (Pulgar-Escobar et al., 2024), and in ALeRCE, the Automatic Learning for the Rapid Classification of Events (Förster et al., 2021), which is extremely useful in identifying transient processes like supernova type Ia explosions, which have huge implications for stellar evolution and even in determining the expansion rate of the Universe. These applications have focused on finding new techniques and tools that help optimize processes that are otherwise very time consuming. For white dwarfs specifically, Badenas-Agusti et al. (2024) developed `cecilia`, which is a neural network-based interpolator that models synthetic spectra for helium white dwarfs, based on their temperature, gravity, and metal abundances; however, `cecilia` still requires an iterative process to fit these spectra to an object, and a strong preprocessing of the data through normalization of the spectra beforehand, which means we still need to do manual work before. For hydrogen white dwarfs -which are the objects studied in this work- Yang et al. (2020) used a combination of neural networks to determine the temperature and gravity of the stars, based on the optical spectra, obtained from the white dwarf catalogue of the Sloan Digital Sky Survey (SDSS). Similarly to `cecilia`, this work requires normalizing and or smoothing the spectra, with which much of the information of the radius of the star, the distance, and the absorption features is lost. Another common problem when implementing ML is the insufficient number of representative samples of the different classes, which makes it difficult to train the algorithms using observational data, which is why we require synthetic models to overcome this problem.

In this work we try to confront the issue of measuring the atmospheric parameters of hydrogen white dwarfs using machine learning algorithms. We start with algorithms less complex than neural networks, based on decision trees, to assess the effect of using different mass-radius relations in the data we use to train each model. Apart from comparing our results to previous work (that determine these parameters in the traditional way), we also use an model-independent constraint: the *Gaia* DR3 distances. *Gaia* is a space telescope that obtained extremely accurate parallax measurements of objects across the sky. By using these measurements we can identify where each mass-radius relation behaves appropriately, by analysing the correlation of the distance prediction with the location of the object in the $T_{\rm eff}$ and $\log(g)$ parameter space.

In Section 1.1 we include an overview of the formation and evolution of white dwarfs. Modelling of their spectral energy distribution is in Section 1.2. Then, in Section 1.3 we introduce the most relevant machine learning concepts used in this work. Finally, we include the process of diagnosing and evaluating the performance of these algorithms in Section 1.4.

1.1 Formation of White Dwarfs

Stars are formed when a large molecular cloud, composed mainly of neutral hydrogen, collapses under its own gravity. The cloud compresses, fragments, and eventually we start seeing the first stellar bodies emit light. The standard way to represent this evolution is in a Hertzprung-Russell (H-R) diagram, which tells us how the temperature and luminosity of the star vary during its life, as shown in Fig. 1.1. At first, this radiation corresponds only to gravitational energy, which is emitted as the star continues to compress during the Pre-Main Sequence phase, but once the required temperature and density are reached at the core, nuclear reactions begin, and the star starts producing energy by the fusion of H into He. This is the Main Sequence (MS) phase, where the star will spend most of its life. In Fig. 1.1 the color represents the age, and we can see that the star stays in the MS for $\sim 10^{10}$ years. Once the star exhausts this initial reservoir of hydrogen, it leaves the MS, and depending on: its mass, it will evolve following different tracks. However, most stars have low mass, and have similar evolutions. Specifically, for masses $\lesssim 8$ $M_\odot$[1] (that is, $\simeq 97\%$ of the stars in the galaxy), the star moves into a Red Giant Branch (see in upper right corner of Fig. 1.1), where it expands its outer layers, and therefore appears cooler, hence as a *redder* colour. On the other hand, the core is still under the effects of the gravitational potential and keeps heating and increasing in density, until it reaches the conditions to start fusing He into C and O, in a new phase of nuclear fusion reactions. Once again, the star will eventually deplete this second fuel and will find itself in the Asymptotic Giant Branch (see in upper right corner of Fig. 1.1). At the end of this phase, the star is not able to maintain equilibrium and will violently expel its outer layers, exposing the core. This core is now a degenerate body, supported by the electron degeneracy pressure, which means that it has an extremely high density. Because of this, its components are stratified; that is, we find that the heavier metals, like C and O, are found in the centre in a degenerate nucleus, surrounded by a He shell, and H in the outermost layer. This is what we call a White Dwarf (WD).

The core of most WDs is composed of C and O, and it constitutes $\sim 99\%$ of its mass. The remaining 1% corresponds to H and He, located in a thin envelope, in ratios that depend on the nuclear burning history of the star. Depending on which element is more abundant, we can have H- or He-rich WDs. The only visible part of the star corresponds to the atmosphere, that can be H- or He-dominated. Depending on the features we see in their spectroscopy, we can classify them into different categories. The most common spectral type is DA (see Fig. 1.2), which corresponds to WDs that only show hydrogen features from the Balmer series in the optical wavelength range ($\sim$4000 to 7000Å). The second most common type are DB WDs, that show He features; these are usually H-deficient, which makes observing He features possible, since He is less opaque than

[1] In astronomy, it is convenient to express some quantities in terms of the Sun, like the mass $M_\odot = 1.998 \times 10^{30}$ kg, radius $R_\odot = 6.957 \times 10^8$ m and luminosity $L_\odot = 3.828 \times 10^{26}$W.

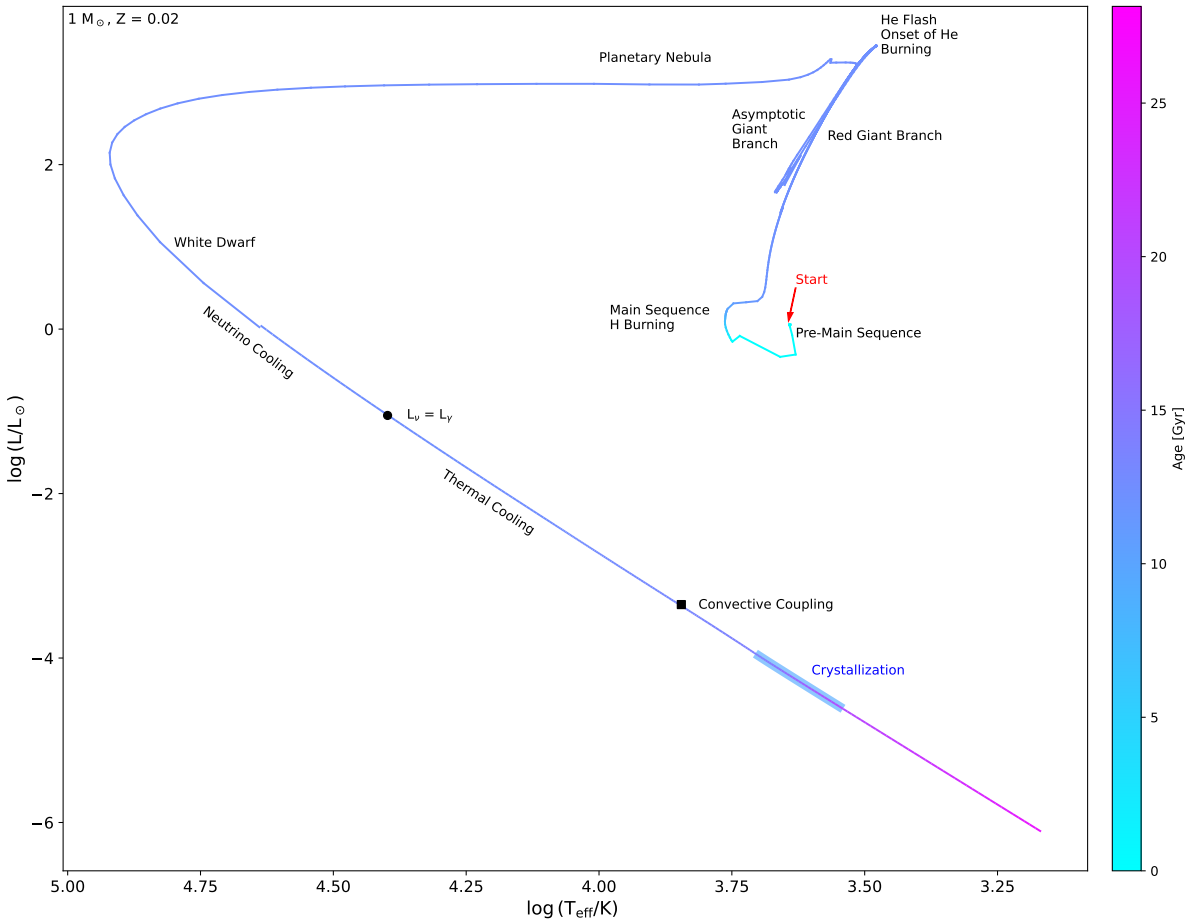


Figure 1.1: Evolution of a 1 $M_\odot$ star, with metallicity $Z = 0.02$, computed with MESA (r22.05.1). The star begins its evolution at the Pre-Main Sequence (red arrow). The final cooling was extracted from the Bédard et al. (2020) evolutionary sequences for a 0.5 $M_\odot$WD. The colorbar indicates the age of the star throughout its evolution. The transition between neutrino and thermal cooling is indicated with a black circle, the onset of convective coupling is shown as a black square, and the crystallization of the core is marked in a thick blue band.

H, which means the outer layers are more transparent to radiation. It is important to note that these types are not necessarily correlated with the composition of the WD, and for example, since He has a lower opacity than H, we can have a WD with a He atmosphere, a DB, but where we only observe H features, like a DA; these WDs are called DABs or DBAs. There are many other categories, depending on the presence of ionized He (DO), neutral, ionized or molecular C features (DQ), the lack of features (DC) or even the presence of metals (DZ), but most of the population corresponds to DA or DB spectral types (see Figure 1.3).

Depending on the WD mass, we could have an He-core when $M_{\rm WD} < 0.3$ $M_\odot$, a hybrid core, with carbon, oxygen and helium, when 0.3 $M_\odot < M_{\rm WD} \lesssim 0.45$ $M_\odot$. For higher masses, between 0.45 $M_\odot \lesssim M_{\rm WD} \lesssim 1.05$ $M_\odot$, we get a C-O core. Finally, for masses higher than 1.05 $M_\odot$we

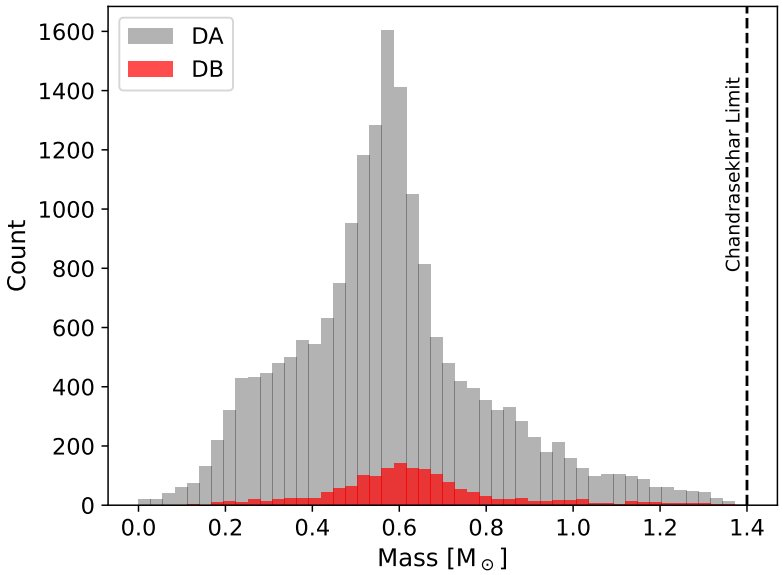


Figure 1.2: Mass distribution of DA (gray) and DB (red) WDs from the Montreal White Dwarf Database. Here we see that most WDs correspond to spectral type DA, and have masses much lower than the Chandrasekhar mass.

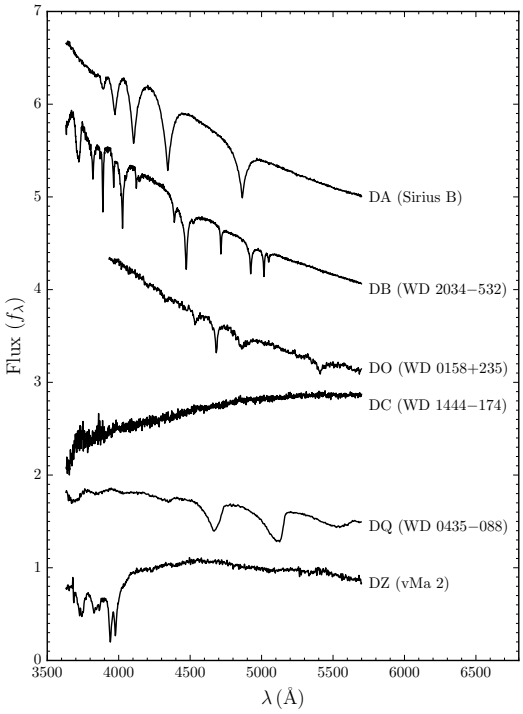


Figure 1.3: Figure taken from Saumon et al. (2022) which displays examples of spectra of DA, DB, DO, DC, DQ, and DZ white dwarfs. Spectra are normalised at 5000 Å and vertically shifted for clarity.

expect an O-Ne-Mg core, produced by the highest mass stars that evolve into WDs. The WD mass distribution for the most common type of WDs, i.e. the types of DA and DB, are shown in Figure 1.2. Early in the 20th century, Chandrasekhar (1931) showed that there exists a maximum mass a non-rotating WD can reach without becoming unstable due to electron degeneracy pressure (Figure 1.2). Thus, solving the equation of state for a ideal Fermi gas, there exist a relation of the WD mass and its radius.

1.1.1 Mass-Radius Relations and White Dwarf Cooling Models

Average mass and radius of a WD are 0.60 $M_\odot$ and 0.013 $R_\odot$ (see Fig. 1.2), which means they have densities of $\sim 10^6$g cm^{-3}, that is almost a million times the density of the Earth, and gravities of $g \sim 10^6 - 10^9$g cm s^{-1} (the Earth's gravity is $\sim 10^2$ g cm s^{-1}). With these conditions, it is clear that all elements inside the WD are ionized and that all matter should be degenerate.

WDs are extremely hot at the beginning of their life, with effective temperatures between 100000 and 150000 K, but since they are no longer producing energy through nuclear reactions, they cool gradually in timescales that can be longer than the age of the universe. The coldest WDs we can observe today have temperatures of $\sim 3000 - 4000$ K. This process can be seen in Fig. 1.1, where the WD cools for most of its life, first by the emission of neutrinos from the core, with a net energy loss, since these do not interact as they move to the stellar exterior, until it reaches a surface temperature of $\sim$25000 K, when it transitions into thermal cooling. In this phase the WD cools by the slow leakage of thermal energy from the ions in the plasma. As the plasma cools, it starts to crystallize. Higher mass WDs crystallize at higher effective temperatures, since they have higher densities. As temperature decreases, the degenerate front moves upwards, while the outer convective layers penetrate deeper. At a $T_{\rm eff}\sim$6000 [K] the convective layers reach the degenerate center, and the WD achieves convective coupling, which allows heat to escape rapidly, and therefore allows an increased cooling rate. The layers only penetrate up to a $M_{\rm H}/M_\star \sim 10^{-6}$ mass fraction. It is important to note that for lower mass WDs, like the one in Fig. 1.1, these phases are inverted, the star experiences convective coupling before crystallization.

Since WDs cool slowly, and following similar evolutionary tracks, they are excellent clocks, because we can estimate their age with precise measurements of $T_{\rm eff}$, and $\log(g)$. These cooling models are built on the basis of mass-radius (M-R) relations. M-R relations are extremely complex, since they depend on multiple parameters, for example Ambrosino (2020) showed the variation of the M-R for different Coulomb interaction values (see their Fig. 15). Carvalho et al. (2016) also shows the effects of studying these stars using general relativity to obtain the M-R, with which they find different results and instability zones depending on if we consider a Newtonian or ultra-relativistic case. Both of these show that the relation varies strongly when multiple

parameters change, which proves its complexity and the need to constrain M-R relations with model-independent methods.

The main M-R relations correspond to the Montreal[2] M-R relation (hereafter referred to as Montreal, M), based on the evolutionary sequences of Bédard et al. (2020), and the La Plata[3] M-R relation (Althaus et al., 2013; Camisassa et al., 2016; Camisassa et al., 2019) (hereafter refered to as La Plata, LP), based on the models of Althaus, L. G. et al. (2005). With these M-R relations we can interpolate values for the radius of the star based only on its $T_{\rm eff}$, and $\log(g)$, and therefore correct the model fluxes to match the observations.

There are a few key differences between Montreal and La Plata cooling models. Montreal models are *relaxed* from an artificial structure at high $T_{\rm eff}$, which means they are built on a pre-WD thermodynamic structure obtained with Modules for Experiments in Stellar Astrophysics[4](MESA) (a software that computes the evolution of a star by numerically solving the differential equations that determine the stellar structure, Paxton et al., 2011, 2013, 2015, 2018, 2019; Jermyn et al., 2023) and an artificial chemical profile that considers C-O in equal quantities, with a thick envelope with a helium mass ratio qHe $= 10^{-2}$ and a hydrogen mass ratio of qH $= 10^{-4}$, and a thin layer with qHe $= 10^{-2}$ and qH $= 10^{-10}$. The WD masses range from 0.3 to 1.3 $M_{\odot}$, while temperatures range from $\log(T_{\rm eff}) \sim 3.18$ to ~ 5.18, and $\log(g)$ between 7.0 and 9.0.

On the other hand, for progenitors with initial masses between 0.9 to 6.7 $M_{\odot}$, meaning WDs with masses between 0.49 to 1.42 $M_{\odot}$, La Plata models compute the full evolution from MS to WD cooling. Since the evolution is computed for each star, the chemical composition (that is, the values of qH and qHe) is determined by the previous evolution. Since La Plata compute the full evolution and do not fix the hydrogen or helium ratio, they find that thickness increases with mass, which in turn affects the radius of the star (Romero, 2022). For initial masses between 9 and 10.5 $M_{\odot}$, we get WDs with masses between 1.1 and 1.29$M_{\odot}$, which correspond to O-Ne cores. In this case, the La Plata models use a similar approach to Montreal, where they select a chemical profile from the AGB phase, and relax a high $T_{\rm eff}$ model into a WD. Here they consider a fixed qH $= 10^{-6}$, while qHe is determined by the previous evolution. These models have temperatures in the range $\log T_{\rm eff} \sim 3.4$ to ~ 4.9, and $\log(g)$ between 7.0 and 9.5. As it could be inferred, these cooling and M-R models are an important input when scaling atmospheric WD models (see below) to reproduce the spectral energy distribution of these stars.

[2] http://www.astro.umontreal.ca/ bergeron/CoolingModels
[3] https://evolgroup.fcaglp.unlp.edu.ar/TRACKS/tracks.html
[4] https://docs.mesastar.org/en/24.08.1/index.html

1.2 Modelling the Spectral Energy Distribution of White Dwarfs

Since WDs are hot objects, most of their emission is located in the UV range. In this section of the spectrum we can see hydrogen absorption features, including Lyman-α absorption. These line profiles are extremely sensitive to the values of the atmospheric parameters, effective temperature ($T_{\rm eff}$) and surface gravity ($\log(g)$), which means we should be able to estimate their values performing UV spectroscopy on DA WDs. To obtain the parameters, we need to compare the received flux with a synthetic atmospheric model.

To obtain a synthetic atmosphere, we should be able to resolve a small section of the star, to obtain the emission from a differential area, and then scale this quantity by the solid angle, which depends on radius of the star and the distance at which the star is located. The models are obtained by solving the radiative transfer equation, by making several assumptions:

- The depth of the atmosphere is much smaller than the radius of the star. All variables (density, temperature, etc.) depend only on the height. The intensity only depends on this height, and the observation angle, but is symmetric along the azimuthal axis.
- The atmosphere is in hydrostatic equilibrium: the gas pressure gradient is in equilibrium with the stars gravity.
- The atmosphere only transports the energy from the lower layers, there is no energy creation or loss. This transport can be through radiation, convection or conduction.
- The matter is in local thermodynamic equilibrium, meaning each layer is in thermal equilibrium with its corresponding temperature. This also helps us constraint the ionization, excitation, and other processes concerning the behavior of the particles, since these processes are well understood under thermal equilibrium. This assumption is well justified for WDs with temperatures lower than 50000 K.

While these assumptions help us reduce significantly the complexity of these models (and therefore the computational cost of building them) it is necessary in some cases to add non-ideal effects, that account for stochastic processes that occur in the surfaces of these stars.

The modelling of the SED of white dwarfs allows us to determine the chemical abundances seen in $\sim 25-50\%$ of WDs, and thus it is an important technique to study the core composition of planetary remnants around WDs (e.g. Zuckerman et al., 2007; Vanderburg et al., 2015; Xu et al., 2014, 2016; Farihi et al., 2022; Vanderbosch et al., 2021, and references therein).

1.2.1 Observations of White Dwarfs

There are two primary techniques to determine the shape of the spectral energy distribution (SED) of WDs: photometry and spectroscopy. Photometry consists in capturing images of the objects, to measure their brightness. In order to do this, we use Charge Coupled Devices (CCDs), that are integrated circuits widely used in digital imaging; here, we detect and *capture* the radiation from a star, as the amount of photons that reach each pixel in the CCD. Each photon interaction frees an electron inside a detector, that is then counted: the brightness of the object will be measured by converting those counts into analogue digital units, and then into apparent magnitude (usually measured in *maggies*). It is important to note that brightness and magnitude are not numerically correlated: a brighter object will have a smaller value of magnitude (for example, the Sun has a magnitude of ~ -26.7, while the brightest star in the night sky only has a magnitude of 1.45).There is a well know relation for the apparent magnitude of an object:

$$m = M + 5\log_{10}(d[\text{pc}]) - 5 \tag{1.1}$$

Here m is the apparent magnitude, that is, the brightness we observe, M is the absolute magnitude, defined as the brightness the object would have if it were at a standard distance $d = 10$ pc. This is the distance modulus, and is extremely useful to compare objects that are not located at the same distances: the apparent magnitude of an object increases as the distance grows, that is, the object gets fainter and fainter. White dwarfs are extremely faint, their absolute brightnesses range from 10-11 magnitudes, which makes them hard to observe. To approximate an SED, we can obtain the magnitudes with different filters, that is, at different wavelength ranges.

On the other hand, spectroscopy, while also using a CCD, employs it in a different manner. In this case, we also use a light refractor, that allows us to separate a ray of light into all its components. The detector receives this separated light along one axis, and each pixel will have its corresponding wavelength. Depending on the quality of the instrument and the refractor, we can have different resolutions and wavelength ranges. The benefit of using spectroscopy, is that it allows us to identify not only how bright the object is (we can also measure magnitude from a spectrum), but it also tells us the composition of the star, through the presence of absorption or emission features. In a WD we primarily look for hydrogen or helium features, to determine what element dominates the stars atmosphere, and any other chemical contamination in the form of absorption. To analyse the composition and properties of a star, we need spectroscopy.

Large-scale spectroscopic surveys, both past and ongoing (e.g., LAMOST, WEAVE, SDSS-V, DESI, and 4MOST), have provided and will continue to provide optical spectroscopy for thousands of WDs.

In Fig. 1.3 we include an example spectra for the different classes in the optical range. For the

DA WD we mainly see the Balmer series, that is, the broad hydrogen features that form when an electron transitions from $n = 2$ to higher levels.

However, a limitation of optical spectroscopy is the flux calibration errors introduced by Earth's atmosphere, an issue that can be significantly mitigated by space-based missions. Thus, optical spectroscopy has proven highly effective for selecting DA white dwarfs to be observed by the *HST*.

In the UV range covered by *HST*, the Lyman series are captured, especially the Lyman$-\alpha$ line (i.e. the transition of an electron from the $n = 1$ orbital to $n = 2$).

Because of the nature of these stars, if we observe them in the ultraviolet, we need to account for the extinction produced by the Interstellar Medium (ISM), since this range is the most affected by the absorption due to the presence of dust. A way to correct these effects is to revert the extinction using the 3D maps from Stilism[5] (Lallement et al., 2014; Capitanio et al., 2017; Lallement, R. et al., 2018) and the wavelength dependent extinction law from Fitzpatrick and Massa (1990); Fitzpatrick (1999).

Up to this point, we have discussed the observations of white dwarfs and the existing theoretical models. However, various methods are available to optimize the search for the best-fitting model that reproduces the observations. These include techniques such as minimizing metrics (e.g., residuals or χ^2), Bayesian analysis (e.g., Markov Chain, Monte Carlo or Gaussian processes), and more recent self-learning approaches (e.g., machine learning, deep learning, and artificial intelligence). In this work, I will utilize machine learning to predict atmospheric parameters of observational data of DA WDs using synthetic models to train the algorithms.

1.3 Supervised Machine Learning: Regression

Nowadays, to model a spectroscopic observation of a star, we require a grid of synthetic models, that are built by solving a simplified version of the radiate transfer equation that dominates the structure of the star (Koester, 2010). These models depend on parameters such as temperature, mass, radius (and therefore, gravity), and the chemical composition of the star.

To *fit* a model to the observed spectrum, the most common technique is to minimize the *difference* between a synthetic spectrum generated from a set of parameters. This difference is defined using the χ^2 statistic defined as:

$$\chi^2 = \sum_i \frac{(F_{\rm obs,i} - F_{\rm syn,i})^2}{\sigma^2_{\rm obs,i}} \tag{1.2}$$

Where $F_{\rm obs}$ and $F_{\rm syn}$ correspond to the observed and synthetic flux, respectively, $\sigma_{\rm obs}$ corresponds

[5] https://stilism.obspm.fr/

to the observed flux error, with the sum spanning all the wavelength range. There is an alternative way to define this statistic, as a reduced χ^2:

$$\chi^2_r = \frac{1}{n} \sum_i \frac{(F_{\rm obs,i} - F_{\rm syn,i})^2}{\sigma^2_{\rm obs,i}} \tag{1.3}$$

Where n corresponds to the number of data points in the spectrum. This second definition ensures that the statistic is weighed by the length and resolution of the spectrum, so that we can accurately compare two observations.

By fitting synthetic models with this technique, we need to have a large grid of models, with multiple parameter combinations, which is computationally expensive, because we need to generate each of these syntetic models, and because we need to manually check each one, using χ^2_r, and then look for the best fitting model, that is, the one with $\chi^2_r \sim 1$. This has not been a huge issue so far, but considering the increasing inflow of data coming from future surveys, we need to start looking for methods that make this process faster and more effective. As a result, we have started to see an uprise in the use of Machine Learning (ML) techniques (see Badenas-Agusti et al. (2024); Yang et al. (2020)), since they offer an alternative that has proven to be much more efficient at recognizing patterns, which is exactly the behavior we expect to see in the spectrum of a star, and is actually what astronomers look for when fitting a model manually. Unfortunately, because of limited computing capacity, and the dimension of the spectrum, that is, the amount of data points, current ML models still need extensive preprocessing of the data, which entails cropping certain regions of the spectrum, or reducing the dimensionality by interpolating or averaging the flux values. This limits strongly what we do with the flux, since we believe there is a loss of information when we manipulate the observed data. Not only this, but preprocessing the data still requires extra work -and time- we think could be better used elsewhere.

A sub-category of ML corresponds to supervised learning, which are algorithms where the model learns from a **training set**, that contains the input or **features** that we want to use to characterize a variable, and **labels** or **targets**, which are the real values of the variable, which we are trying to predict. Labels are related to classification algorithms, where, as the name says, the algorithm simply learns to categorize **instances** (that is, each object) according to its features. Targets however correspond to continuous values, that are related to **regression problems**, where the algorithm learns to model the data and returns a value that optimizes the error between its predictions and the real data. Here we will only use regression models.

Once the model has been trained, we need to evaluate its performance. This is done using a **validation set** which should contain a sample that the model has not encountered in training. We predict the targets from the validation set using the model, and compare the results with the real

target values. There are several tests we can perform with both of these sets, to see how robust the model is, by for example, evaluating how well it performs when the training set changes, either by varying which models are in the set, or how large the set is. In any case, we expect the performance to remain constant, to prove that the model holds up independently of the data sample we use.

Finally, the model should be used to predict values in a **test set**. In this final set, we do not know the target values, which is why we need to ensure that the model is performing well before we introduce this new data. In our case, the training and validation sets are built with synthetic models, while the test set corresponds to observational data.

In order to quantify how good the predictions are for continuous variables (that is, in regression problems) in either the validation or test sets, we can use several **metrics**, such as the Mean Squared Error (MSE):

$$\text{MSE} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \tag{1.4}$$

Or the Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{1.5}$$

Where n is the number of points, y_i corresponds to the expected value, and $\hat{y}_i$ corresponds to the predicted value. We can also use the Mean Absolute Error (MAE):

$$\text{MAE} = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{1.6}$$

Another option is to use the R^2 score, that is defined as:

$$\text{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{1.7}$$

Where $\bar{y}_i$ corresponds to the mean value of the variable or feature. Both MAE and RMSE are superior in terms of interpretability, since they will have the same units as the original variable, and are therefore directly comparable with the original values. On the other hand, R^2 can be a good indicator of the performance of the model, but unfortunately if we have data with a high level of dispersion, the right side of the equation approximates 0, and $R^2 \simeq 1$, even if the model is not performing well. Usually we evaluate and compare the performances of the algorithms with more than one metric.

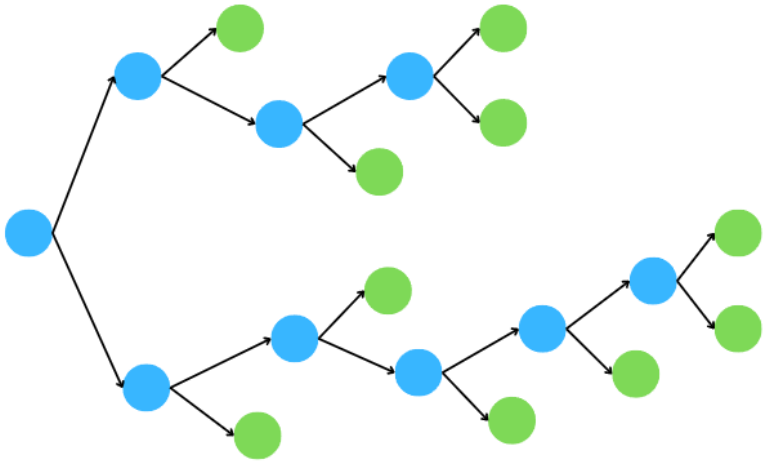


Figure 1.4: Simplified representation of a decision tree. The blue circles represent the nodes, where each decision is made. Black arrows correspond to the different outcomes, or branches. Finally, the green circles correspond to the leaves, the final result found by the algorithm.

1.3.1 Ensemble Methods

One type of machine learning algorithm that works with supervised learning are **decision trees**. As the name says, these models have a tree-like structure, with nodes that represent a test made on a certain feature, branches that represent the possible outcomes, and leaves that represent the final value of the regression parameters. We show a very simple representation of a decision tree in Fig. 1.4, where the blue circles represent nodes, where we are splitting the input data according to a condition imposed by the algorithm. The final regression values will be found in the leaves, in this case corresponding to the green circles. Although decision trees can perform well for certain problems, they tend to be extremely noisy, and to over-fit the data if there is no control over the *growth* of the tree (meaning, without the proper hyperparameters, described below).

To obtain a more robust algorithm, we can implement **ensemble methods**, which allow us to aggregate more than one estimator in a single model. In this case, we chose two decision tree based ensemble methods: **Random Forest**, based on **bootstrap aggregating or bagging** and **Extreme Gradient Boosting**. In the following Section, we will describe each of them, based on the `sklearn` documentation, since they were both implemented from `Python`.

1.3.1.1 Random Forest

Bootstrap AGGregating, or bagging, corresponds to a two part method, used to improve the performance of ensemble models. First, the bootstrapping generates randomized learning sets by re-sampling the data, and then an averaging or aggregating the prediction of the individual models

to obtain an ensemble prediction. This method assumes that each indiviual tree has a higher bias (meaning it models the training set well) and a high variance. By bagging we expect to reduce the variance of the model without increasing the bias too much, so that the total error is lower (Acquaviva, 2023).

If we use bagging with decision trees as the estimators, we obtain a Random Forest[6] (RF) algorithm. Now, we have mentioned that ensemble methods help reduce the variance in the model, but since the trees created with bootstrap sampling are independent, they should all have more or less the same data distribution, which means their individual bias is the same, and we can only improve by reducing the variance. RF algorithm tries to improve the variance reduction of bagging by randomly selecting the input for each tree, that is, selecting different samples of features to build the regression. Some of the hyperparameters that control the way an RF algorithm is built are:

- Number of Estimators (`n_estimators`): number of individual decision trees used to build the forest.
- Maximum number of features (`max_features`): maximum number of features that can be used during the training process.
- Maximum depth (`max_depth`): maximum amount of splits the tree can make; the deeper the tree, the more it will fit the data.
- Maximum number of leaf nodes (`max_leaf_node`): corresponds to the maximum number of leaves a tree can have: these correspond to the most *pure* class, so a tree with many leaves will also fit the training data very well.

1.3.1.2 Extreme Gradient Boosting

While Extreme Gradient Boosting[7] (XGBoost, XGB) is also an ensemble method based on decision trees, there is a fundamental difference with RF, since for the latter the trees are trained all at once, independently from each other, while for XGB the trees are trained sequentially, correcting the errors on each tree based on the performance of the previous one (Hastie et al., 2017; Friedman, 2001).

Gradient boosting works by adjusting the target function we are trying to predict, in this case, a representation of the residuals, that is, the difference between the prediction and the real value. Recalling that these models are trying to minimize an error function by fitting the residuals, this

[6] https://scikit-learn.org/1.5/modules/generated/sklearn.ensemble.RandomForestRegressor.html
[7] https://xgboost.readthedocs.io/en/latest/python/index.html

means that at each step it will correct the residuals and add them to the previous model, to improve the performance of the algorithm.

Some hyperparameters that control XGB models are:

- Number of Estimators (`n_estimators`): number of individual decision trees used to build the forest.
- Maximum depth (`max_depth`): maximum amount of splits the tree can make; the deeper the tree, the more it will fit the data.
- Shrinkage (`learning_rate`): As we have mentioned, gradient boosting implies we will add new corrections to the target function at each iteration; in order to limit the contribution of each of these, we include a shrinkage or learning rate, between 0 and 1, to regulate how fast or slow the model learns: the smaller the learning rate, the longer the model takes to achieve a certain accuracy.
- Subsampling (`subsample`): To achieve different levels of regularization we can subsample with or without replacement from the data. A subsample value of 0.5 will mean we will select 50% of the training instances.
- Column Sampling (`colsample_bytree`): Use a random number of columns, determined by a fraction of 0 to 1, to construct each tree. This parameter is similar to the behavior for bagging algorithms, to reduce variance.

1.4 Diagnostics and Evaluation

Once we have built and trained the models, it is necessary to analyse their performance on the validation set, before introducing the test data. To do this, we evaluate how different metrics change when we modify the training set and its size, to assess how the model depends on the data, to know if we are under- or over-fitting the training set. Apart from this, we can also interpret how much each feature *weighs* in the decision making.

1.4.1 Bias and Variance

There are two important properties that we constantly examine when using ML algorithms: bias and variance. The **bias** corresponds to the difference in the predicted and real values of the targets, and informs us if the model is under- or over-fitting the data. The variance corresponds to the amount that the performance of an algorithm changes when we change the training set;

in other words, we estimate how sensitive the algorithm is to the training data. There are a few combinations of bias and variance that indicate what is happening with the model:

- A high bias could indicate that the model cannot capture the trends of the data (under-fitting), and can not predict the values correctly. If we also have low variance, it means all predictions are similar, because the model is not complex enough.
- A high variance could mean that the model is over-fitting the training data, and is unable to generalize to the validation or test set. If the bias is low, we will find that some of the predictions made are correct.
- A high bias and high variance could mean that the model is under-fitting the data, and is not able to capture the underlying patterns, while also being sensitive to the training set.

The ideal case is to have low bias and low variance, where the model is complex enough to capture the patterns of the training data, without being sensitive to it or over-fitting; this is not realistic, and in most -if not all- cases, impossible. It is important to know that these combinations could very well have different interpretations depending on the properties of the data; there is an amount of variance we expect if the data is noisy by nature, which will affect the performance of the algorithm and will be noticeable when evaluating the bias and variance.

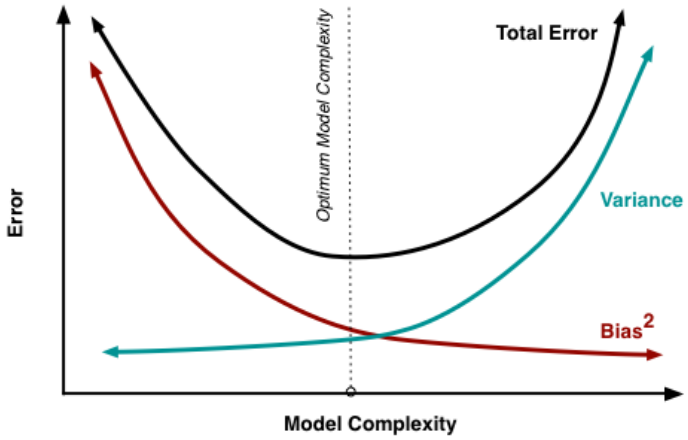


Figure 1.5: Representation of the bias-variance trade-off. We can see that as the model becomes more complex, bias (red line) decreases, but the variance (green line) increases, meaning that we will fit the data more accurately, but we will introduce more error, since the model will not be able to generalize itself to the validation set. We also note that there is a minimum for the total error (black line), where we find the optimum complexity from the model. Figure taken from Fortmann-Roe (2012)

The **bias-variance trade off** is a well known issue in ML algorithms. Fig. 1.5 illustrates this problem. What we mean when we say that there is a bias-variance trade off is that for a model,

in order to reduce bias we need to increase the complexity, which in turn increases variance, and viceversa, if we want less variance, the model has to be less complex, which makes the bias increase. There is fortunately, an optimum complexity, that can be found through cross-validation, hyperparameter optimization and evaluating the overall performance through metrics, for example, with the learning curve of the model, which we will describe in the following Sections.

1.4.1.1 Learning Curves and Cross-validation

Figure 1.6: Learning curve examples adapted from https://utkuufuk.com/2018/05/04/learning-curves/. The centre panel shows the ideal case, where the validation error (blue line) decreases as the number of instances (that is, the training set size) increases, with the testing error (red line) increases slightly and reaches a plateau. The left panel shows a case of over-fitting, where the bias is high, which means the error is high in both cases. The right panel shows a case of high variance, where the curves are very separated.

To assess the level of bias and variance in a model, we can implement two validation methods. First are the *learning curves* that evaluate how the performance of an algorithm changes as we increase the number of elements in the training set. The performance can be evaluated with RMSE, MSE, MAE, R^2 or other metrics, depending on if we are dealing with a regression or classification problem. These values are calculated for both the training and the validation set for each iteration of the size of the training set. The objective is to see how the models behave when the size of the training set size changes, and what we expect is to see similar (but not equal!) values of the metric at the largest sizes. The error for the validation set should decrease as the training set gets bigger, while the error for the training set should increase slightly, since we expect to reduce over-fitting by introducing a larger sample to train the model. Additionally, if the data is not noisy, we expect both curves to saturate and remain constant when we reach a certain training set size, as shown in the center panel of Fig. 1.6. Here, the left-hand panel shows a case of high bias, where the error of both the training (blue) and validation (red) sets is high, while the right-hand panel shows a case of high variance, where the training error is extremely low while the validation error is high, indicating that the model is most likely over-fitting the training data and is therefore unable to correctly predict the validation set targets.

Another test we can perform on ML algorithms is *cross-validation*, which consists of dividing the data into several subsets or *folds*, choosing one of them as the validation set, and training the algorithm with the remaining data. This is repeated until each fold has been used as the validation set. Then the scores for each iteration are averaged to obtain a better idea of the models performance. We expect the standard deviation of these scores to be small; however, if there is a lot of variation, it could mean that the model's performance depends too strongly on the choice of training set, indicating that it is over-fitting the training data.

1.4.2 Feature Importance

The great benefit of using decision tree-based algorithms is that they are highly interpretable. Because of this, we can access the *feature importance*, that is, how large was the effect each feature had in the decision making process. This is algorithm-dependent, and does not necessarily mean that the feature is important; for example, correlated features tend to split the importance, which means we might have lower scores for features that are being used constantly by the algorithm, but have similar *meanings*.

In decision trees, and ensemble methods based on them, more important features are usually at the top, in the first few splits, since they are most effective in reducing the residuals in regression problems. If we wanted the feature importance we would have to sum the decrease of error each time the feature was used in a split.

So far we have introduced the main properties of WDs and ML used in this work. The great benefit of using supervised learning is that we can teach the algorithm using synthetic models we think are correct. If they are not, we might be able to notice it by using this type of algorithm, since ML does not learn physics at any point: it is only capable of recognising patterns and features of the data. We use ensemble methods because they are one of the ways to reduce variance and improve the performance of the model. And as if this was not enough, decision trees are highly interpretable, one quirk neural networks do not have, even though they are much more powerful. Because of this, we can easily see what features are being used in the model, and how much they weigh inside the algorithm. They are also extremely easy to implement, and very malleable. This makes them great non-computationally expensive alternatives. Even though they have not been widely used with spectroscopic data, given their limitations to the size of the input features, they proved to be robust algorithms. In the following sections we will describe the way they are used in this work, their training and optimization process, and finally, the observational data results.

Chapter 2

Methods

In this Chapter we explain the methodology followed in this work. First, we describe the source and characteristics of the observational and synthetic data used to model the DA WDs. Then, we go into detail into how the ML algorithms were built, evaluated, and optimized, and finally used the predicted $T_{\rm eff}$, $\log(g)$, and distance of 262 DA WDs, using two different algorithms, and two different M-R relations. The results from the ML predictions are then used to compare the M-R relations, and determine in which zones of the $T_{\rm eff}$ and $\log(g)$ parameter space the predictions of distance are within the expected ML errors, and the *Gaia* errors, which places a hard constraint in the performance of our models. The methodology is also schematized in the flowchart at the end of this Chapter, in Fig. 2.10.

2.1 Observations

We chose a sample of 262 DA WDs with available Hubble Space Telescope UV spectroscopy. The selected targets (see Table E.1 in the Targets) were drawn from a sample studied by Sahu et al. (2023) (programs 12169, 12474, 13652, 14077, 15073, 16011, and 16642), who determined $T_{\rm eff}$ and $\log(g)$ by fitting synthetic models with χ^2 minimization. This study will provide measurements that can be compared with these previous results. All these targets have available *Gaia* DR3 parallaxes. *Gaia* provides us with extremely precise measurements of the position and distance to these stars, which allows us to place strict constraints on their size, since we can limit the solid angle the radiation is originating from. The solid angle allows us to determine their radius, with which we can compare and test the different M-R relations. The UV range is highly obscured, and thus having gas and dust in the ISM is something important to consider when working with UV data. We include their sky distribution along the Galactic extinction in Fig. 2.1. Most of the targets are above or below the Galactic plane, but there are a few of them that are located in more opaque regions.

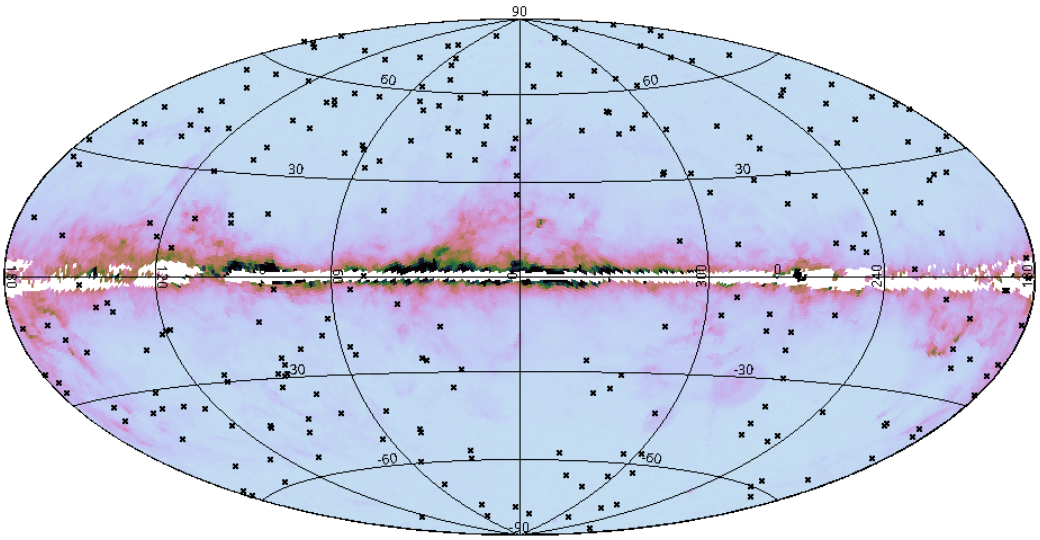


Figure 2.1: Sky distribution in Galactic coordinates of the observed DA WDs, over the Galactic extinction according to *Gaia* DR3.

2.1.1 Hubble Space Telescope

The *HST* is a space telescope that features a 2.4 meter mirror. It has five main instruments that can observe in wavelength ranges from the UV to the near-infrared. Because it is located above the Earth's atmosphere, it can obtain extremely high-resolution images. For this work we selected spectroscopy performed with the Cosmic Origins Spectrograph (COS), with the G130M grating at the 1291 Å central wavelength, covering from 1130–1430 Å, with a gap at 1278–1288 Å due to the physical separation of the FUV and NUV detectors. The exposure times of the COS observations ranged from 400 to 2000 s, with a median of 1225 s, and a median Signal to Noise Ratio (SNR)[1] of 12.04 between 1350 and 1400 Å. To accurately compare with the results presented by Sahu et al. (2023), we have also used the flux-calibrated spectra retrieved from the *HST* archive, which are processed with the COS pipeline CALCOS (v.3.3.4).

2.1.2 Spectrum Overview

In Fig. 2.2 we show one of the WDs used in this study. Among the main features of this spectrum (and of WD spectra in general) are the broad absorption feature centred around ∼1215.6702 Å, corresponding to the Ly$-\alpha$ feature. We can also see the *continuum*, usually modelled as a blackbody -a theoretical object that radiates all absorbed energy, with the colour or wavelength of maximum emission depending only on its temperature- dominating the emission from ∼1300 Å to longer wavelengths. For a summary of the analysis of the observational data see the flowchart

[1]SNR is calculated as the median flux divided over its flux error.

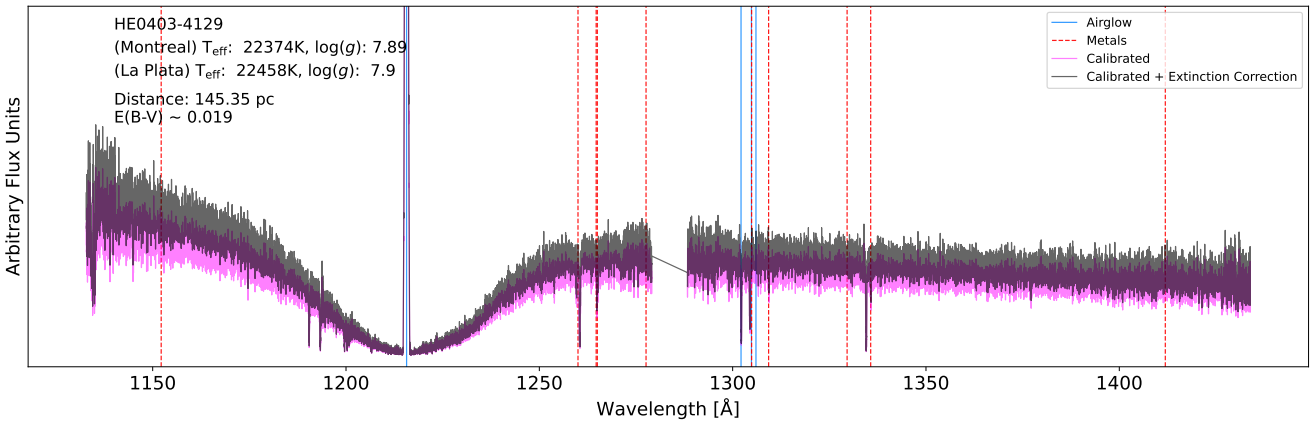


Figure 2.2: Spectrum of one of the WDs used in this work. Details about the object such as temperature ($T_{\rm eff}$), surface gravity ($\log(g)$), distance and excess color in[dex] ($E(B-V)$) used in the extinction correction are included in the upper left corner. The magenta spectrum corresponds to the calibrated data obtained from the *HST* archive, while the black one corresponds to the result of applying our reddening correction (Fitzpatrick, 1999) to the flux. The vertical blue lines show the wavelengths at which we expect geocoronal airglow emission lines, and the vertical red dashed lines show the location of common metals that pollute WDs and produce absorption features. The gray areas defines the region where typical metal absorption and airglow emission lines are detected in WD spectroscopy (see Table 2.1).

in Fig. 2.10.

The Ly$-\alpha$ emission is extremely important in determining the $T_{\rm eff}$ and $\log(g)$ of the star, since its shape is highly sensitive to $T_{\rm eff}$. As $T_{\rm eff}$ decreases, the absorption broadens at the core. On the other hand, the effect of $\log(g)$ is slightly more difficult to observe, since higher gravity values will flatten the *wings* of the absorption, that is, the edges of the feature where it starts to "dilute" with the continuum. By observing in the UV, which is the range the observation in Fig. 2.2 is in, we can determine these parameters by comparing with synthetic models, however, there are several corrections we need to make in order to obtain accurate results. First of all, we have elements polluting the surface of the star, which is not necessarily a problem, as we will see in Section 2.1.2.1, but is something we need to analyse separately. We also have, unfortunately, contamination from other sources, like the one that is produced by our own Sun on the Earth's atmosphere, that produces geocoronal airglow emission features, right at the centre of our most important absorption feature, at the Ly$-\alpha$ wavelength (see Section 2.1.2.2). Finally, as we showed in Fig. 2.1, we also need to account for extinction due to the ISM, since the UV range is highly affected by this absorption; these corrections are addressed in Section 2.1.2.3.

2.1.2.1 Metal Absorption Lines

One objective of studying WDs is to determine, apart from their own stellar parameters, the abundance of each contaminant that we detect in their atmosphere. These contaminants are nothing more than elements heavier than hydrogen and helium, or as we usually call them in astronomy, metals. These metals show up in the spectrum as thin absorption lines, since they are able to absorb photons at certain frequencies, although they were not noticeable in all of our spectra. Most techniques mask these features prior to fitting, since they are hard to account for in the atmosphere mdoels. The wavelengths of these absorptions are included in Table 2.1.

Element	Wavelength [Å]
Si II	1260.4
Si II	1263.7
Si II	1265.0
Si II	1309.3
C I	1277.6
C I	1329.6
C II	1335.7
N I	1411.9
O I	1152.2
O I	1304.9
Ly α	1215.6702
O I	1302.1685
O I	1304.8576
O I	1306.0286

Table 2.1: Metals lines which are masked during the feature sampling in the ultraviolet. Wavelengths are in vacuum. The second block lists the geocoronal airglow emission lines

2.1.2.2 Geocoronal Airglow Emission

HST observations can be affected by emission features in the far UV that are spectral features produced at the outermost region of the Earth's atmosphere dominated by neutral elements. For example, atoms like hydrogen and oxygen absorb and re-emit sunlight solar UV radiation through resonance fluorescence. Thus, they are prominent during daylight observations. These correspond to the Ly-α feature at 1216.6702 Å and the O I triplet at 1302.17, 1304.86, and 1306.03 Å. Since these lines are not produced by the observed objects, and have varying intensity depending on the telescope's position, we need to take them into account in our corrections.

The maximum intensity of the O I line is very weak, and since it was not prominent in most spectra, we decided to preserve this region and not mask it. However the Ly-α emission is significant in most spectra, as we can see in Fig. 2.2, and so it was necessary to correct it, since WDs have

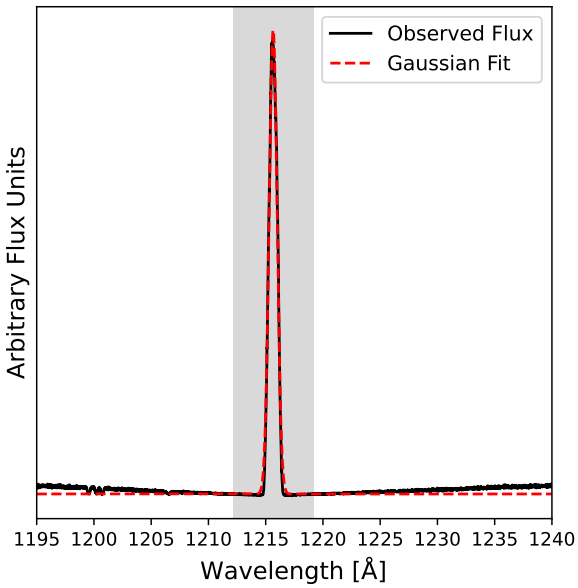


Figure 2.3: Observed geocoronal airglow Ly$-\alpha$ emission feature (black line) with its corresponding Gaussian fit (dashed red line). The mask applied to all spectra is shown as the filled gray region. The y-axis is in arbitrary units.

a very prominent absorption feature at this same wavelength, that is extremely sensitive to $T_{\rm eff}$ and $\log(g)$. While there are available templates[2] in the *HST* documentation to correct for this emission feature, we chose to perform a statistical analysis of the emission line's properties, by fitting a Gaussian curve to the emission profile. In Fig. 2.3 we show one of the fits performed around the airglow feature. Here we see the accuracy of the fit (red line), and the mask chosen to eliminate this section of the flux. In Fig. 2.4 we include the distribution of the amplitude of the line for the 262 WDs, where we find a median amplitude of 3.869×10^{-13} erg/cm^2/s/Å, and a median width (that is, the value of σ in the Gaussian fit) of 0.351 Å. Just as a comparison, in the COS handbook the reported intensity varies between 6.3×10^{-14} and 6.3×10^{-13} erg/cm^2/s/arcsec^2. We decided to mask the line with an exaggerated width, to completely avoid receiving any external flux or noise coming from this contamination, and chose to eliminate the regions spanning $\pm10\sigma$ around the Ly-α emission (grey band in Fig. 2.3).

Once again, it is important to note that the reason we focus on correcting the emission feature at this wavelength in a statistically significant way is to ensure that the ML algorithms we will use will not consider this feature as important, because it is not related to the radiation from the objects since it is produced by Sunlight fluorescence. More so, it falls in an extremely important region for determining the values of $T_{\rm eff}$ and $\log(g)$, which are two of the parameters we are looking

[2] https://www.stsci.edu/hst/instrumentation/cos/calibration/airglow, outlined in (Bourrier, V. et al., 2018)

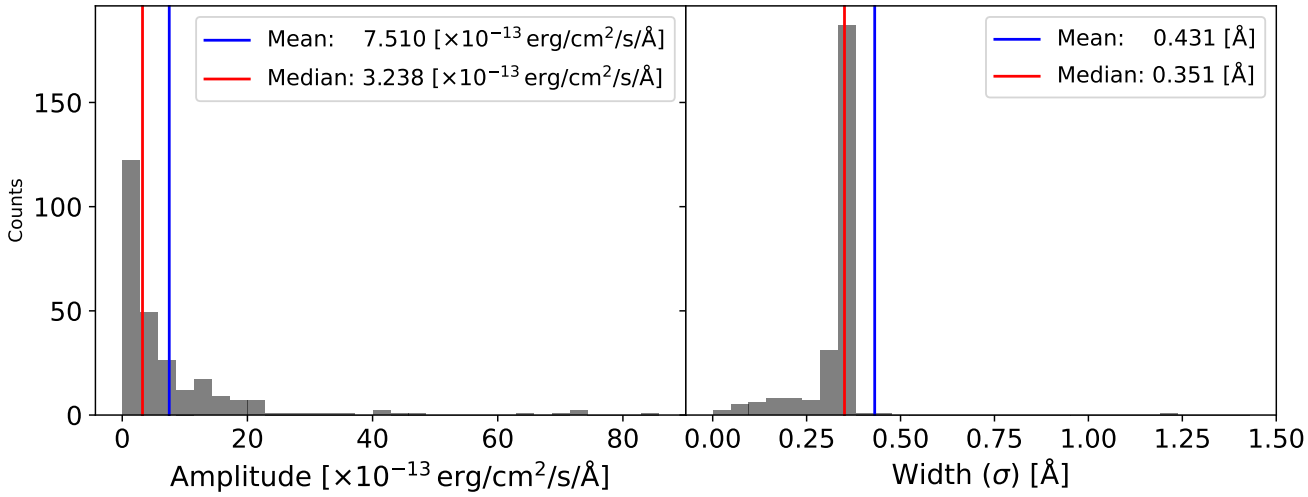


Figure 2.4: *Left*: Distribution of the amplitude of the Ly$-\alpha$ airglow emission feature in all spectra, the mean and median value are included as the blue and red vertical lines, respectively. *Right*: Distribution of the width σ of the line for the observed spectra, the mean and median values are shown as blue and red vertical lines, respectively. Since there are a few outliers in this distribution, we chose to use the median to set the mask for the airglow emission line, and we chose to eliminate the flux spanning $\pm$Å, corresponding to $\sim 10\sigma$ in each direction.

to predict with the ML models.

2.1.2.3 Galactic Extinction

As mentioned in the Introduction, we need to correct the observations to account for the extinction due to the ISM. Usually when fitting models the process is opposite: the model is reddened to match a single observation, but in this case, it was more appropriate to correct all observations, because all objects had different sky positions and distances, which meant they all had different excess colour indexes $E(B-V)$, which would make creating a single ML algorithm for all of them impossible. In order to *un-redden* the observations, we use the relation:

$$F_{\rm corr} = F_{\rm obs} 10^{0.4\times({\rm ext}+R(V))\times E(B-V)} \tag{2.1}$$

Where $F_{\rm corr}$ is the corrected flux, $F_{\rm obs}$ is the observed flux, and $R(V)$ is the ratio of the total to selective extinction given by $R(V) = A(V)/E(B-V) = 3.1$ (Sahu et al., 2024). The extinction (ext) corresponds to the wavelength dependent extinction from Fitzpatrick (1999), and the $E(B-V)$ value is obtained from the Stilism[3] 3D extinction maps (Lallement et al., 2014; Capitanio et al., 2017; Lallement, R. et al., 2018).

[3] https://stilism.obspm.fr/

In Fig. 2.2 we can clearly see the difference between the flux extracted directly from the *HST* archive (black line) and the one that has been corrected to account for the Galactic extinction (magenta line).

2.2 Synthetic Models

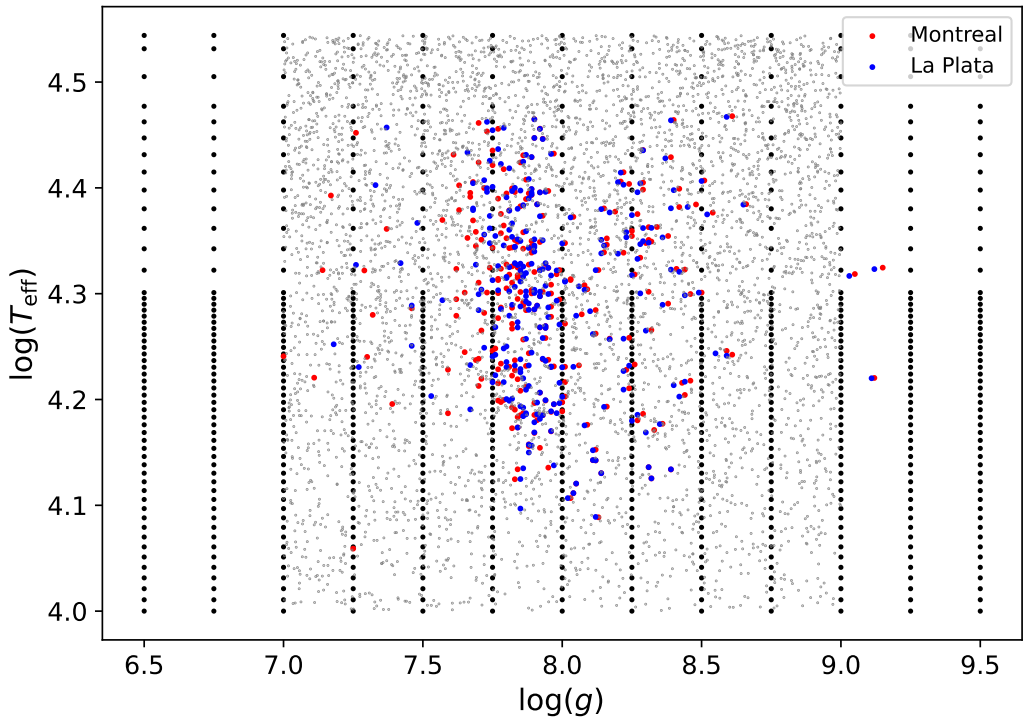


Figure 2.5: $T_{\rm eff}$ and log(g) from the DA WD grid from Koester (2010) (black dots). The observed objects with their corresponding $T_{\rm eff}$ and log(g) calculated by Sahu et al. (2023) are included for the Montreal (red) and La Plata (blue) M-R relations. The gray dots correspond to the $T_{\rm eff}$ and log(g) values used to create the synthetic spectra for the training set.

A summary of these sections can be found in the flowchart in Fig. 2.10 (see top box). In order to train the algorithms we used the synthetic DA WD models from Koester (2010), downloaded from the Theoretical Spectra section of SVO2[4] . From this grid we selected temperatures from 10000 to 35000 K, and log(g) from 6.5 to 9.5 [dex], corresponding to the black dots shown in Fig. 2.5. Here we also include $T_{\rm eff}$ and log(g) values calculated by Sahu et al. (2023), using two different Mass-Radius relations, that we will talk more in depth in the following sections. The relevant detail here is that most WDs are concentrated in log(g) values between 7.5 and 8.5 [dex], with a single outlier with low mass (see one WD with log(g) $\sim$7.0 and $T_{\rm eff}\sim$18000) and few outliers with

[4] http://svo2.cab.inta-csic.es/theory/newov2/

extremely high masses (three WDs with $\log(g) > 9.0$). The fluxes are given in $4\pi \times H_{\rm Edd}$, where $H_{\rm Edd}$ corresponds to the Eddington flux, defined as the first moment of the specific intensity:

$$H_{\rm Edd} = \frac{1}{4\pi} \int I_\nu \cos\theta d\Omega \tag{2.2}$$

Where $d\Omega = \sin\theta d\theta d\phi$ is the differential solid angle, and I_ν is the specific intensity:

$$I_\nu = \frac{dE_\nu}{\cos\theta d\Omega dA dt d_\nu} \tag{2.3}$$

That is, the fraction of energy (dE_ν) that travels through a projected surface area $\cos\theta dA$, with a solid angle $d\Omega$ in a time dt and with frequency $d\nu$.

It is important to note that the models had wavelengths ranging from ∼900 to ∼4000 [Å], but were interpolated linearly to match specific points using `interp1d` from `scipy` (`python`).

2.2.0.1 Grid Interpolation

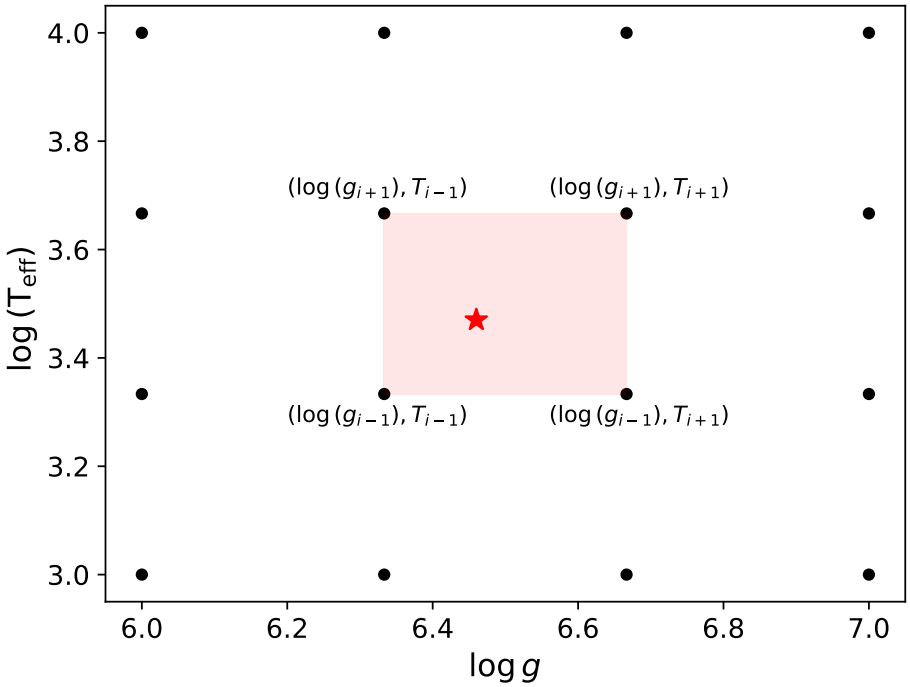


Figure 2.6: Schematic representation of the points the `LinearNDInterpolator` chooses to interpolate a new spectrum from the original synthetic model grid (black dots). The grid from Koester (2010) appears to be regular in the parameter range shown here, which means the points are equally spaced in a log − log plot, but this is not the case, and the spacing changes at higher $T_{\rm eff}$ values, which is why we chose this specific interpolator, since it can handle irregularly spaced data.

In order to obtain models with intermediate values of $T_{\rm eff}$ or $\log(g)$, we implemented an N-

dimensional linear interpolator in `python`, using `LinearNDInterpolator` from `scipy`. In our case, the interpolation is 2-dimensional, to obtain a specific $T_{\rm eff}= T_i$ and $\log(g)= \log(g_i)$, the function needs four coordinate pairs, as shown in Fig. 2.6. The benefit of using `LinearNDInterpolator` is that it does not require the grid to be regular, which is the case for the synthetic models we are using (as shown in Fig. 2.5), and that it can handle interpolation on arrays instead of single numbers, so it is able to perform the interpolation on the full spectrum all at once.

2.2.1 Eddington correction

In order to obtain synthetic data that represents the observed fluxes correctly, it was necessary to correct the synthetic fluxes, given in terms of the Eddington flux $\mathrm{H_{Edd}}$, to observed flux $F_{\rm obs}$, using the following relation:

$$F_{\rm obs} = 4\pi \mathrm{H_{Edd}} \left(\frac{R_{\rm WD}}{D}\right)^2 \tag{2.4}$$

Where R corresponds to the radius of the WD, and D corresponds to the distance, that is:

$$D = \frac{1000}{\omega}[\mathrm{pc}] \tag{2.5}$$

With ω corresponding to the *Gaia* parallax in mili-arcseconds ([mas]).

It is important to note that the radius of the WD will be determined entirely by $T_{\rm eff}$ and $\log(g)$, and the M-R relation chosen, and so the flux for the same synthetic model, with the same temperature and gravity will change depending on if we chose Montreal or La Plata to determine the radius.

2.2.2 Simulating Instrumental Noise

In order to reproduce the observational data in a realistic way, we added artificial noise to the synthetic spectra. In observations, this noise is due to the electronic components of the detector, the background contamination, and many other effects. It is usually approximated to a Gaussian distribution, and can be characterized with the Signal-to-Noise ratio (SNR). An observation with a very high SNR will have, as the name says, much more signal than noise, which means the spectrum will appear smoother. One available pipeline to calculate SNR of *HST* spectra is `DER_SNR` (Stoehr et al., 2008), that calculates the SNR as:

$$\mathrm{SNR} = \frac{\mathrm{signal}}{\mathrm{noise}} = \frac{\mathrm{median}(\mathrm{F_{SYN}})}{\frac{1.482602}{\sqrt{6}}\mathrm{median}(f_i + f_{i-2} + f_{i+2})} \tag{2.6}$$

Here the signal is the median value of the synthetic flux $F_{\rm SYN}$, and the noise is the median of $f_i + f_{i-2} + f_{i+2}$ (f_i is the ith point of the synthetic flux). This result is equivalent to the well known method of calculating the ratio of the median flux and its corresponding error, but we chose to use this expression since it only depends on the flux and SNR values, and can actually return an approximation of the noise. So, knowing that the observations had a median SNR of ∼12, we added a SNR of 10 to the observations, and we use the noise (that is, signal/SNR) as a proxy for the standard deviation of the flux error. We then obtained a random Gaussian distribution of these errors centred around 0, considering this value as the standard deviation, and added them to each point of the spectra.

2.2.3 Our Synthetic Sample

To train the models we randomly generate 5000[5] combinations of $T_{\rm eff}$, $\log(g)$ and distance, shown as grey dots in the background of Fig. 2.5. With these combinations we interpolate a synthetic flux (see Section 2.2.0.1) for each set of variables, which are then used to build the training and validation sets. As mentioned in Section 1.1.1, the Montreal M-R relation only spans from $\log(g)$ 7.0 to 9.0, and in order to train all algorithms based on the same parameter sample, that is, the same combination of $T_{\rm eff}$, $\log(g)$ and distance, we decided to keep these same limits, even though there are three WDs with $\log(g) > 9.0$. The reason is that we are unable to interpolate values from the Koester (2010) grid without a radius value, because we need to amplify the flux by $(R_{\rm WD}/D)^2$. For $T_{\rm eff}$ we chose a range from 10000 to 35000 K, and distance from 5 to 305 pc, which covers all our observed objects.

2.2.4 Feature Sampling

Feature sampling refers to the selection of specific wavelengths at which the ML model will attempt to make predictions. Most of our observed spectra have $\gtrsim$ 25000 wavelength points, due to the very high resolution of *HST*. Unfortunately, providing the full range of wavelength points is computationally prohibitive and will overkill the process. Therefore, it is needed to reduce the dimensionality of our data and the selection must focus on the minimum number of data points that encapsulate the essential information from the spectrum. As a reference, some studies (private communication) have successfully modelled the Balmer lines using 100 data points. In this work, we chose to sample some of the flux points to preserve no more than 1% of the original data. To

[5]The number of samples has been adjusted to ensure that the ML model reaches the "plateau" in the learning curves (described later) while remaining manageable for the computer used in the thesis. However, considering the uncertainties in the three parameters (i.e., 5% and 1% in $T_{\rm eff}$and $\log(g)$, respectively, as per (Barstow et al., 2001), and assuming a conservative uncertainty of 1 [pc] in distance), we estimate that approximately 3 million points would be required to adequately sample the parameter cube.

avoid performing additional calculations on the spectra, such as linear interpolation, we opted to divide the wavelength range into bins with varying resolutions (i.e. different intervals) and selected the median wavelength value; we then select the flux at this same location. This approach helps to mitigate the influence of absorption lines. However, we explicitly masked the Ly$-\alpha$ emission line, located around $\sim$ 1216 Å (as shown in Fig. 2.2), since its strength is too significant to be effectively diluted. The sampling resulted in 98 wavelength points distributed as described in Table 2.2. In Fig. 2.7 we see these wavelength points in blue and how this sampling differs for one observed spectrum (black). The figure also shows that there is a break in the spectrum due to the physical separation of the detectors.

	Wavelength range [Å]	$\Delta\lambda$ [Å]
Sampling 1	1130 to 1190	5
	1190 to ∼1212.07	0.5
	∼1219.27 to 1240	0.5
	1240 to 1450	20

Table 2.2: Wavelength ranges for the samplings used to bin the fluxes in the training, validation and test sets for our ML models.

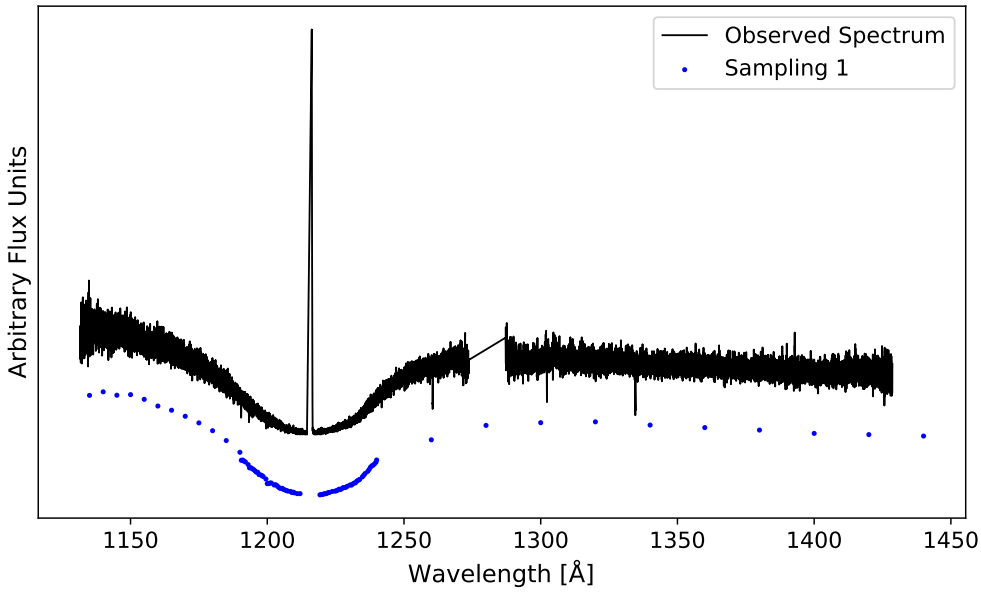


Figure 2.7: Observed spectrum used in this study (black line) with the feature sample used as input for the ML algorithms (blue dots). The flux is in arbitrary units and has been displaced upwards for better visualization.

It is important to note that in order to have the exact same amount and location of points to train the algorithm, we had to find and eliminate any wavelength bin where at least one of the

spectra had no flux points, which is why the sampling has some missing values. These same considerations were taken when processing the synthetic spectra, and we ended up with two different synthetic datasets, one for each M-R relation.

As shown in Figure 2.7, the sampling is primarily concentrated on the core of the Lyman α absorption line, which is the key feature providing information about $T_{\rm eff}$ and $\log(g)$, and less data points cover the remaining spectrum which is less sensitive to these parameters.

2.3 Machine Learning implementation

In this Section we describe the training, performance evaluation and optimization of the two ML algorithms chosen: Random Forest and Extreme Gradient Boosting. A scheme of these sections can be found in the flowchart in Fig. 2.10 (see centre box).

2.3.1 Training and Validation

	M-R	ML	R^2		RMSE			MAE			$d_{\rm Gaia}$ [pc]
			Train	Validation	$T_{\rm eff}$ [K]	$\log(g)$ [dex]	d [pc]	$T_{\rm eff}$ [K]	$\log(g)$ [dex]	d [pc]	
NO OPT	M	RF	0.999	0.994	549	0.25	23.21	404	0.19	17.30	
		XGB	0.999	0.993	607	0.21	19.95	451	0.16	14.84	
	LP	RF	0.999	0.994	563	0.26	24.42	413	0.20	18.50	
		XGB	0.999	0.992	616	0.22	20.16	461	0.16	14.62	0.24
OPT	M	RF	0.875	0.869	2573	0.53	61.62	1991	0.45	52.13	
		XGB	>0.999	0.994	553	0.18	17.822	398	0.13	12.77	
	LP	RF	0.871	0.863	2623	0.52	62.59	2035	0.44	53.05	
		XGB	>0.999	0.994	571	0.19	19.01	407	0.14	13.75	

Table 2.3: R^2 score, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for the sampling described in Section 2.2.4 for the M-R relations (Montreal, M, and La Plata, LP) and ML algorithm (RF and XGB), obtained in the training and validation process, that is, where the algorithms were only applied to synthetic data. These metrics are shown pre- and post optimization (see text for details). For comparison purposes, also shown is the median of the distance errors calculated from *Gaia* parallaxes.

From the 5000 synthetic model sample we use 4000 spectra to train the algorithms, and 1000 as a validation set. We include the results of a first try, with the default parameters included in `sklearn` for RF and in the XGB package, in the first row of Table 2.3. The top rows label as NO OPT in Table 4.3 we see that the training score for all the models is 0.999, with mean validation R^2 scores of 0.994. This could indicate that the model is either predicting the parameters very accurately, or that the data is noisy (that is, there is a lot of dispersion). The RMSE and MAE metrics are similar for all four models in each variable, but the Montreal M-R seems to introduce

less scatter in all three variables. Overall, it seems that RF is also performing better than XGB. However, this is not final, since we also went ahead with optimizing these models to improve their performance, with the aim to reduce the distance error, given that it is much higher than the *Gaia* value.

2.3.2 Optimization of Hyperparameters

Since previous tests were not conclusive on which performs the best, we performed a hyperparameter optimization on both algorithms, using both M-R relations.

Model	**Hyperparameter {Values}**
RF	`n_estimators`: {25, 50, 100, 150} `max_features`: {`sqrt`, `log2`, `None`} `max_depth`: {3, 6, 9} `max_leaf_nodes`: {3, 6, 9}
XGB	`n_estimators`: {50, 100, 200} `max_depth`: {3, 6, 9} `learning_rate`: {0.01, 0.1, 0.2} `subsample`: {0.8, 0.9, 1.0} `colsample_bytree`: {0.8, 0.9, 1.0}

Table 2.4: Hyperparameter grids for the optimization of the RF and XGB based models. Each combination of parameters was used to build the model, and test its performance on synthetic data, using the `GridSearchCV` package from `python`.

Optimizing means we implemented an exhaustive search for the best combination of hyperparameters using `GridSearchCV` from the `sklearn` package in `python`. Even though RF and XGB have common hyperparameters, since they are both based on decision trees, the way they are built is different, and so we established two different parameter grids, that is, two different sets of combinations, specific to each model. These are included in Table 2.4, while the definition of each parameter can be found in Sections 1.3.1.1 and 1.3.1.2 for RF and XGB, respectively. The optimization was performed with a 5-fold cross-validation, that is, the synthetic data was split into 5 parts, which means that we trained the model five times, to predict values for one of the subsamples each time being trained with the remaining 4 parts, each time changing the training set.

Once we obtained the optimized models we assessed their performance and analysed their bias and variance using learning curves, to select the best algorithm overall, which was used to predict $T_{\rm eff}$, $\log(g)$ and distance of the observed fluxes.

The resulting metrics after optimizing the models are included in bottom rows label as OPT in Table 2.3, with the obtained hyperparameters included in Table 2.5. From these results it is clear

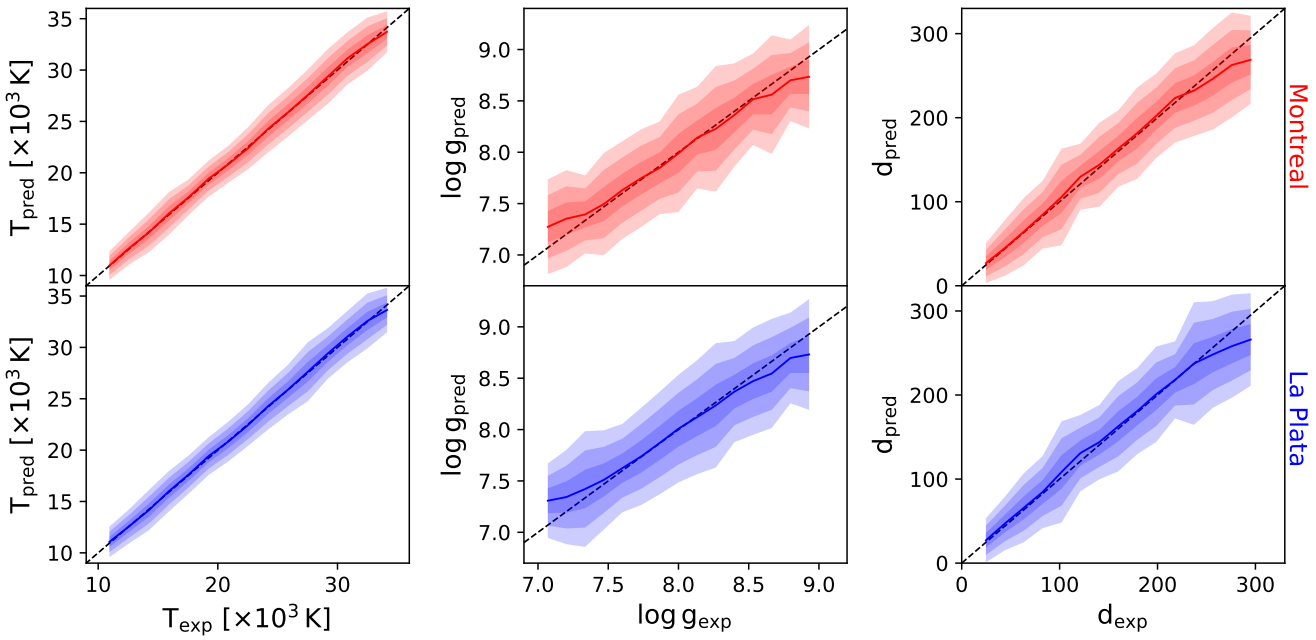


Figure 2.8: Expected and predicted $T_{\rm eff}$, $\log(g)$and distances for the Montreal (upper panels, red) and La Plata (lower panels, blue). The means are shown as solid lines for both M-R relations. The expected values correspond to the real target values used to interpolate the synthetic fluxes that are in the validation set. We see that for $\log(g)$ the predictions worsen for the smallest and largest values, while $T_{\rm eff}$ are somewhat accurately predicted in the whole range. The largest distances are underestimated for both M-R models.

that after optimizing, the best performing model is XGB, independent of the M-R relation used, given that it consistently has the highest R^2 scores, and the lowest errors. Because of this, we use this model to obtain the final predictions on the observational data.

In Fig. 2.8 we include the behavior of the validation set, as the blue or red line (Montreal and La Plata, respectively), with the shaded regions corresponding to a deviation of 1, 2 and 3σ. From the validation set results, we see that $T_{\rm eff}$ is the best predicted parameter, and it is not clear if there are regions where the temperature is being consistently under- or overestimated. For $\log(g)$ we see that both the Montreal and La Plata ML models struggle with the lowest and highest values, where we overestimate low $\log(g)$, and underestimate high $\log(g)$. For distance, once again the models behave similarly, and tend to underestimate the highest distances for the validation set.

2.3.3 Learning Curves

To add another performance test, we compute the learning curves of both optimized models, shown in Fig. 2.9, with scores corresponding to the MAE calculated for the three variables $T_{\rm eff}$, $\log(g)$and distance at the same time. This means that the score is not comparable to any of the metrics

Model	ML	Parameter	Original Value	Optimized Value
RF	M	n_estimators	100	100
		max_features	1.0	None
		max_depth	None	6
		max_leaf_nodes	None	9
	LP	n_estimators	100	150
		max_features	1.0	None
		max_depth	None	6
		max_leaf_nodes	None	9
XGB	M	n_estimators	100	200
		max_depth	6	9
		learning_rate	0.3	0.2
		subsample	1	1
		colsample_bytree	1.0	1.0
	LP	n_estimators	100	200
		max_depth	6	9
		learning_rate	0.3	0.1
		subsample	1	1
		colsample_bytree	1.0	0.9

Table 2.5: Hyperparameter values obtained from the optimization for RF and XGB. We also include the default parameters, used in the previous training and validation for comparison.

shown in previous tables, and is only used to determine how the performance of the model is changing..

Here we see that contrary to what we expect, the error increases for RF as the training set size increases, showing that this model is not performing well for our data. As we mentioned before (see Section 1.4.1.1), the expected behaviour for a learning curve is to stabilize after a certain training set size, when the model can no longer learn anything else, however, if the data has a high amount of noise, like this case, this might not happen for certain algorithms.

On the other hand, for XGB, the learning curves have the expected behaviour, with a training error that increases and a validation error that decreases with the training set size. The curves show a large separation, which indicates a high variance, while the low training error indicates low bias, which means our model has a high complexity, and its most likely over-fitting the training set. This could be fixed by limiting the complexity of the model thrugh the optimization of the hyperparameters. It is also worth noting that we recognize that there are other issues with the implementation (see Chapter 4), and it is possible that the performance of the algorithm is being limited by that, and the obtained scores are the highest possible value considering the underlying problems.

All pre-processing of the data, implementation and training othe ML algorithms, was done on

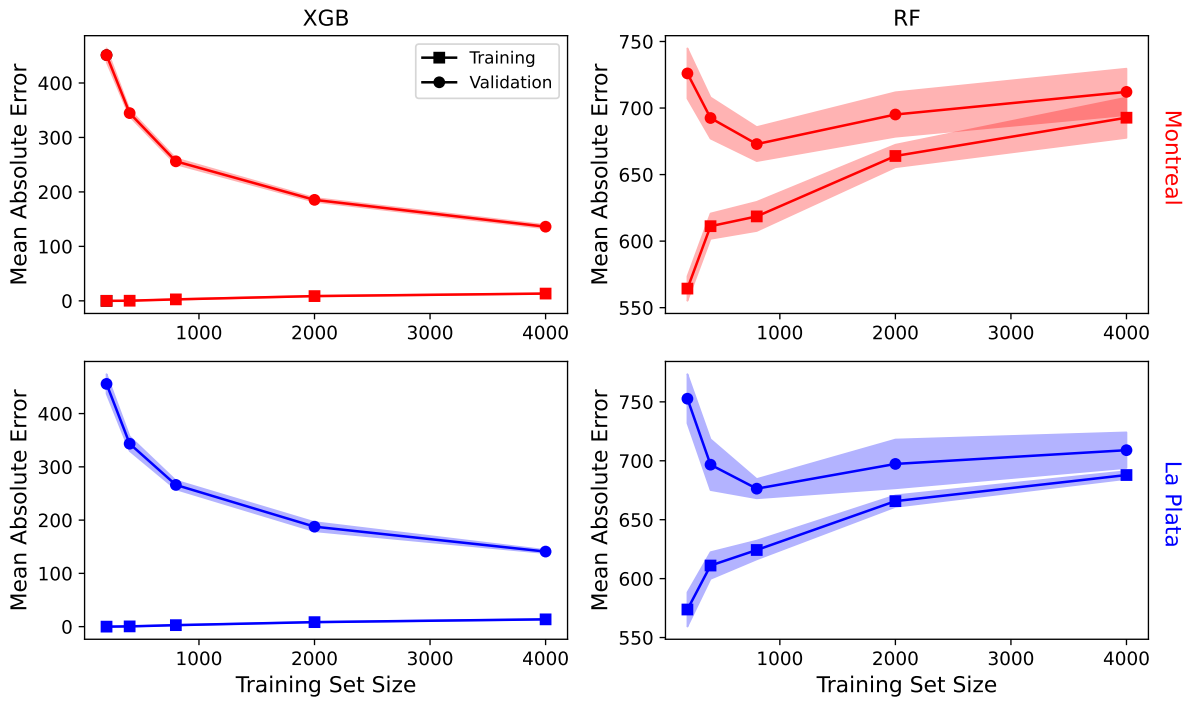


Figure 2.9: Learning Curves obtained for the RF (right) and XGB (left) ML algorithms, combined with the Montreal (red) and La Plata (blue) M-R relations, measured with the Mean Absolute Error of the validation set. From here we see that RF has a high bias for both M-R relations, while XGB has a higher variance, with a smaller training error, which means it performs better in this case.

a Intel® Core™ i5-10300H CPU with 8 GB of RAM.

Based on the results shown in this Section, we see that XGB is the best performing model in this case, with validation R^2 of $\sim$0.994 for both Montreal and La Plata. It also achieves the smallest distance MAE, of $\sim$12.5 on average. With our ML framework established, we now use the trained algorithms to predict atmospheric parameters of observational UV spectroscopy of 262 WDs, applying the corrections described in Section 2.1.

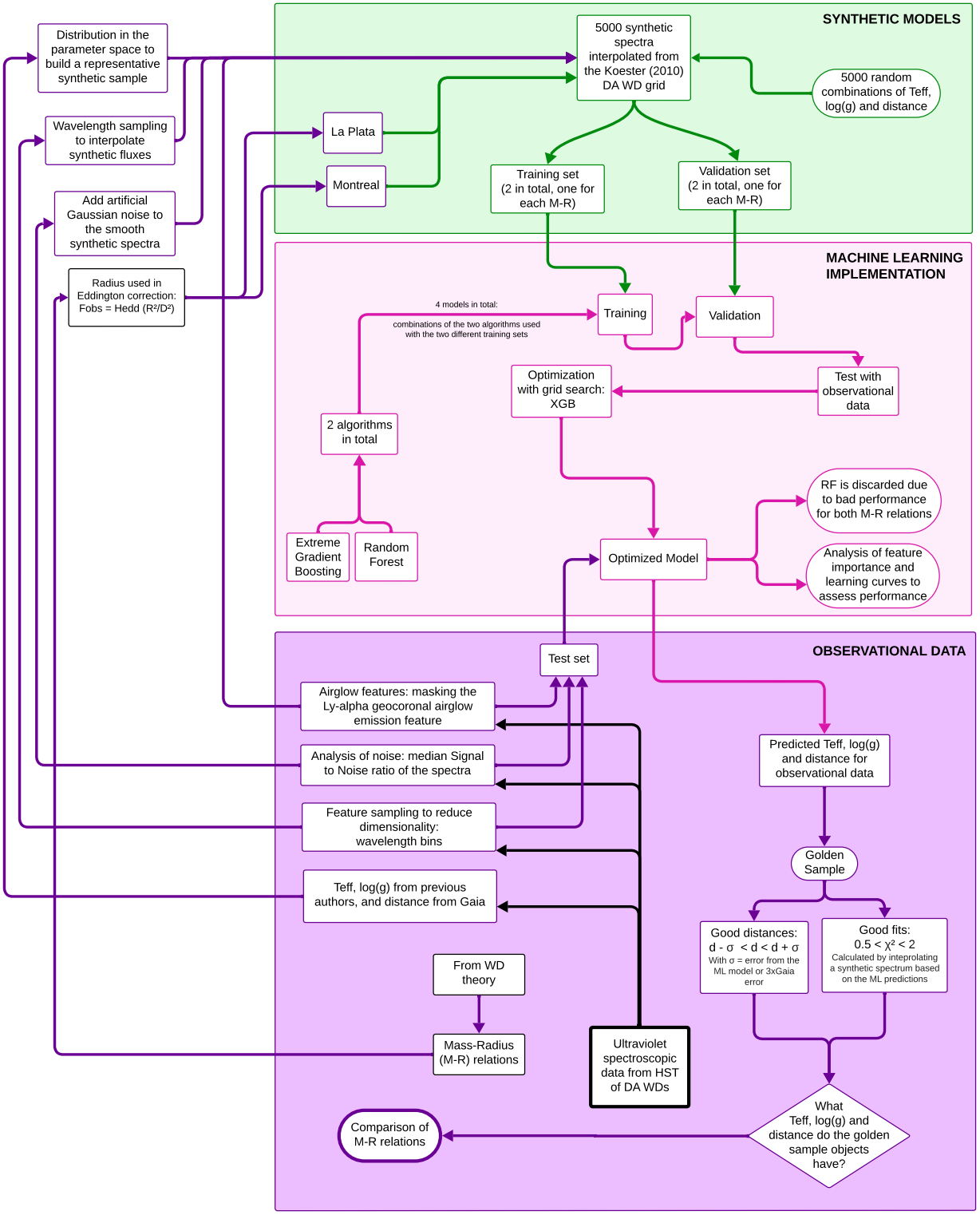


Figure 2.10: Flowchart that represents the methodology of this work, including the pre-processing of the observational data (sampling and masking airglow emission features), the building of the synthetic training and validation sets based on the main characteristics of these objects ($T_{\rm eff}$, $\log(g)$, distance and SNR), and the following training, optimization and final predictions of the model on the observational data. The figure is divided into three main sections, that represent the three main blocks of this work, the top box (green) contains the synthetic data, the centre box (pink) corresponds to the implementation of the ML algorithms and the bottom box (purple) contains the observational data.

Chapter 3

Results

In this Chapter we present the main results of this work, consisting of the obtained metrics for the optimized ML model, and the comparison of the M-R relations. From the comparison we determine which zone of the parameter space contains the best fits, comparing the observed spectra with its corresponding synthetic model, based on the ML predictions; with those results we define a golden sample, considering also which objects the ML distance predictions agrees with the *Gaia* values.

3.1 Best Performing Model: XGBoost

As mentioned in Section 2.3.2, the best performing model after optimizing the hyperparameters is XGB, for both Montreal and La Plata. In the following sections we include the feature importance obtained for this model.

While feature importance is used sometimes as a first assessment to select features that will go into the training of the algorithm, in our case it is only analysed to check that the model is taking into consideration all the input in a way that relates to the physics behind the data.

3.1.1 Feature Importance

Regarding the behaviour of the XGB algorithm, we include Fig. 3.1, which contains the ratio of feature importance (measured with the F-score in XGBoost, which reflects the frequency with which each feature is used for splits during training) between the XGB model trained with Montreal and La Plata. In the upper panel we include a synthetic example spectrum, with which we see that the most important feature, for this specific model is the first point of the flux. We believe this might be related to the fact that this is the point with the highest value in each spectrum, and is probably being used to determine the distance and $T_{\rm eff}$, since the order of magnitude of the flux changes drastically when one of these variables is varied, however we insist that the actual value

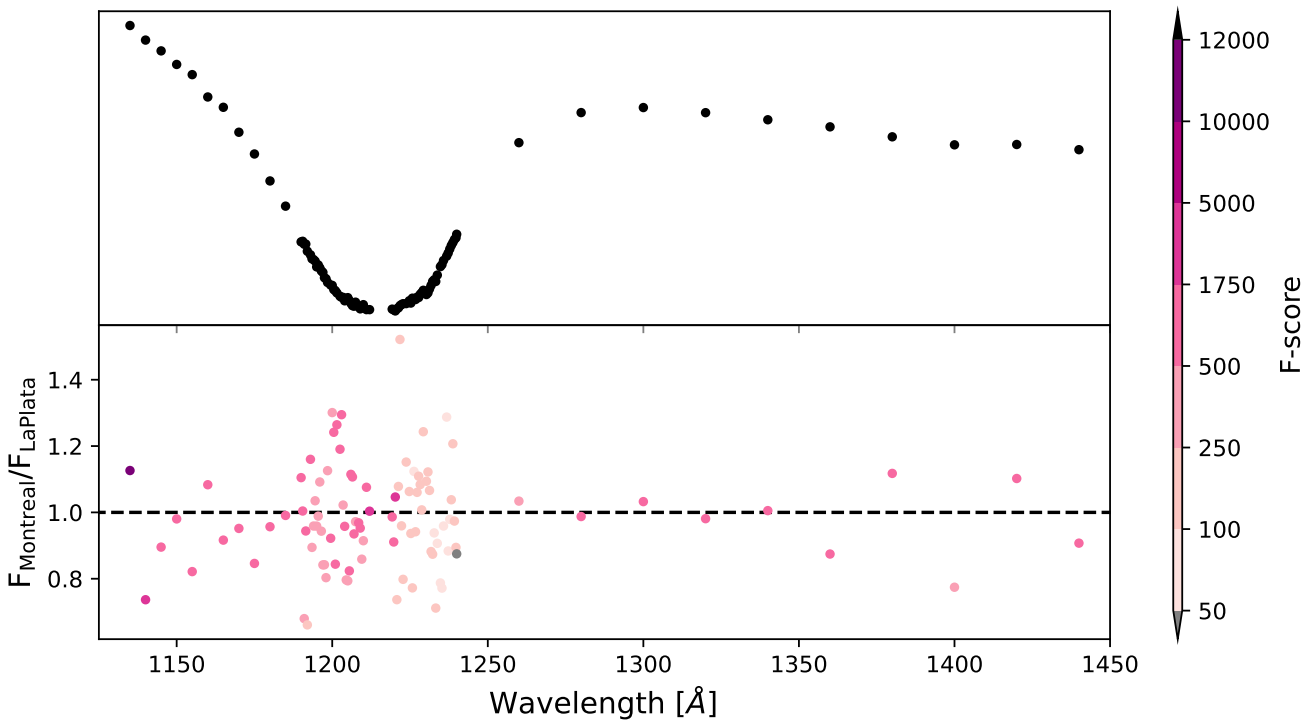


Figure 3.1: Ratio of the feature importance for Montreal and La Plata (lower panel). The colour of each dot represents the F-score (that is, how many times the feature was used in a split) for the Montreal XGB model. The black dashed line shows a ratio of 1, that is, when the importance was equal in both models. In the upper panel we include a reference spectrumm with which we can see that the importance is the highest for the first point, and its also considerable in the Ly-α absorption feature, which is consistent with what we expected from the physics we know affect the spectra: $T_{\rm eff}$ changes the slope and level of the continuuum, which affect the first point of the spectra, while $\log(g)$ changes the shape of the absorption.

of the score does not necessarily mean that this feature will always be this important, it would most likely change for another algorithm or even another iteration of the same model. On the other hand, we see that the core of the Ly-α absorption line is also important for both models. We chose this sampling knowing that the physics behind this type of star would modify the shape of this absorption and the overall shape of the spectrum when $T_{\rm eff}$, $\log(g)$ and distance had different values, and from this feature importance it seems that the algorithm was able to pick up some of that information, and is using most if not all of the features with a relatively high importance. It is also worth noting that La Plata and Montreal have very similar importances, as shown in the lower panel of Fig. 3.1.

	M-R	ML	RMSE	MAE
			d [pc]	d [pc]
NO OPT	M	RF	25.50	14.90
		XGB	27.87	18.06
	LP	RF	25.70	15.53
		XGB	29.97	21.32
OPT	M	RF	-	-
		XGB	26.77	19.71
	LP	RF	-	-
		XGB	29.03	19.35

Table 3.1: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for each M-R relation (Montreal, M, and La Plata, LP) and ML algorithm (RF and XGB), obtained with the test set, that is, where the algorithms were applied to observational data. The results are separated for the non-optimized models (NO OPT, upper row) and optimized results (OPT, lower row); we note that since RF shows a worse performance than XGB after optimization it was not used to predict values for observed spectra.

3.2 *HST* spectroscopy with XGBoost

Once the model was trained and optimized, we introduce the test set, that is, the observational data from *HST* spectroscopy, to obtain the predictions. It is important to note that this data was not introduced in the optimization or training, to avoid *data leakage*. Additionally, we do not have model-independent measurements of $T_{\rm eff}$ and $\log(g)$, and all comparison with previous work is included in the Discussion. The only parameter of which we know the values of is distance, and we include in Table 3.1 the resulting metrics for this variable. Note that R^2 is not included in this case, since it is calculated from the results of all the variable outputs, and we do not compare $T_{\rm eff}$ and $\log(g)$ with previous results in this Section. It is important to note that Table 3.1 does not include results on observational data for RF, because it did not perform well in synthetic data (see Section 2.3.2). For XGB however, we see that the distance error decreases after optimizing.

3.2.1 Comparing the Mass-Radius relations

Results of our predictions are in Fig. 3.2, which shows the distribution of $T_{\rm eff}$ and $\log(g)$, for the Montreal (red circles) and La Plata (blue dots) models. Predictions for the same object are joined with a black dashed line. At the top we show the distribution of the difference between $T_{\rm eff}$ for Montreal and La Plata. The mean (pink solid line) shows that the Montreal model predicts slightly higher temperatures than La Plata for 17000[K] $\lesssim T_{\rm eff}$ and the standard deviation (pink band) shows that the differences between the predictions diverge towards hotter white dwarfs ($>$ 25000 K), which is the temperature at which radiative levitation becomes relevant and yet is

crudely included in atmospheric models. In the right side panel, we include the distribution of the difference between the predicted $\log(g)$ value for Montreal and La Plata, here the mean (pink line) shows that at larger $\log(g)$ the Montreal prediction is higher than La Plata. At $\log(g)\sim 8$ this relation inverts, and we start seeing that for the lowest $\log(g)$, the La Plata prediction is larger. This trend is somewhat unexpected, as the models are similar except in the low $\log(g)$ region (see figure 3.3) which la Plata models predicts white dwarfs with larger radii for the same $\log(g)$. However, the differences are within the standard deviation of the $\log(g)$ measurements (pink band) and further improvements are required (see Future work). Fig. 3.2 also reflects the amount of uncertainty that our ML models have, given the low number of synthetic spectra provided during the training process. The connected points do not follow a clear distribution nor a pattern nor a trend. Indeed, the estimated values are significantly different, as shown in Fig. 3.4, where the colour of each point represents the difference between the Montreal and La Plata prediction, which clearly shows large discrepancies. This is in stark contrast with previous findings, in which the results in $T_{\rm eff}$ and $\log(g)$ agree within uncertainties for values around $\log(g)\simeq 7.75-8.25$ [dex] and $T_{\rm eff}\simeq 10000-30000$ [K](Sahu et al., 2023, a detailed comparison with this work is provided in section 4.3).

We also analyse the difference in the distance predictions between Montreal and La Plata, shown in Fig. 3.5. Here we see that the estimated distances are very similar for white dwarfs located between $50-150$ pc, and are within the ML errors, $\sim$12 [pc] (see distance MAE in Table 2.3) . For smaller distances it seems that La Plata predicts higher values, while at larger distances Montreal obtains larger distances. This can be seen clearly in the right panel of Fig. 3.5, where the difference between the predictions moves from negative, to close to 0, to positive, as distance increases. From the right panel in the figure also we can see that the predicted values diverge significantly at larger distances.

In Table 2.8 we showed that the distance MAE for Montreal and La Plata was $\sim$12.8 [pc] and $\sim$13.8 pc, respectively. These values are significantly larger than the *Gaia* error which for the sample analysed here it is of 0.241 ± 0.869 pc. Indeed, very few works have measured distances for WDs, in which they provide the *Gaia* distance as a prior. The work of Sahu et al. (2023) have reported distance errors and their mean and standard deviation are $\sim0.465\pm0.523$ [pc] and $\sim0.545\pm0.685$ [pc], for Montreal and La Plata, respectively. Their work shows high precision, with significantly smaller errors compared to ours. However, similar as for the other parameters, our large uncertainties might be due to the low number of points that sample the parameter cube (i.e. $T_{\rm eff}$, $\log(g)$ and Distance) when training the ML algorithm, or even the nature of the algorithm itself.

We can also analyse the implications the differences in the cooling models have on our estimates. In Fig. 3.3 we show the relation between radius and $T_{\rm eff}$ for Montreal (red) and La Plata (blue).

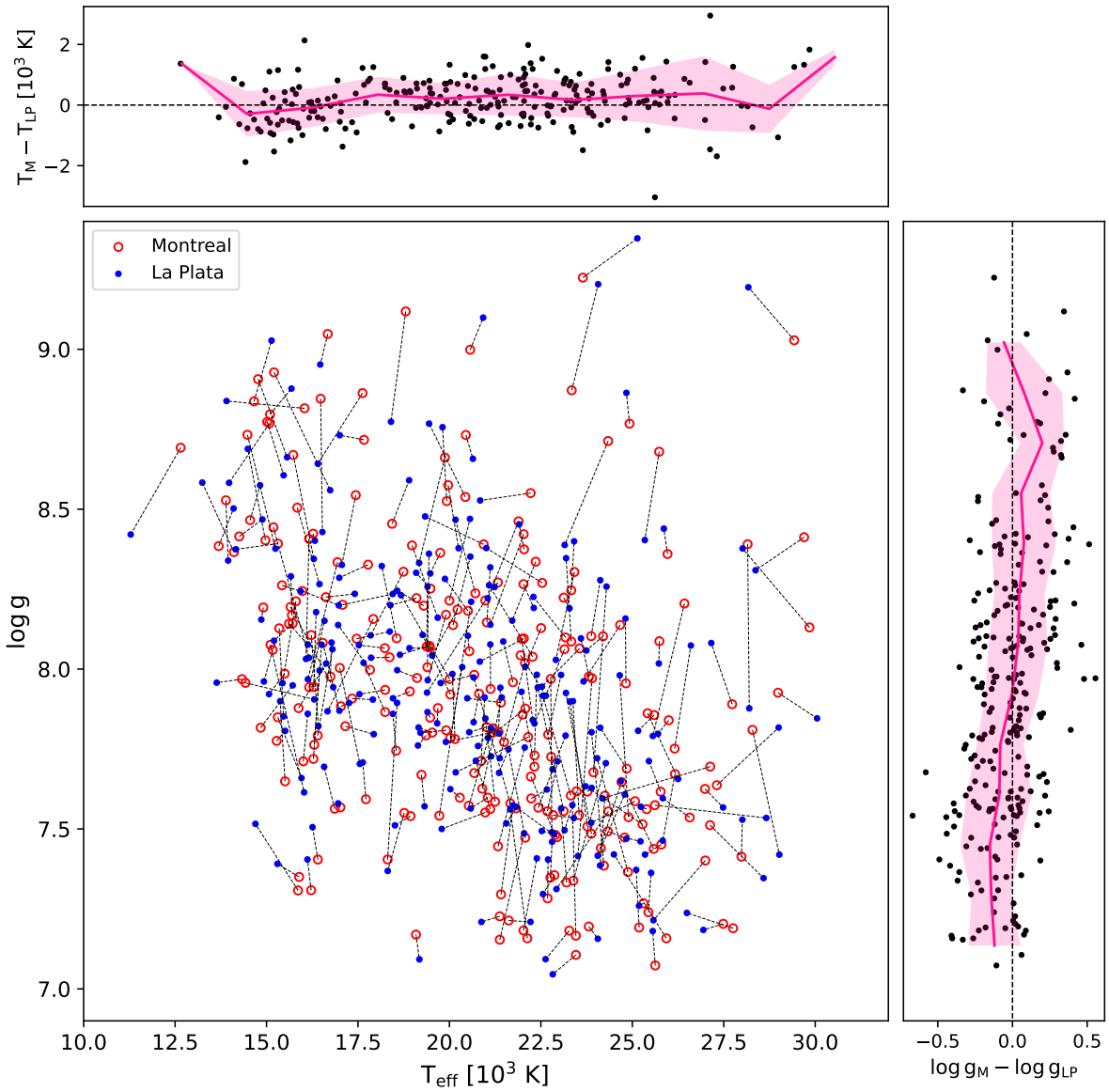


Figure 3.2: Predictions of $T_{\rm eff}$ and $\log(g)$ using Montreal (red circles) and La Plata (blue dots) cooling models. The results for the same white dwarf spectrum are joined by a black dashed line. At the top and left side we include the distribution of the difference between the prediction for Montreal and La Plata (black dots) with the mean at each $T_{\rm eff}$ or $\log(g)$ value (solid pink line), plotted with respect to the Montreal results. The filled pink band corresponds to a deviation of $\pm\sigma$ for each $T_{\rm eff}$ and $\log(g)$ value. As seen, the scatter of the data points is stochastic in the parameter space.

For $\log(g) \gtrsim 8$ we see that La Plata has larger radii values than those from Montreal. In addition, we can deduce from Eq. (2.4) that the radius and distance are tightly coupled. Thus, if the radii are consistently larger, as in the case of la Plata models, it is expected that the distance predictions will be large too. However, the right panel of Figure 3.5 shows that for the low $\log(g)$ values (i.e.

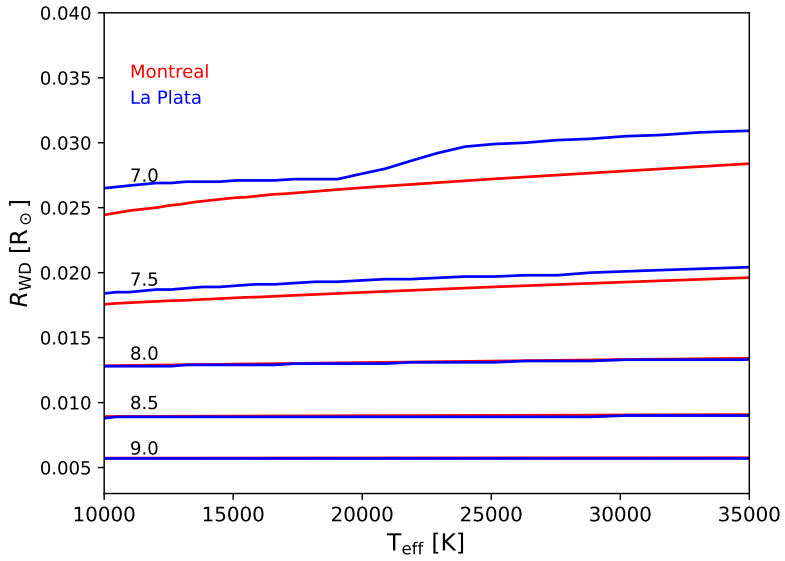


Figure 3.3: Radius for different $T_{\rm eff}$ and $\log(g)$ as predicted by the Montreal (red) and La Plata M-R relations. For $\log(g) \gtrsim 8.0$, the La Plata models calculates larger radii values than those from Montreal models.

those at distances farther than $\sim$150 [pc] due to observational biases which will be explained in section 4.1) it results the opposite, and we see that Montreal models predict larger distances for the WDs. However, this seems to be an issue that was propagated during the training process, since we also see this trend in the validation set results (see A.1 in the Validation Set Results). Indeed, the work by Yang et al. (2020), which utilized a sample of white dwarfs from SDSS (i.e. optical spectroscopy), applied neural networks to predict $T_{\rm eff}$ and $\log(g)$ from the shape of the Balmer lines. The observed spectra were normalized, meaning that distance information was not required. This normalization removes the coupling with the MR relations, allowing the analysis to rely solely on atmospheric models.

However, thanks to the Gaia parallaxes, we have independent distance measurements to identify accurate predictions of our method. In the top panels of Fig. 3.6 we show the difference in the distance prediction for Montreal (left) and La Plata (right); here we see that there are a few points where the difference is > 25 [pc] (2σ of ML error, magenta dots), to that of the *Gaia* distance. It is also important to note that this points do not seem to follow any trend in their location in the parameter space. If we focus on the points that have smaller differences, that is with predictions within +25[pc] (violet and purple), we see that they are mostly concentrated at the centre, and towards higher $\log(g)$ and lower $T_{\rm eff}$ values. Finally, the points with differences of < -25[pc] (blue and cyan) are located at the lower right corner, that is at the lowest $\log(g)$ and highest $T_{\rm eff}$ values. Here the distances are constantly being overestimated for both cooling models. So, we identify the region at low $\log(g)$ as problematic, and somewhat expected as it is the region in which the

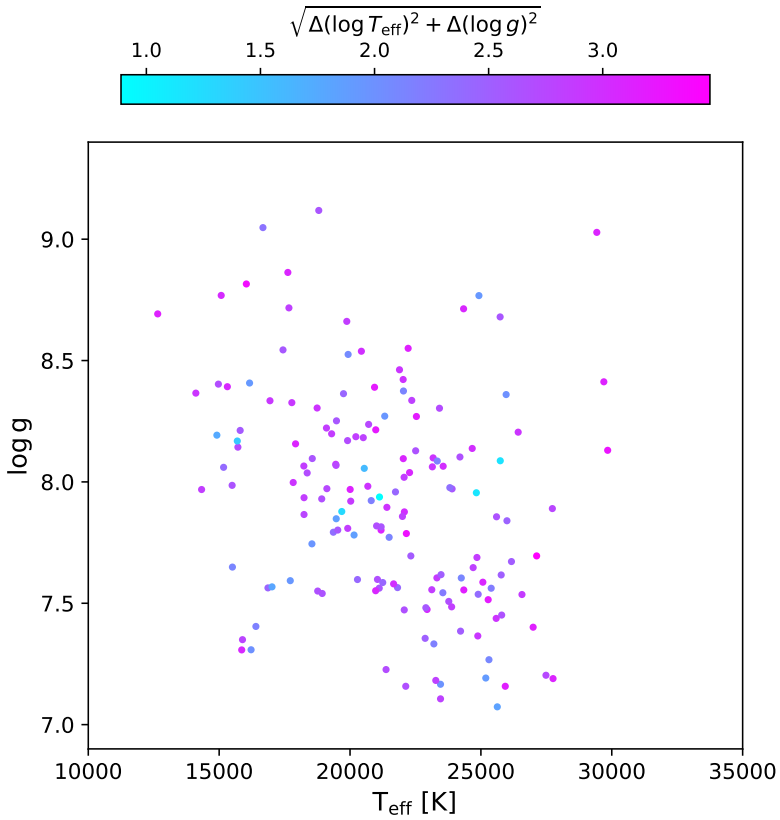


Figure 3.4: Absolute distance of the measurements predicted using Montreal and La Plata models in the plane of $T_{\rm eff}$ and $\log(g)$. The smallest distance is presented by cyan colour of the dot, which means that the more similar are the measurements from the different models. However, we see that the estimates are significantly different.

cooling models diverge significantly and thus it requires revision.

In addition, we observe a systematic trend in the predicted distances for both cooling models. The middle panel of Figure 3.6 shows that larger distances correspond to the hotter and lower $\log(g)$ region of the parameter space. This trend is once again influenced by the radius: smaller white dwarfs appear less luminous due to their reduced surface area, which limits the radiation emitted, and this limits their detection distances. Consequently, in this parameter space, we can detect distant white dwarfs with larger radii, or conversely low $\log(g)$.

Since our observations were in the UV range, we need to determine if the error was concentrated in the most reddened stars. The bottom panels in Fig. 3.6 show the reddening $E(B-V)$, for Montreal (left) and La Plata (right). From this figure we see that there exist a few cases in which the stars are highly reddened (i.e. $E(B-V) > 0.05$; magenta and violet points in the lower panels), which correspond to the stars with the largest distances (> 100 pc; magenta and violet points in the centre panel, indeed the six white dwarfs lie between $\simeq \pm 30$ degrees, and two of them are in the galactic plane; see the distribution in Fig 2.1). This comes from the fact that the

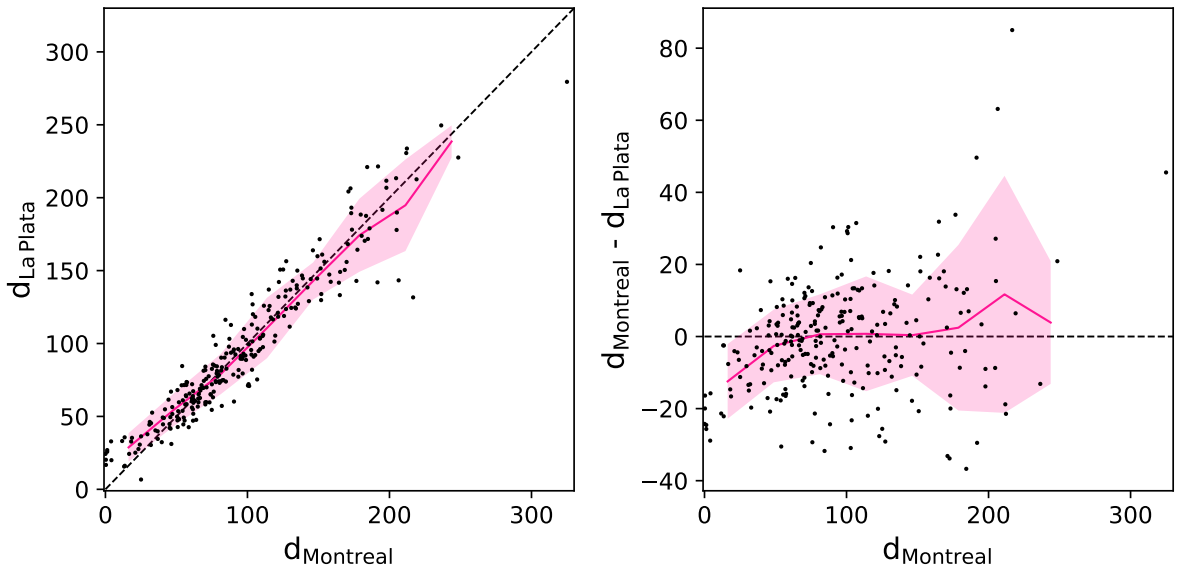


Figure 3.5: Predicted Montreal and La Plata distances (black dots). The left panel shows the correlation between these two predictions, and we can see that apparently both models predict similar distances for the objects. The pink line represents the mean La Plata distance at each Montreal value, and the pink filled region corresponds to a 1σ deviation. In the right panel we show the difference between Montreal and La Plata predictions. Once again, the pink line represents the mean value, and the filled region a 1σ deviation.

further away a star is, the more ISM the light has to travel through, and therefore, the more it its absorbed in its journey to us. What is interesting, is that these stars are located at low $\log(g)$ and high $T_{\rm eff}$, which means they also should be the brightest, and their large flux requires larger surface area e.g. larger radii or alternatively lower $\log(g)$, and the WD needs to be hotter as it radiates more flux.

As already pointed out in order to draw robust conclusions we should analyse the predictions which agree with the *Gaia* distances. Therefore, we have a motivation to select a sub-sample that includes only good fits, that is, those with χ^2 close to 1, and those where the distance prediction was close to the real (*Gaia*) distance.

3.2.2 Golden Sample

As mentioned, we should draw conclusions from a reliable sample, in order to determine how many of these results give us accurate predictions of $T_{\rm eff}$, $\log(g)$ and distance. Hence, we chose to filter the sample using two criteria: first, we analysed the *goodness of fit*, by calculating the χ^2_r between the observed flux, and a synthetic flux, at the same wavelength points, interpolated from the Koester (2010) DA WD grid, considering the M-R relation of the ML model and the predictions

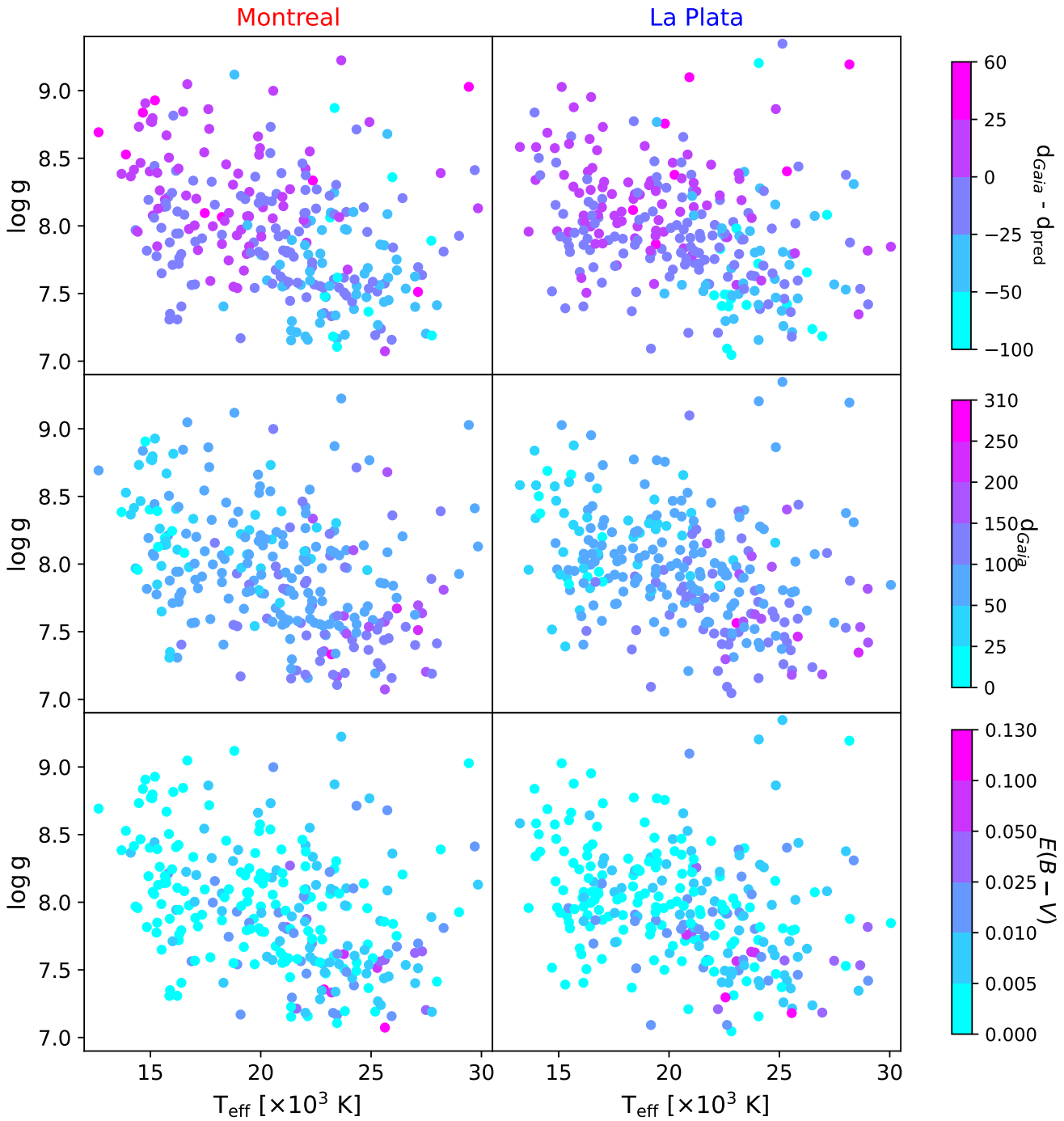


Figure 3.6: $T_{\rm eff}$ and $\log(g)$ values for Montreal (left) and La Plata (right). In the upper panels the color represents the distance, and for the bottom panels it corresponds to the reddening of each object $E(B-V)$. Here we see that the most reddened stars correspond to the ones with the largest distance, with is expected since we would have mode CSM interrupting the radiation's path. The furthest stars have greater temperatures, and lower $\log(g)$, which makes them the least massive, but the biggest in size, since the M-R for WDs is inverted.

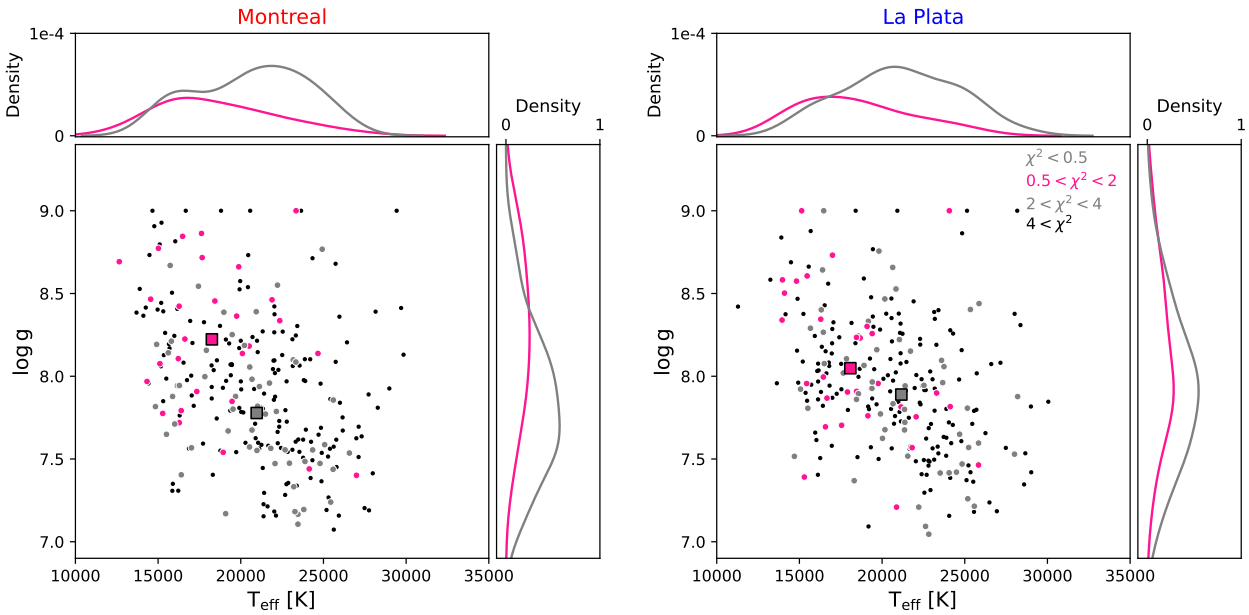


Figure 3.7: Predicted $T_{\rm eff}$ and $\log(g)$ for Montreal (left) and La Plata (right). The largest panels show where the best fits (pink dots) are located, that is, those where the $0. < \chi^2 < 2.0$ between the sampled flux that was the input to the ML model, and the interpolated flux obtained with the predicted targets. We also include bad fits ($\chi^2 < 0.5$ and $2.0 < \chi^2 < 4$) as grey dots, and predictions that did not reproduce the observed flux (black dots). For each M-R relation, we also include the Kernel Density Estimate (KDE), in the direction of each axis, for the good and bad fits (pink and gray line, respectively). The median of these distributions is plotted as a pink or gray square, depending on the category. It is important to note that, since each model has its own M-R relation, and therefore different results, the filtering process for each category was done independently, and we do not expect the same object to necessarily have good predictions in both M-R relations.

made. We consider good fits[1] as those with $0.5 < \chi^2_r < 2$, while bad fits had results lower than 0.5 or higher than 2. In this case, we decided to add a third classification, to contain all the fits that are simply incorrect; these correspond to those with $\chi^2_r > 4$. After filtering the results, we obtained the objects shown in Fig. 3.7. Here the larger panels show the distribution of the points in the parameter space, and the upper and side panels show the Kernel Density Estimate (KDE) for each category (good, bad and incorrect fit as magenta, grey and black dots, respectively). These distributions clearly indicate the presence of two distinct groups of data points that can be separated. The median value for each distribution (squared in the figure) separates, with good fits being located at higher $\log(g)$and lower $T_{\rm eff}$, and bad fits being in the same problematic zone

[1]It is important to note that most works consider masks over the metal and airglow features before calculating the χ^2 value, since these features will affect the results of the metric; here we only masked the Ly$-\alpha$ airglow feature, so we need more refinement to determine if these metal absorption and other emission features could be causing us to mis-label some objects as bad fits.

as before: large $T_{\rm eff}$ and small $\log(g)$. It is also important to note that the filtering was done separately for each MR relation, since resulting predictions, and the interpolated flux are different.

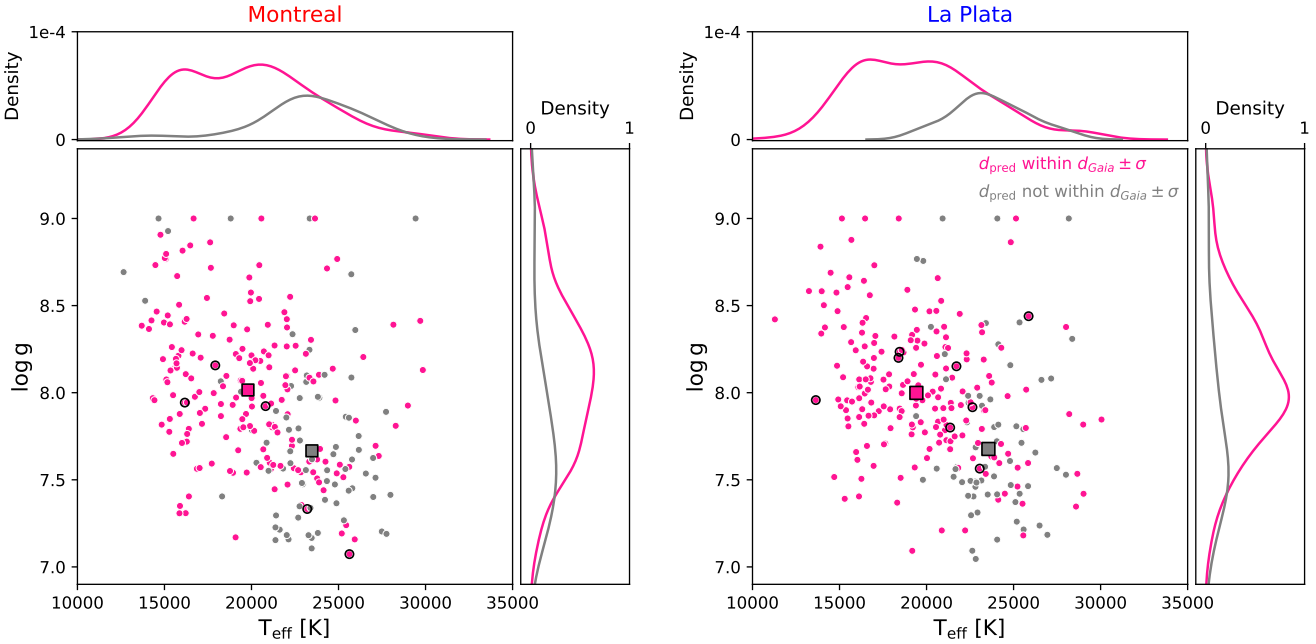


Figure 3.8: Predicted $T_{\rm eff}$ and $\log(g)$ for Montreal (left) and La Plata (right). The largest panels show where the good distance predictions are located, that is those were the d_{Gaia} distance is within the ML model MAE. The bad distance predictions are included as gray dots. For each M-R relation, we also include the Kernel Density Estimate (KDE), in the direction of each axis, for the good and bad predictions (pink and gray line, respectively). The median of these distributions is plotted as a pink or gray square, depending on the category. As an additional constraint, we have outlined the predictions where the distance was also within the *Gaia* error, which is much smaller than the MAE of our algorithm. It is important to note that, since each model has its own M-R relation, and therefore different results, the filtering process for each category was done independently, and we do not expect the same object to necessarily have good predictions in both M-R relations.

Our second criterion was to determine which objects had good distance predictions, that is, which of the predicted distances fell within 3σ (considering σ as the distance MAE for each M-R relation from Table 2.3) of the expected (*Gaia*) value. In Fig. 3.8 we include the results for both M-R relations. Here, the gray dots are the predictions where the distance was not accurately predicted. In pink we have the good distance predictions, and we can see from the KDE in the upper and left panels that once again these cases are located at higher $\log(g)$ and lower $T_{\rm eff}$ values, while bad predictions are still displaced to the lower right corner. Additionally, we include the cases where the distance prediction was within $3\sigma_{Gaia}$ (black outline), that is, the parallax error of each object. This is an important constraint, since it means that the algorithm is returning a reliable results, since *Gaia* errors are much smaller than what we find with our ML algorithms.

The spectrum for all objects in the golden sample can be found in the Observed Fluxes: Montreal Golden Sample (section 3.2.2).

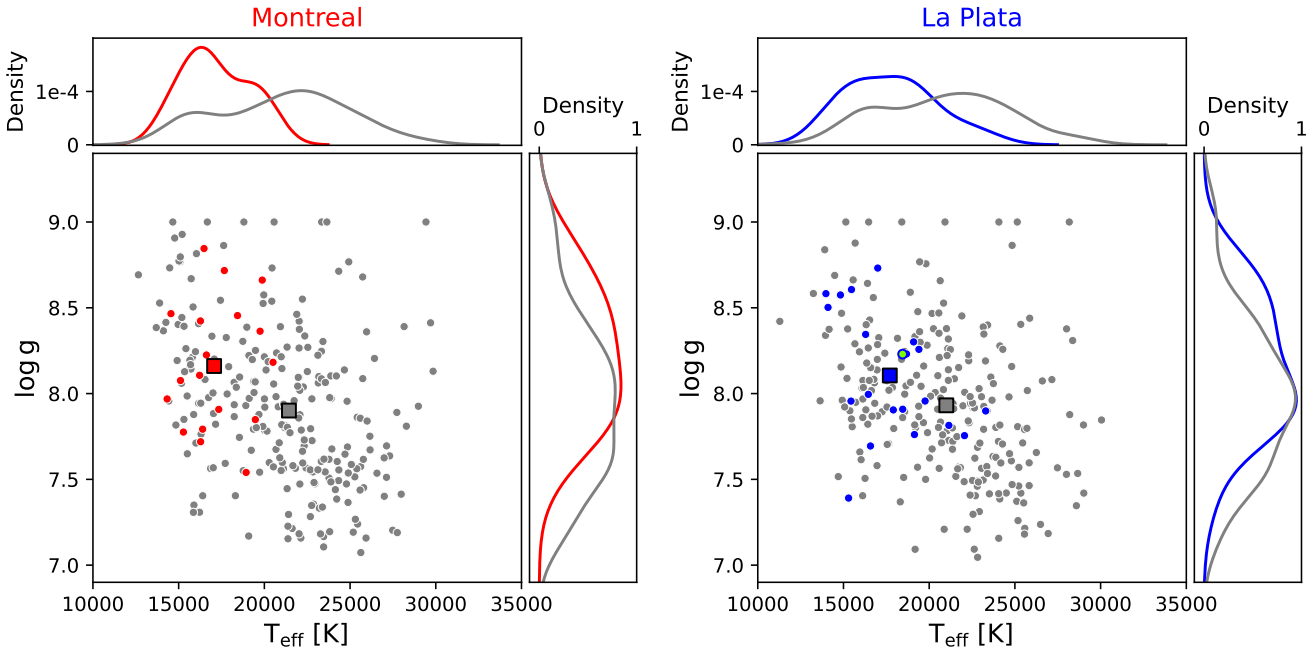


Figure 3.9: Predicted $T_{\rm eff}$ and $\log(g)$ for Montreal (left) and La Plata (right). The upper and side panels show the KDE of the golden sample (red and blue solid lines) and all other objects (gray lines). We also include the object from La Plata that has a distance prediction within $3\sigma_{Gaia}$ as a cyan dot with a blue outline.

In Fig. 3.9 we show the $T_{\rm eff}$ and $\log(g)$ of the golden sample for Montreal (left, red dots) and La Plata (right, blue dots). Here we see that the sample in both cases is displaced towards higher $\log(g)$ and lower $T_{\rm eff}$ values (red and blue squares, for Montreal and La Plata, respectively), while the rest of the points are displaced towards the problematic zone: higher $T_{\rm eff}$ and low $\log(g)$, agreeing with our previous results.

The obtained final sample consist of 18 objects for Montreal (spectra in Appendix B) and 21 for La Plata (spectra in Appendix C). Two objects are both in the Montreal and La Plata golden sample. Their spectra are shown in Fig. 3.10. The bottom panel shows WD1449+513, and from the figure we can see that both models estimate somewhat accurately the flux, with Montreal (red dashed line and dots) giving a slightly higher value, given that the predicted temperature is ~ 700 [K]higher than the temperature predicted by La Plata. Nonetheless, both fits are contained in the noise of the observed spectra, and we also see that the obtained $\log(g)$ value is extremely similar for both, as well as the predicted distance. The upper panel shows HS0400+1451, and in this case the results for $T_{\rm eff}$ and $\log(g)$ are similar for both M-R relations. They show prediction distances with a difference of ~ 5 pc, which probably is due to the fact that La Plata has a higher $T_{\rm eff}$

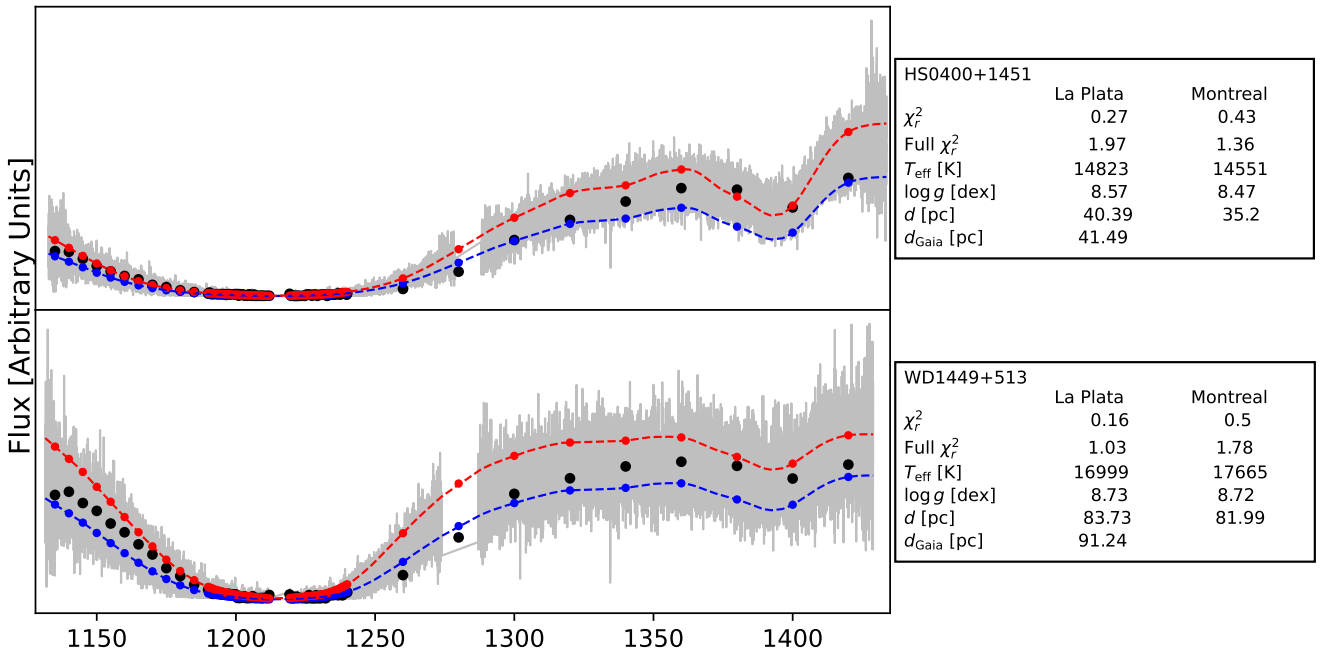


Figure 3.10: Observed flux (gray line) with the flux sampling used as input in the ML model (black dots). The predictions obtained with Montreal (red) and La Plata (blue) are included as the full spectrum (dashed lines) and in the sampling used in the models (dots). The information of the object and the predicted values for $T_{\rm eff}$, $\log(g)$ and distance for each model are included in the box at the right side. Both objects are in the golden sample of both M-R relations.

prediction, which means the ML places it farther away to scale the flux.

It is important to note that the reason both of these objects show Montreal fluxes higher than La Plata, even when they have similar parameters, is due to the fact that for $\log(g)$ higher than ~ 8.0 [dex] the La Plata radii are smaller than Montreal, which means that for the same object the flux will be brighter in the Montreal grid. Note also that the effect of the three parameters is somewhat degenerate: $T_{\rm eff}$ scales the flux because a hotter object will be brigther, $\log(g)$ is closely related to the radii, so smaller $\log(g)$ will mean a larger radii, and thus a larger surface that radiates, and finally, the distance of the object determines how bright it appears, farther away objects will seem fainter even if they have the same atmospheric parameters.

In Fig. 3.11 we show one of the objects where the predicted distance with the La Plata M-R was close to the *Gaia* value, and the χ^2 was within our good fit limits. Here the Ly-α emission at $\sim$1216Å was masked. The observed flux corresponds to the gray line, and we can see that the predictions not only try to match the input (magenta dots), but they also fit the rest of the wavelength range somewhat accurately. We see that the Ly-α absorption feature is very well represented by the synthetic flux (dashed line and blue dots). The χ^2_r value was calculated considering only the wavelength sampling that was used as input for the ML model (magenta and

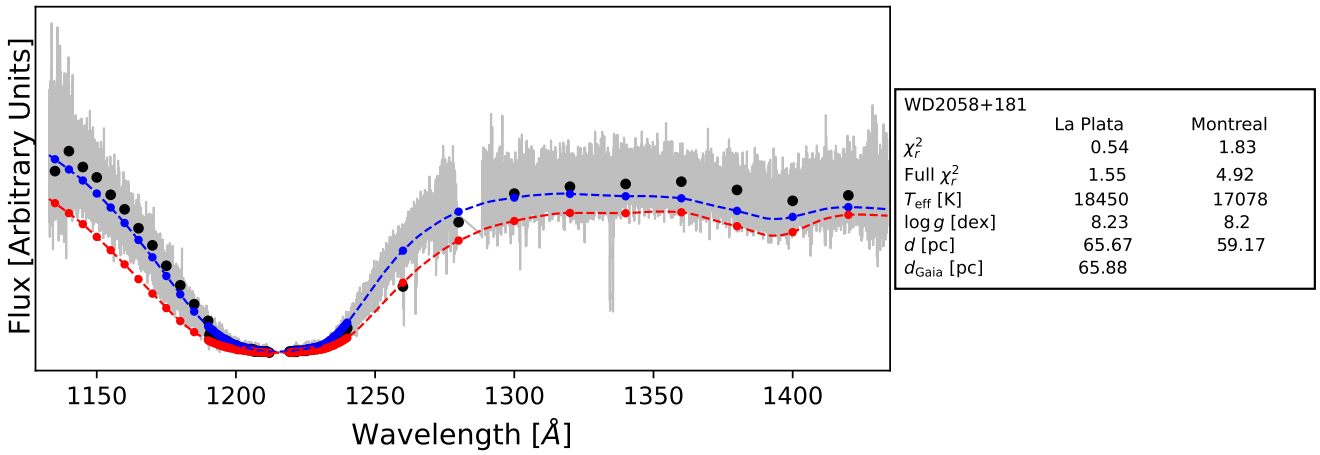


Figure 3.11: Observed flux (gray line) with the flux sampling used as input in the ML model (black dots). The predictions obtained with Montreal (red) and La Plata (blue) are included as the full spectrum (dashed lines) and in the sampling used in the models (dots). The information of the object and the predicted values for $T_{\rm eff}$, $\log(g)$ and distance for each model are included in the box at the right side. This object has a La Plata distance prediction within $\sim 1.45\sigma_{Gaia}$, meaning the distance value obtained is extremely accurate.

blue dots), and with the full spectrum, since we considered a much higher resolution at the Ly-α core than in the rest of the continuum. For this object, these values correspond to 0.54 and 1.55, respectively.

3.2.3 Mass Distribution

From the prediction of $T_{\rm eff}$ and $\log(g)$ we can also obtain the mass for each WD, following the same procedure to interpolate as we do for the grid of synthetic models (see Section 2.2.0.1). The obtained mass distribution is showed in Fig. 3.12. Both Montreal (blue, left panel) and La Plata (red, right panel) show asymmetric distributions, given by the fact that there are WDs with higher masses. Our mean value of $0.60 \pm 0.18 M_\odot$for the Montreal M-R is in perfect agreement with the results shown by Sahu et al. (2023); Tremblay and Bergeron (2009). Our La Plata mass is slightly higher than the result from Sahu et al. (2023), with a mean value of 0.62 ± 0.22 $M_\odot$; their value ($\sim 0.61 M_\odot$) is however, inside our standard deviation.

The golden sample for each M-R (shown as the darker histogram in Fig. 3.12) follows a different trend. To understand this, we can think of the mass in terms of the $\log(g)$ and radius of the star: for the same $T_{\rm eff}$, a smaller mass WD will have a lower $\log(g)$ value, and a bigger radii. This means that these WDs are most likely not entering the golden sample because they are in the problematic zone, since the best fits are displaced towards higher $\log(g)$, that is, larger masses (see Fig. 3.9). On the other hand, we are able to recover some of the most massive stars in this

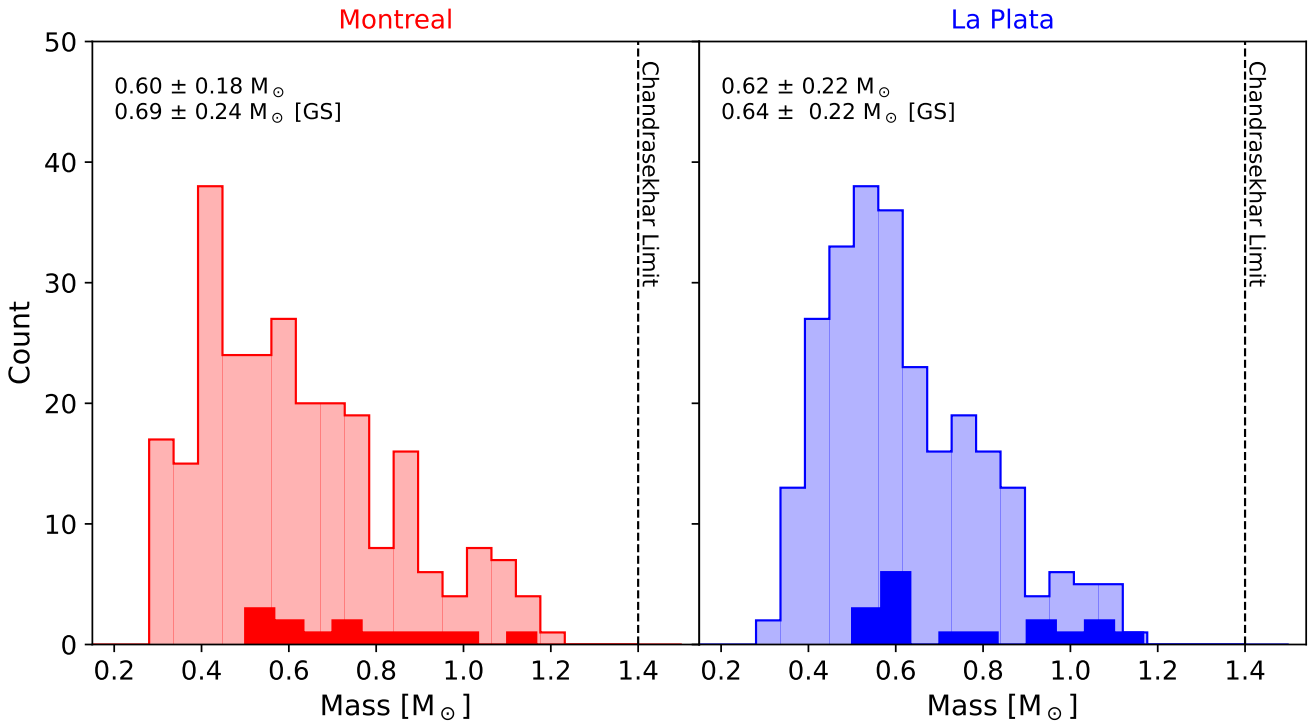


Figure 3.12: Distribution of WD masses for Montreal (red) and La Plata (blue), calculated from the predicted value of $\log(g)$ and the radius obtained from the respective M-R relation. The golden sample is included as the darkened histogram.

samples. According to the literature, the M-R should be bimodal, with a peak at $\sim 0.55 M_{\odot}$, and a second one at $\sim 0.8\ M_{\odot}$. This distribution is caused by the different evolutionary tracks a MS star follows depending on its mass: larger WDs are expected to have O-Ne-Mg cores, that are formed by the higher mass end of the progenitors. There is also a small population of WDs we expect to see at $\sim 1\ M_{\odot}$, formed by the merger of WDs; this is however, the least common population. In the golden sample case, for both M-R relations we see an almost uniform distribution, which is not representative of the full sample. These results can be, as we mentioned, partly due to the fact that the stars that obtain good fits are able to do so because of their placement in the parameter space, but it is also possible that the algorithm is predicting a golden sample with uniform distribution because it does not know that there is a small dearth of WDs between the two main peaks (see Figure 9 of Sahu et al. (2023)), since it is being trained with a random uniform distribution of spectra.

Thus, we can confidently present the results obtained from the full sample, which can be summarized as:

- XGBoost provides more precise results compared to Random Forest. This work demonstrates the feasibility of using machine learning in the spectroscopy of white dwarfs, with room for

improvement, particularly by increasing the amount of data used during training.

- Incorporating synthetic models helps address several machine learning challenges, such as the lack of representative spectra.
- *Gaia* distances can serve as independent measurements to highlight potential discrepancies in cooling models. This contrasts with Bayesian analysis, where the directionality of the optimal parameters tends to be diluted.
- Highly reddened white dwarfs are typically located near the galactic plane and at significant distances. For these stars to be detectable, they must be both large in size and sufficiently hot.
- Cooling models require revision, especially for systems with low $\log(g)$, which is largely affected by observational bias, which is also a known issue by other works. And thus this region is very much sensitive to the physics implemented in the models.
- The obtained mass distribution is representative of the expected results for a population of WDs, with the mean value for both M-R relations being similar to the results of other works.

Chapter 4

Discussion

In this Chapter we discuss the main problems we encountered during the development of this work, including limitations due to the nature of the observations, and the implementation of the machine learning algorithms.

4.1 Observational bias

Due to the faint nature of these objects, we have a strong observational bias. This means that their intrinsic properties limit our capabilities of observing them. Assuming a uniform integration time of the same instrument for all distances, we can detect cool white dwarfs that are relatively close. As a result, most of the cool white dwarfs, located in the bottom-right region of the HR diagram (the same diagram as Fig. 1.1 in Section 1.1, with different axis), and are at closer distances, as indicated by the color bar in the figure. In contrast, as we move towards larger distances, we can only detect the brightest and large in size stars. Thus in Fig. 4.1) we see that the brightest WDs (those at the upper left corner of the WD sequence) are also the ones that are the furthest (magenta points). As a result, the observational bias becomes evident along the white dwarf sequence, transitioning from hot, distant white dwarfs to cool, nearby ones. This trend is reflected in the color gradient of the sequence shown in Fig. 4.1. The observational bias identified in our work (see Section 4.1) strongly supports conducting analyses using volume-limited samples, an approach that has been explored in previous studies. For example, McCleery et al. (2020) assembled a 40 [pc] sample, with some extensions reaching up to 100 [pc] within the SDSS footprint (Kilic et al., 2020). In this study, an ideal volume limit is around 50 [pc] assuming the average exposures times for *HST* spectroscopy.

In this same figure, we observe that the white dwarf sequence splits into two distinct groups. While most white dwarfs align along a well-defined diagonal, a second group is located below this sequence. The white dwarfs in this lower group must have smaller radii to occupy this region

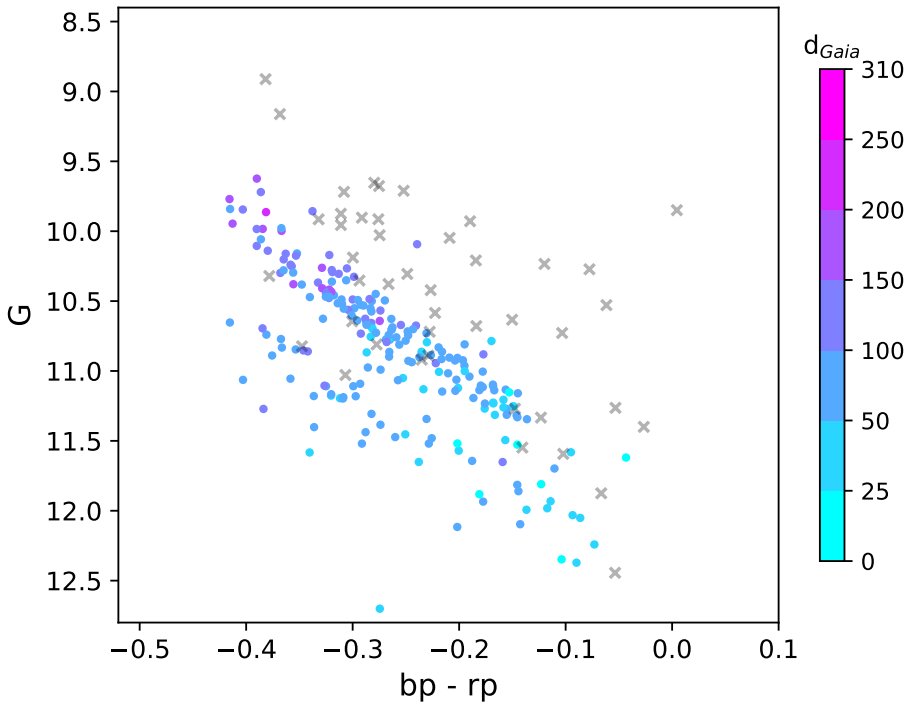


Figure 4.1: Color-Magnitude diagram for the observed objects. Absolute magnitude in the G-band (G) was calculated from the apparent magnitude (g) in the same filter; aparent magnitudes g, bp and rp where obtained from *Gaia*. Objects classified as mass outliers (binaries, WDs with disks, and others) by Sahu et al. (2023) are marked as gray crosses. The coloured dots show the distance of each object, according to *Gaia*. From here we see that the brightest WDs (upper left corner) also have the largest distances.

in the HR diagram. Consequently, as dictated by the MR relation, they are likely the most massive white dwarfs, with $\log(g)$ values generally exceeding 8.5 dex. These massive white dwarfs likely contribute to the asymmetry observed in the mass distribution of white dwarfs shown in Figure 3.12.

All the statements mentioned above introduce problems when we try to implement an ML algorithm that can predict atmospheric parameters because:

- On one hand, we were unable to confidently train the ML algorithm with observational spectroscopy, not only due to the limited number of white dwarfs in the sample (only 262), but also because certain regions of the parameter cube ($T_{\rm eff}$, $\log(g)$, and Distance) lack a balanced and representative distribution of white dwarfs.

- On the other hand, all of these objects, as we have shown in previous sections (see Section 3.2.1, Fig. 3.6) are located in the problematic zone in the parameter space, which could be the reason why we struggle so much in constraining the M-R relations for this region, since it has the hardest to obtain and probably more biased observations.

4.2 Interstellar Lyman-alpha absorption

In regions of high reddening, it has been found that there exists significant amount of neutral hydrogen atoms in the interstellar medium (ISM). These atoms can absorb ultraviolet radiation at the Lyman α wavelength—approximately 1216 Å (Jenkins, 1971).

The work of Diplas and Savage (1994) shows that there is a straight correlation of between neutral hydrogen and reddening. Thus, we expect the most reddened objects to also present interstellar Ly$-\alpha$ absorption. If this is not accounted for in the models, the $T_{\rm eff}$ of the WD will be underestimated. As an extreme example, the study by Werner et al. (2020) showed that accounting for interstellar hydrogen absorption when modelling the UV spectrum of the white dwarf in the binary system UCAC2 46706450 causes the effective temperature to increase significantly, from 24000 [K]to $T_{\rm eff} = 105000 \pm 5000$ K. The primary impact on temperature arises from the fact that, in a typical, unabsorbed white dwarf spectrum, the core at the central wavelength of $\sim$ 1216 Å is expected to be steep. The shape of this core is predominantly determined by $T_{\rm eff}$. However, when additional absorption from the interstellar medium is present, the core becomes saturated and takes on a square-like shape. The stronger the interstellar absorption, the more the wings of this saturation extend. Our conservative mask around Ly$-\alpha$ mitigates this effect, as the wings of the atmospheric Ly$-\alpha$ absorption are relatively far from the central core and remain unaffected by this saturation. Since colder WDs will have broader cores, if we do not consider this external effect, it is very likely that the fit will predict a lower temperature, because the square core appears broadened. We showed in Fig. 3.6 that there are 6 highly reddened stars. Sahu et al. (2023) showed that one of this objects has considerable interstellar Ly-α absorption (see their Figure 14).

4.3 Literature comparison: Sahu et al. 2023

Several studies have determined atmospheric parameters for large samples of white dwarfs. Among them is the work by Sahu et al. (2023), which we used to select our targets and obtain reference values for constructing our training and validation sets. Building on this foundation, it is natural to compare our ML predictions with the results from this prior study, as we do in this section. Sahu et al. also extensively compares to other authors, who used different methods: Liebert et al. (2005); Koester et al. (2009); Gianninas et al. (2011) use optical spectra of the same objects to determine the atmospheric parameter by fitting the Balmer line profiles (see Fig. 1.3); Kilic et al. (2020); Dufour et al. (2017); Gentile Fusillo et al. (2021); Jiménez-Esteban et al. (2023) use photometric data, along the *Gaia* parallaxes. Sahu et al. reiterates that some of their results use the same objects of these previous studies, and in some cases even the parameters used are fixed to the results of previous works, which is why they insist that the uncertainty values published in

the literature should not be taken as realistic.

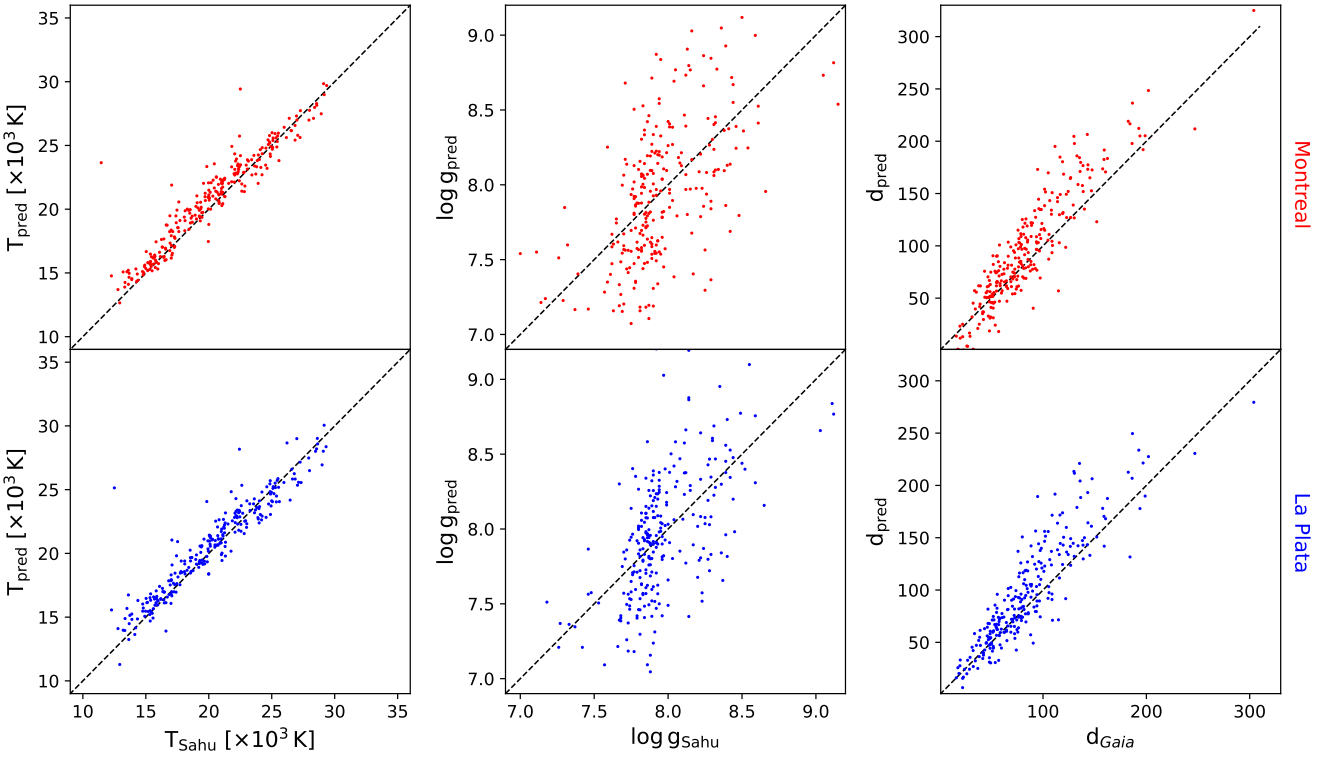


Figure 4.2: Predicted $T_{\rm eff}$ (left panel) and $\log(g)$ (centre panel), compared to the Sahu et al. (2023) calculated values. The right panel shows the predicted distance compared to the real *Gaia* value. The upper row corresponds to the Montreal (red) M-R relation, while the bottom row corresponds to La Plata.

In Table 4.1 we include the metrics obtained when predicting $T_{\rm eff}$, $\log(g)$ and distance for the observed spectra with the optimized model, for each M-R relation, considering the results from Sahu et al. (2023) as the correct ones for $T_{\rm eff}$ and $\log(g)$. Because this data is noisy, and is contaminated by emission and absorption features, we expected lower R^2 scores for the test set than for the validation set, that had obtained a score of 0.993 (see Section 2.3.2). We see that for the La Plata M-R this score is lower than for Montreal. La Plata has lower MAE values for $\log(g)$ and distance, while Montreal has lower RMSE values for $T_{\rm eff}$, $\log(g)$ and distance, however these differences are not large enough to consider one M-R as better.

In Fig. 4.2 we show the predicted and expected (that is, the values calculated by Sahu et al. 2023) for the observational spectra (black dots). Here we see that the best results are for $T_{\rm eff}$ (left panels), where we have the least dispersion. There are a few points where temperature is strongly overestimated, but this is due to the architecture of the models: they will always give us an output, and if it cannot find an adequate prediction, it will do its best, and sometimes the best is really bad. For $\log(g)$ (centre panels), we see that there is much more dispersion, and as expected, the

ML	M-R	R^2	RMSE			MAE			Sahu et al. 2023	
		Test	$T_{\rm eff}$ [K]	$\log(g)$ [dex]	d [pc]	$T_{\rm eff}$ [K]	$\log(g)$ [dex]	d [pc]	$T_{\rm eff}$ [K]	$\log(g)$ [dex]
XGB	M	0.877	1397	0.40	26.77	907	0.31	19.71	50	0.02
	LP	0.867	1455	0.37	29.03	927	0.28	19.35		

Table 4.1: R^2 score, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for the optimized models, for the test set (observed spectra) comparing with the expected value, that is the results from Sahu et al. (2023). Here we see that the metrics are similar for both M-R relations. Focusing on the MAE values, we see that for La Plata the error is lower for $\log(g)$ and distance, while $T_{\rm eff}$ is lower for Montreal. The last column corresponds to the errors reported by Sahu et al. (2023), who used χ^2 minimization to fit the models.

highest values that are outside of the training set range are also underestimated, which is what we expected from the validation set results. For distance (right panels), we see that both models overestimate the value of the largest distances. This will be analysed in the following sections. Apart from this, we can also see that the results for this target have a lower dispersion at smaller distances.

Because of the large dispersion of $\log(g)$ values we can not interpret much of where the algorithm is failing the most, but since both M-R relations behave similarly, it is very likely that is an error that has seeped through the training process, since we also saw that this variable had the largest dispersion.

Temperatures are mostly over estimated up to 20000 K, where we see that they start to get underestimated for both M-R relations. Sahu et al. (2023) compared their work with multiple other spectroscopic and photometric studies. One of these is Gianninas et al. (2011), who performed an analysis of the atmospheric parameteris of bright DA WDs using spectroscopy in the optical range. Sahu et al. finds temperatures consistenly lower than Gianninas et al., which means that we could be finding closer results to the latter. This is, however, subject to the performance of our ML algorithm, and might not stand after further analysis and improvement, which we will discuss in the following sections.

The errors reported in Sahu et al. (2023) are included in the last column of Table 4.1. As we have mentioned, our errors are significantly higher, and this result shows that there is a need for improvement before this technique can be confidently used to analyse spectroscopic data.

4.4 Attempting a Second Feature Sampling

An obvious question is how much our results would change if the feature sampling were different. To explore this, we defined an alternative set of wavelength points, detailed in Table 4.2. This

sampling differs from the previous one both in resolution, since we select a broader region with high resolution to cover the wings of the Ly$-\alpha$ absorption feature, and in wavelength ranges. Given that most of the WDs showed metal absorption lines (see Section 2.1.2.1), and a second airglow feature (see Section 2.1.2.2), we felt that the results could improve if we removed these zones, since they were the zones where the different spectra varied without dependance on the three variables we were predicting.

	Wavelength range [Å]	$\Delta\lambda$ [Å]
Sampling 1	1130 to 1190	5
	1190 to ∼1212.07	0.5
	∼1219.27 to 1240	0.5
	1240 to 1450	20
Sampling 2	1130 to 1150	5
	1160 to ∼1212.07	0.3
	∼1219.27 to 1255	0.3
	1280 to 1300	3
	1340 to 1450	5

Table 4.2: Wavelength ranges for the second sampling (Samplig 2) used to bin the fluxes in the training, validation and test sets for our ML models. For comparison purposes also shown is the first sampling (Sampling 1).

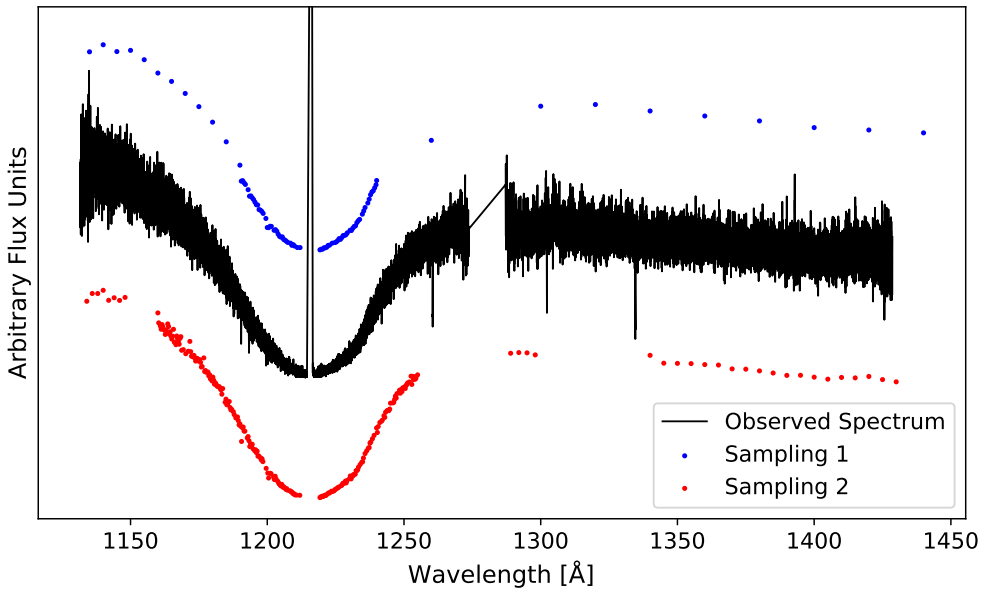


Figure 4.3: Same as Figure 2.7, but the second sampling is shown in red dots.

In Table 4.3 we see that the training score for all eight models is 0.999, while the validation score

is slightly higher for Sampling 1 (see Fig. 4.3). However, since the difference is not significant, we went ahead with testing all these models with the observational data, to determine which sampling gave us a better performance. Since training and validation were done on the same synthetic data with different wavelength samplings, we expected the results to be similar; for observational data however, we wanted to see which of the two samplings was able to obtain a better performance, since the observations did contain effects like the metal absorption lines and the airglow features that we did not account for in the training process.

	M-R	ML	R^2		RMSE			MAE		
			Train	Validation	$T_{\rm eff}$ [K]	$\log(g)$ [dex]	d [pc]	$T_{\rm eff}$ [K]	$\log(g)$ [dex]	d [pc]
S1	M	RF	0.999	0.994	549	0.25	23.21	404	0.19	17.30
		XGB	0.999	0.993	607	0.21	19.95	451	0.16	14.84
	LP	RF	0.999	0.994	563	0.26	24.42	413	0.20	18.50
		XGB	0.999	0.992	616	0.22	20.16	461	0.16	14.62
S2	M	RF	0.999	0.993	607	0.26	23.56	423	0.19	17.74
		XGB	0.999	0.992	619	0.20	16.99	447	0.15	12.28
	LP	RF	0.999	0.993	599	0.28	24.83	421	0.19	18.90
		XGB	0.999	0.992	647	0.21	19.03	457	16.02	14.09

Table 4.3: Same as Table 2.3. For comparison purposes also Sampling 1 (S1) is shown. The table lists R^2 score, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for the new sampling (Sampling 2, S2) M-R relation Montreal, M and La Plata, LP) and ML algorithm (RF and XGB), obtained in the training and validation process, that is, where the algorithms were only applied to synthetic data. The best performing model overall is RF with S1 and the Montreal M-R relation.

We input the observational data to the non-optimized models, and compared the performance. Those metrics are included in Table 4.4. Here we see that the new sampling has far worse metrics than Sampling 1, with R^2 scores for RF that are negative for both M-R relations, which indicates that the predictions are not accurate. In Fig. 4.4 we include the comparison of the predictions of $T_{\rm eff}$ and $\log(g)$, compared to the results from Sahu et al. (2023), and the distance compared to the *Gaia* values. Here it is clear that this sampling does not give good results.

It is still unclear why this happened, and is part of the reason we believe there is still work to do in terms of how the sampling is made for this type of ML application, since its clear that what we thought would be a better sampling, since it covered more of the Ly$-\alpha$ absorption which is fundamental in determining $T_{\rm eff}$ and $\log(g)$, actually worsened the results. One speculation is that the addition of more points (Sampling 1 had 98, Sampling 2 has 256 points) caused the algorithm to place less *weight* on the Ly$-\alpha$ feature that is determined by $T_{\rm eff}$, $\log(g)$.

	M-R	ML	R^2	RMSE			MAE		
			Test	$T_{\rm eff}$ [K]	log(g) [dex]	d [pc]	$T_{\rm eff}$ [K]	log(g) [dex]	d [pc]
S1	M	RF	0.905	1227	0.26	25.50	818	0.19	14.90
		XGB	0.858	1502	0.40	27.87	987	0.30	18.06
	LP	RF	0.907	1212	0.24	25.70	816	0.17	15.53
		XGB	0.842	1582	0.41	29.97	1017	0.31	21.32
S2	M	RF	-0.258	4472	0.50	50.64	3937	0.40	40.27
		XGB	-0.277	619	0.20	16.99	447	0.15	12.28
	LP	RF	-0.222	440	0.46	49.43	3863	0.38	39.96
		XGB	0.140	3696	0.59	55.12	3106	0.49	45.15

Table 4.4: R^2 score, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for the eight combinations of sampling (Sampling 1, S1, and Sampling 2, S2) M-R relation (Montreal, M, and La Plata, LP) and ML algorithm (RF and XGB), obtained with the test set, that is, where the algorithms were applied to observational data. Note that these metrics are calculated for all parameters considering the results from Sahu et al. (2023) as the correct ones. Here we see that Sampling 2 (S2) does not perform well for any models, and has consistently the worst R^2 independent from the M-R or algorithm used.

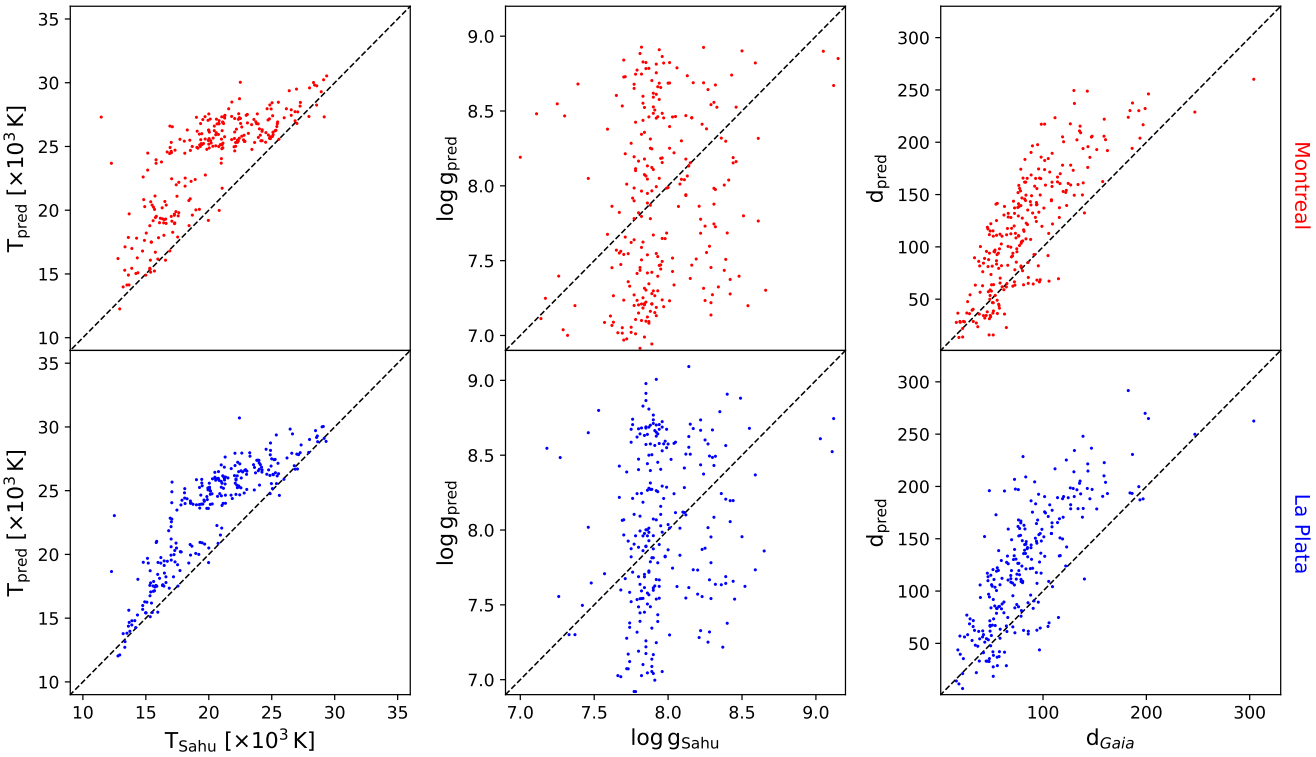


Figure 4.4: Same as Fig. 4.2, but for Sampling 2.

4.5 Future work

In this section we will describe the main issues and the improvements we could make on our current results.

4.5.1 More Feature Samplings

From our results, it is clear that there is a need to keep exploring other kinds of feature sampling. There are unsupervised ML algorithms, like `selectKbest` from `sklearn` that could help us select features without strong pre-processing, like different resolution wavelength binning as we have done here. This algorithm specifically performs clustering over the data; since it is unsupervised it does not require a training set. It is based on the assumption that by selecting some features we could be able to separate the data by performing a reduction of the dimensionality, that is, linearly combining some of the features to reduce the overall number. By using unsupervised techniques, the selection would be done entirely by the ML, and that selection could then be fed as a training set to the algorithm that predicts the atmospheric parameters.

Other than that, we were severely limited by the computational costs of building more complex models. Because of this, we would like to explore in the future a uniform binning of higher resolution. In this work, both samplings tested prioritized the Ly$-\alpha$ absorption, which meant the sampling in the continuum had to be more sparse; we believe a more accurate representation of the spectra would entail a high-resolution and uniform selection of points, since that is the way that the real, raw data, is obtained.

We also believe that this is a variable we need to look into further, because of the great difference we obtain when we compare the two different samplings.

4.5.2 Size of training and validation sets

While we argued that from the learning curve shown in Section 1.4.1.1 it is apparent that the algorithm has sufficient points, which is also visible in the validation R^2 scores from the optimized model, we believe that in order to reduce the error we also need to increase the number of points in the hyperparameter space. In this work all algorithms were trained with the same 4000 combinations of T_{eff}, $\log(g)$ and distance, but in order to obtain a resolution similar to the error of traditional methods, the number could increase to up to 3 million synthetic spectra. Now, this is an extremely high number, that would be impossible to implement in an efficient way, but we do believe that it is necessary to increase the training set size to over 10000 spectra, to obtain better results. For example, Yang et al. (2020) employs $\sim$ 6000 training examples, while Badenas-Agusti et al. (2024) use over 20000; both of them, however, use neural networks, which are much more

efficient at handling large amounts of data.

This last reason is why we believe that while there might be room for improvement, it is very likely that the metrics obtained with decision tree ensemble methods, like XGBoost, will not reach the error values of manual fits of SEDs, because of their own capacity of processing the data. Nonetheless, it is a test we will perform in the future, in hopes to refine our results. This will be done using the cluster of the Simulations group at Universidad Técnica Federico Santa María (Casa Central).

If we were unable to improve the computational capacity, we could take into account the reduction placed by a volume limited sample, since it will reduce the distance range, which will increase the resolution of points for T_{eff} and $\log(g)$. Once these changes are applied, we believe this work could be published.

4.5.3 Exploring other ML algorithms

Finally, we also believe it is necessary to explore other ML algorithms, specially neural networks in the future. Most spectroscopic data analysis done with ML uses neural networks because of the benefits it has in managing large amounts of data. The obvious counterpoints are that they are not interpretable due to their *black box* architecture, they are much more complex and implementation and optimization is much more sophisticated, which makes them not only more difficult to manage, but also very computationally expensive. Because of this we believe it is more appropriate to exhaust all resources with the algorithms we have used so far, since they have shown good performance.

Chapter 5

Conclusion

White dwarfs are extremely important in the understanding of the fate of most stars in our galaxy, the end of planetary systems, and even the age of the Milky Way. Because of this, we have been studying their properties for decades; this has not been an easy task, since they are very faint objects. Upcoming surveys will provide us with unimaginable amounts of data, that we need to process in an efficient way. Because of this, we have seen an uprise in the use of machine learning techniques.

In this work we successfully used decision tree-based ensemble algorithms to predict $T_{\rm eff}$, $\log(g)$ and distance for 262 WDs using UV spectroscopy from *HST*. To do this, we used a synthetic DA WD model grid, with which we obtained validation MAE on average of $\sim$ 403 [K] for $T_{\rm eff}$, $\sim$ 0.14 [dex] for $\log(g)$, and $\sim$ 13 [pc] for the distance. For the observational data, we could only determine the performance based on the model-independent variable: the *Gaia* distance. In this case, we found an average MAE of 19.53 [pc] for both M-R relations. The average difference between Montreal and La Plata is $\sim 145 \pm 738$ [K] for $T_{\rm eff}$, $\sim -0.019 \pm 0.21$ [dex] for $\log(g)$, and $\sim -0.46 \pm 15.72$ [pc] for the distance. Even with a model that is still sensitive to noise and has a considerable error, we still recover slight differences in the results for the M-R relations, which proves the sensibility of ML models to the training data, and the *inherent* biases these contain.

During the development of this project we also confirmed the need for a volume limited sample, because of the amount of error we obtained in the ML predictions. By limiting the maximum distance of the objects we can have a more representative sample in the the $T_{\rm eff}$-$\log(g)$ parameter space, which in turn can help us reduce the parameter ranges the training set has to cover, which would be helpful when improving the resolution of the *cube*, since we believe this also needed to reduce the error.

From our predictions we were able to obtain several results by comparing the M-R relations:

1. We need to improve the models for the lowest mass WDs, that is, those with the smallest $\log(g)$ values, since this range of $\log(g)$ proved to contain the most error in both M-R

relations.

2. In the range of the parameter space where the M-R relations give good results (that is, for $\log(g)$ higher than ~ 8.0 [dex] and temperatures lower than 25000 K) the algorithm performs well, and we obtain similar results to previous works (Sahu et al., 2023; Gianninas et al., 2011; Tremblay and Bergeron, 2009), and we are able to obtain a golden sample of 18 objects for Montreal and 21 for La Plata, using the difference between the *Gaia* distance and our distance prediction as a constraint, as well as the χ^2_r of the spectrum to determine which objects have good fits. Not only are we able to model these spectra somewhat accurately, but we also reproduce the results from the full sample. As the mass distribution of WDs, which proves that even with the amount of error of our algorithms, we are obtaining a good performance.

3. The fact that we are able to introduce the *Gaia* distance as a variable that is predicted allows us to also preserve crucial information contained in the spectra, since we do not need to normalize it. By removing this step from the pre-processing of the data, we are able to expand the capabilities of these algorithms, and we are also able to add a constraint that is not dependent in the M-R relation used or in the existence of previous labels calculated in the traditional way. Even though the selected targets for this work were based on the work of a previous author (Saha et al., 2023), we could have just evaluated the performance of the algorithm -not even the M-R relations- just based on the predictions, since we can model a synthetic atmosphere based on those values.

4. In order to mitigate observational bias that could skew the results, further analyses should be conducted on volume-limited samples. In the case of the *HST* sample, a 50 [pc] volume would provide a representative sample of white dwarfs.

5. Considering the errors associated with different methods in the literature compared to the larger errors in our ML approach, it is still preferable to fit spectroscopy using conventional methods when the sample size is relatively small (fewer than 1000 objects).

Finally, the computational cost of this algorithms is much lower than the traditional methods. Obtaining a synthetic atmosphere model for a single object can take upwards of hours, since we need to numerically solve the stellar structure. By using a grid and a linear multi-dimensional interpolator we can speed up the process significantly, because we do not need to create a specific model for each object, but rather a sample of models that have enough resolution in the parameter space so that the model can be trained to an optimal complexity and the lowest error possible, without over-fitting. More so, since these algorithms are not the most complex ML models, they do not take more than 2 minutes to be trained and validated, and the observational data predictions

take seconds in a Intel® Core™ i5-10300H CPU with 8 GB of RAM. The most time consuming process is the optimization, that takes around 30 minutes with the aforementioned hyperparameter grid. Nonetheless, this is extremely fast, which is why we believe with some improvements we could develop a very useful tool to predict the atmospheric parameters of DA WDs in an efficient way.

Bibliography

Acquaviva, V. (2023). *Machine Learning for Physics and Astronomy*. Princeton University Press.

Althaus, L. G., Miller Bertolami, M. M., and Córsico, A. H. (2013). New evolutionary sequences for extremely low-mass white dwarfs. Homogeneous mass and age determinations and asteroseismic prospects. *A&A*, 557:A19.

Althaus, L. G., Miller Bertolami, M. M., Córsico, A. H., García-Berro, E., and Gil-Pons, P. (2005). The formation of da white dwarfs with thin hydrogen envelopes. *A&A*, 440(1):L1–L4.

Ambrosino, F. (2020). White Dwarf mass-radius relation. *arXiv e-prints*, page arXiv:2012.01242.

Ardila, D. R., Shkolnik, E., Ziemer, J., Swain, M., Owen, J. E., Line, M., Loyd, R. O. P., Sellar, R. G., Barman, T., Dressing, C., Frazier, W., Jewell, A. D., Kinsey, R. J., Liebe, C. C., Lothringer, J. D., Martinez-Sierra, L. M., McGuire, J., Meadows, V., Murray-Clay, R., Nikzad, S., Peacock, S., Schlichting, H., Sing, D., Stevenson, K., and Wu, Y.-H. (2022). The UV-SCOPE mission: ultraviolet spectroscopic characterization of planets and their environments. In den Herder, J.-W. A., Nikzad, S., and Nakazawa, K., editors, *Space Telescopes and Instrumentation 2022: Ultraviolet to Gamma Ray*, volume 12181 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 1218104.

Badenas-Agusti, M., Viaña, J., Vanderburg, A., Blouin, S., Dufour, P., Xu, S., and Sha, L. (2024). cecilia: a machine learning-based pipeline for measuring metal abundances of helium-rich polluted white dwarfs. *MNRAS*, 529(2):1688–1714.

Barstow, M. A., Holberg, J. B., Hubeny, I., Good, S. A., Levan, A. J., and Meru, F. (2001). A comparison of DA white dwarf temperatures and gravities from Lyman and Balmer line studies. *MNRAS*, 328(1):211–222.

Bédard, A., Bergeron, P., Brassard, P., and Fontaine, G. (2020). On the Spectral Evolution of Hot White Dwarf Stars. I. A Detailed Model Atmosphere Analysis of Hot White Dwarfs from SDSS DR12. *ApJ*, 901(2):93.

Bourrier, V., Ehrenreich, D., Lecavelier des Etangs, A., Louden, T., Wheatley, P. J., Wyttenbach, A., Vidal-Madjar, A., Lavie, B., Pepe, F., and Udry, S. (2018). High-energy environment of super-earth 55 cancri e - i. far-uv chromospheric variability as a possible tracer of planet-induced coronal rain. *A&A*, 615:A117.

Camisassa, M. E., Althaus, L. G., Córsico, A. H., De Gerónimo, F. C., Miller Bertolami, M. M., Novarino, M. L., Rohrmann, R. D., Wachlin, F. C., and García-Berro, E. (2019). The evolution of ultra-massive white dwarfs. *A&A*, 625:A87.

Camisassa, M. E., Althaus, L. G., Córsico, A. H., Vinyoles, N., Serenelli, A. M., Isern, J., Bertolami, M. M. M., and García–Berro, E. (2016). The effect of 22ne diffusion in the evolution and pulsational properties of white dwarfs with solar metallicity progenitors. *ApJ*, 823(2):158.

Capitanio, L., Lallement, R., Vergely, J. L., Elyajouri, M., and Monreal-Ibero, A. (2017). Three-dimensional mapping of the local interstellar medium with composite data. *A&A*, 606:A65.

Carvalho, A., MarinhoJr, R. M., and Malheiro, M. (2016). Mass-radius relation for white dwarfs models at zero temperature. *Journal of Physics: Conference Series*, 706(5):052016.

Chandrasekhar, S. (1931). The Maximum Mass of Ideal White Dwarfs. *ApJ*, 74:81.

Diplas, A. and Savage, B. D. (1994). An IUE Survey of Interstellar H i Lyman-Alpha Absorption. II. Interpretations. *ApJ*, 427:274.

Dufour, P., Blouin, S., Coutu, S., Fortin-Archambault, M., Thibeault, C., Bergeron, P., and Fontaine, G. (2017). The Montreal White Dwarf Database: A Tool for the Community. In Tremblay, P. E., Gaensicke, B., and Marsh, T., editors, *20th European White Dwarf Workshop*, volume 509 of *Astronomical Society of the Pacific Conference Series*, page 3.

Farihi, J., Hermes, J. J., Marsh, T. R., Mustill, A. J., Wyatt, M. C., Guidry, J. A., Wilson, T. G., Redfield, S., Izquierdo, P., Toloza, O., Gänsicke, B. T., Aungwerojwit, A., Kaewmanee, C., Dhillon, V. S., and Swan, A. (2022). Relentless and complex transits from a planetesimal debris disc. *MNRAS*, 511(2):1647–1666.

Fitzpatrick, E. L. (1999). Correcting for the Effects of Interstellar Extinction. *PASP*, 111(755):63–75.

Fitzpatrick, E. L. and Massa, D. (1990). An Analysis of the Shapes of Ultraviolet Extinction Curves. III. an Atlas of Ultraviolet Extinction Curves. *ApJS*, 72:163.

Förster, F., Cabrera-Vives, G., Castillo-Navarrete, E., Estévez, P. A., Sánchez-Sáez, P., Arredondo, J., Bauer, F. E., Carrasco-Davis, R., Catelan, M., Elorrieta, F., Eyheramendy, S., Huijse, P., Pignata, G., Reyes, E., Reyes, I., Rodríguez-Mancini, D., Ruz-Mieres, D., Valenzuela, C., Álvarez-Maldonado, I., Astorga, N., Borissova, J., Clocchiatti, A., De Cicco, D., Donoso-Oliva, C., Hernández-García, L., Graham, M. J., Jordán, A., Kurtev, R., Mahabal, A., Maureira, J. C., Muñoz-Arancibia, A., Molina-Ferreiro, R., Moya, A., Palma, W., Pérez-Carrasco, M., Protopapas, P., Romero, M., Sabatini-Gacitua, L., Sánchez, A., San Martín, J., Sepúlveda-Cobo, C., Vera, E., and Vergara, J. R. (2021). The Automatic Learning for the Rapid Classification of Events (ALeRCE) Alert Broker. *AJ*, 161(5):242.

Fortmann-Roe, S. (2012). Understanding the bias-variance tradeoff.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189 – 1232.

Gentile Fusillo, N. P., Tremblay, P. E., Cukanovaite, E., Vorontseva, A., Lallement, R., Hollands, M., Gansicke, B. T., Burdge, K. B., McCleery, J., and Jordan, S. (2021). VizieR Online Data Catalog: Catalogue of white dwarfs in Gaia EDR3 (Gentile+, 2021).

Gianninas, A., Bergeron, P., and Ruiz, M. T. (2011). A spectroscopic survey and analysis of bright, hydrogen-rich white dwarfs*. *The ApJ*, 743(2):138.

Hastie, T., Friedman, J., and Tisbshirani, R. (2017). *The elements of Statistical Learning: Data Mining, Inference, and prediction*. Springer.

Jenkins, E. B. (1971). A Closer Look at Interstellar Lyman-Alpha Absorption. *ApJ*, 169:25.

Jermyn, A. S., Bauer, E. B., Schwab, J., Farmer, R., Ball, W. H., Bellinger, E. P., Dotter, A., Joyce, M., Marchant, P., Mombarg, J. S. G., Wolf, W. M., Sunny Wong, T. L., Cinquegrana, G. C., Farrell, E., Smolec, R., Thoul, A., Cantiello, M., Herwig, F., Toloza, O., Bildsten, L., Townsend, R. H. D., and Timmes, F. X. (2023). Modules for Experiments in Stellar Astrophysics (MESA): Time-dependent Convection, Energy Conservation, Automatic Differentiation, and Infrastructure. *The Astrophysical Journal Supplement Series*, 265(1):15.

Jiménez-Esteban, F. M., Torres, S., Rebassa-Mansergas, A., Cruz, P., Murillo-Ojeda, R., Solano, E., Rodrigo, C., and Camisassa, M. E. (2023). Spectral classification of the 100 pc white dwarf population from Gaia-DR3 and the virtual observatory. *MNRAS*, 518(4):5106–5122.

Kilic, M., Bergeron, P., Kosakowski, A., Brown, W. R., Agüeros, M. A., and Blouin, S. (2020). The 100 pc White Dwarf Sample in the SDSS Footprint. *ApJ*, 898(1):84.

Koester, D. (2010). White dwarf spectra and atmosphere models. *Mem. Societa Astronomica Italiana*, 81:921–931.

Koester, D., Gänsicke, B. T., and Farihi, J. (2014). The frequency of planetary debris around young white dwarfs. *A&A*, 566:A34.

Koester, D., Voss, B., Napiwotzki, R., Christlieb, N., Homeier, D., Lisker, T., Reimers, D., and Heber, U. (2009). High-resolution UVES/VLT spectra of white dwarfs observed for the ESO SN Ia Progenitor Survey. III. DA white dwarfs. *A&A*, 505(1):441–462.

Kulkarni, S. R., Harrison, F. A., Grefenstette, B. W., Earnshaw, H. P., Andreoni, I., Berg, D. A., Bloom, J. S., Cenko, S. B., Chornock, R., Christiansen, J. L., Coughlin, M. W., Wuollet Criswell, A., Darvish, B., Das, K. K., De, K., Dessart, L., Dixon, D., Dorsman, B., El-Badry, K., Evans, C., Ford, K. E. S., Fremling, C., Gansicke, B. T., Gezari, S., Goetberg, Y., Green, G. M., Graham, M. J., Heida, M., Ho, A. Y. Q., Jaodand, A. D., Johns-Krull, C. M., Kasliwal, M. M., Lazzarini, M., Lu, W., Margutti, R., Martin, D. C., Masters, D. C., McKernan, B., Naze, Y., Nissanke, S. M., Parazin, B., Perley, D. A., Phinney, E. S., Piro, A. L., Raaijmakers, G., Rauw, G., Rodriguez, A. C., Sana, H., Senchyna, P., Singer, L. P., Spake, J. J., Stassun, K. G., Stern, D., Teplitz, H. I., Weisz, D. R., and Yao, Y. (2021). Science with the Ultraviolet Explorer (UVEX). *arXiv e-prints*, page arXiv:2111.15608.

Lallement, R., Vergely, J. L., Valette, B., Puspitarini, L., Eyer, L., and Casagrande, L. (2014). 3D maps of the local ISM from inversion of individual color excess measurements. *A&A*, 561:A91.

Lallement, R., Capitanio, L., Ruiz-Dern, L., Danielski, C., Babusiaux, C., Vergely, L., Elyajouri, M., Arenou, F., and Leclerc, N. (2018). Three-dimensional maps of interstellar dust in the local arm: using gaia, 2mass, and apogee-dr14. *A&A*, 616:A132.

Liebert, J., Wickramsinghe, D. T., Schmidt, G. D., Silvestri, N. M., Hawley, S. L., Szkody, P., Ferrario, L., Webbink, R. F., Oswalt, T. D., Smith, J. A., and Lemagie, M. P. (2005). Where are the magnetic white dwarfs with detached, nondegenerate companions? *The Astronomical Journal*, 129(5):2376.

McCleery, J., Tremblay, P.-E., Gentile Fusillo, N. P., Hollands, M. A., Gänsicke, B. T., Izquierdo, P., Toonen, S., Cunningham, T., and Rebassa-Mansergas, A. (2020). Gaia white dwarfs within 40 pc II: the volume-limited Northern hemisphere sample. *MNRAS*, 499(2):1890–1908.

Parsons, S. G., Gänsicke, B. T., Marsh, T. R., Ashley, R. P., Bours, M. C. P., Breedt, E., Burleigh, M. R., Copperwheat, C. M., Dhillon, V. S., Green, M., Hardy, L. K., Hermes, J. J., Irawati, P., Kerry, P., Littlefair, S. P., McAllister, M. J., Rattanasoon, S., Rebassa-Mansergas, A., Sahman,

D. I., and Schreiber, M. R. (2017). Testing the white dwarf mass–radius relationship with eclipsing binaries. *MNRAS*, 470(4):4473–4492.

Paxton, B., Bildsten, L., Dotter, A., Herwig, F., Lesaffre, P., and Timmes, F. (2011). Modules for Experiments in Stellar Astrophysics (MESA). *The Astrophysical Journal Supplement Series*, 192:3.

Paxton, B., Cantiello, M., Arras, P., Bildsten, L., Brown, E. F., Dotter, A., Mankovich, C., Montgomery, M. H., Stello, D., Timmes, F. X., and Townsend, R. (2013). Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rotation, and Massive Stars. *The Astrophysical Journal Supplement Series*, 208:4.

Paxton, B., Marchant, P., Schwab, J., Bauer, E. B., Bildsten, L., Cantiello, M., Dessart, L., Farmer, R., Hu, H., Langer, N., Townsend, R. H. D., Townsley, D. M., and Timmes, F. X. (2015). Modules for Experiments in Stellar Astrophysics (MESA): Binaries, Pulsations, and Explosions. *The Astrophysical Journal Supplement Series*, 220:15.

Paxton, B., Schwab, J., Bauer, E. B., Bildsten, L., Blinnikov, S., Duffell, P., Farmer, R., Goldberg, J. A., Marchant, P., Sorokina, E., Thoul, A., Townsend, R. H. D., and Timmes, F. X. (2018). Modules for Experiments in Stellar Astrophysics (MESA): Convective Boundaries, Element Diffusion, and Massive Star Explosions. *The Astrophysical Journal Supplement Series*, 234:34.

Paxton, B., Smolec, R., Schwab, J., Gautschy, A., Bildsten, L., Cantiello, M., Dotter, A., Farmer, R., Goldberg, J. A., Jermyn, A. S., Kanbur, S. M., Marchant, P., Thoul, A., Townsend, R. H. D., Wolf, W. M., Zhang, M., and Timmes, F. X. (2019). Modules for Experiments in Stellar Astrophysics (MESA): Pulsating Variable Stars, Rotation, Convective Boundaries, and Energy Conservation. *The Astrophysical Journal Supplement Series*, 243(1):10.

Pulgar-Escobar, L. M., Henríquez-Salgado, N. A., Cerulo, P., and Mennickent, R. E. (2024). Characterizing NGC 6383: A study of pre-main sequence stars, mass segregation, and age using Gaia DR3 and 2MASS. *arXiv e-prints*, page arXiv:2405.09145.

Romero, A. (2022). Differences between the wd cooling models.

Saha, A., Axelrod, T., Olszewski, E., and DA White Dwarf SED Standards Team (2023). All Sky Faint DA White Dwarf Spectrophotometric Standards. In *American Astronomical Society Meeting Abstracts*, volume 241 of *American Astronomical Society Meeting Abstracts*, page 366.08.

Sahu, S., Gänsicke, B. T., Tremblay, P.-E., Koester, D., Hermes, J. J., Wilson, D. J., Toloza, O., Hoskin, M. J., Farihi, J., Manser, C. J., and Redfield, S. (2023). An HST COS ultraviolet

spectroscopic survey of 311 DA white dwarfs - I. Fundamental parameters and comparative studies. *MNRAS*, 526(4):5800–5823.

Sahu, S., Tremblay, P.-E., Lallement, R., Redfield, S., and Gänsicke, B. T. (2024). Ultraviolet extinction correlation with 3D dust maps using white dwarfs. *MNRAS*, 535(1):1147–1162.

Saumon, D., Blouin, S., and Tremblay, P.-E. (2022). Current challenges in the physics of white dwarf stars. *Phys. Rep.*, 988:1–63.

Stoehr, F., White, R., Smith, M., Kamp, I., Thompson, R., Durand, D., Freudling, W., Fraquelli, D., Haase, J., Hook, R., Kimball, T., Kümmel, M., Levay, K., Lombardi, M., Micol, A., and Rogers, T. (2008). Der_snr: A simple & general spectroscopic signal-to-noise measurement algorithm. *PASP*.

Tremblay, P. E. and Bergeron, P. (2009). Spectroscopic Analysis of DA White Dwarfs: Stark Broadening of Hydrogen Lines Including Nonideal Effects. *ApJ*, 696(2):1755–1770.

Vanderbosch, Z. P., Rappaport, S., Guidry, J. A., Gary, B. L., Blouin, S., Kaye, T. G., Weinberger, A. J., Melis, C., Klein, B. L., Zuckerman, B., Vanderburg, A., Hermes, J. J., Hegedus, R. J., Burleigh, M. R., Sefako, R., Worters, H. L., and Heintz, T. M. (2021). Recurring Planetary Debris Transits and Circumstellar Gas around White Dwarf ZTF J0328-1219. *ApJ*, 917(1):41.

Vanderburg, A., Johnson, J. A., Rappaport, S., Bieryla, A., Irwin, J., Lewis, J. A., Kipping, D., Brown, W. R., Dufour, P., Ciardi, D. R., Angus, R., Schaefer, L., Latham, D. W., Charbonneau, D., Beichman, C., Eastman, J., McCrady, N., Wittenmyer, R. A., and Wright, J. T. (2015). A disintegrating minor planet transiting a white dwarf. *Nature*, 526(7574):546–549.

Werner, K., Reindl, N., Löbling, L., Pelisoli, I., Schaffenroth, V., Rebassa-Mansergas, A., Irawati, P., and Ren, J. (2020). An extremely hot white dwarf with a rapidly rotating K-type subgiant companion: UCAC2 46706450. *A&A*, 642:A228.

Xu, S., Jura, M., Dufour, P., and Zuckerman, B. (2016). Evidence for Gas from a Disintegrating Extrasolar Asteroid. *ApJ*, 816(2):L22.

Xu, S., Jura, M., Koester, D., Klein, B., and Zuckerman, B. (2014). Elemental Compositions of Two Extrasolar Rocky Planetesimals. *ApJ*, 783(2):79.

Yang, Y., Zhao, J., Zhang, J., Ye, X., and Zhao, G. (2020). Estimating Atmospheric Parameters of DA White Dwarf Stars with Deep Learning. *AJ*, 160(5):236.

Zuckerman, B., Koester, D., Melis, C., Hansen, B. M., and Jura, M. (2007). The Chemical Composition of an Extrasolar Minor Planet. *ApJ*, 671(1):872–877.

Zuckerman, B., Koester, D., Reid, I. N., and Hünsch, M. (2003). Metal Lines in DA White Dwarfs. *ApJ*, 596(1):477–495.

Appendix A

Validation Set Results

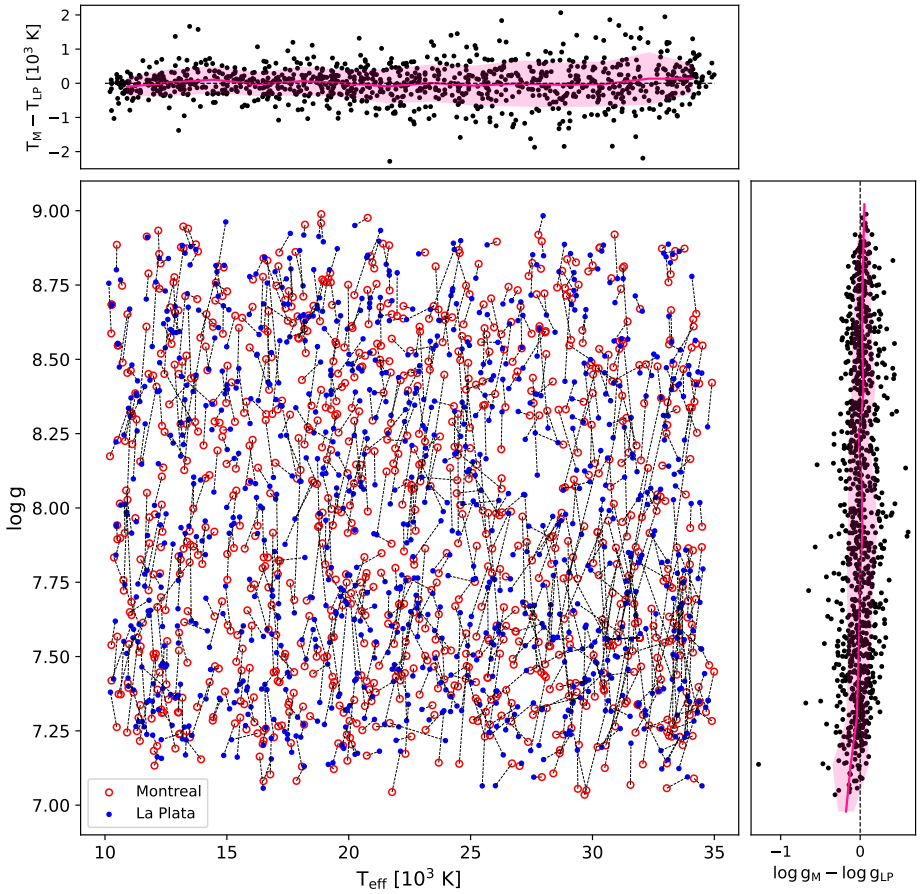


Figure A.1: Predictions of $T_{\rm eff}$ and $\log(g)$ using Montreal (red circles) and La Plata (blue dots) cooling models for the synthetic validation set. The results for the same instance are joined by a black dashed line. At the top and left side we include the distribution of the difference between the prediction for Montreal and La Plata (black dots) with the mean at each $T_{\rm eff}$ or $\log(g)$ value (solid pink line), plotted with respect to the Montreal results. The filled pink band corresponds to a deviation of $\pm\sigma$ for each $T_{\rm eff}$ and $\log(g)$ value. As seen, the scatter of the data points is stochastic in the parameter space.

Appendix B

Observed Fluxes: Montreal Golden Sample

In Figures Figures B.1 to B.3 we show the golden sample observed spectrum according to the Montreal predictions, with their corresponding synthetic spectra for each M-R relation.

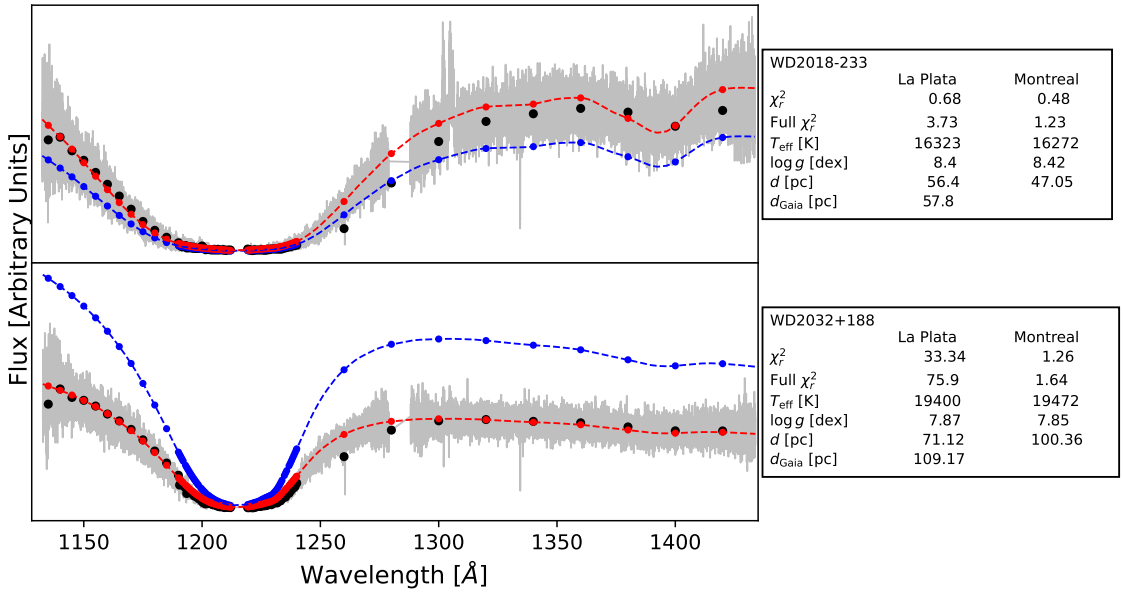


Figure B.1: Observed spectrum (gray) with its corresponding sampling (black dots). The fits obtained with Montreal (red dots and dashed line) and La Plata (blue dots and dashed line) are also included. The name of the object, the χ^2 value for the sampling and the full spectrum, the predicted $T_{\rm eff}$, $\log(g)$, and distance, and the corresponding *Gaia* distance are included in a text box at the right of each golden (according to the Montreal distance prediction and full χ^2 value) spectrum.

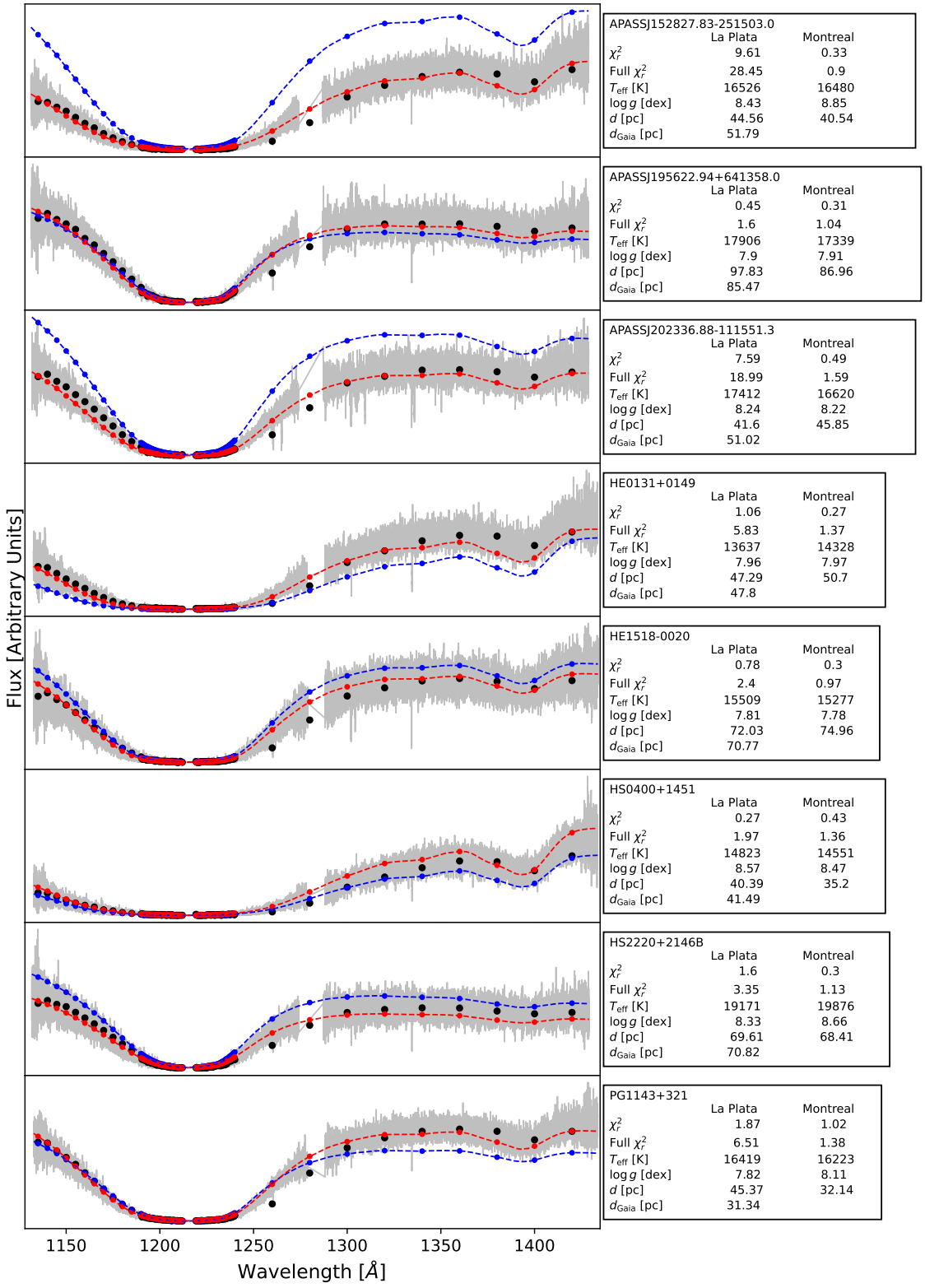


Figure B.2: Same as Fig. B.1.

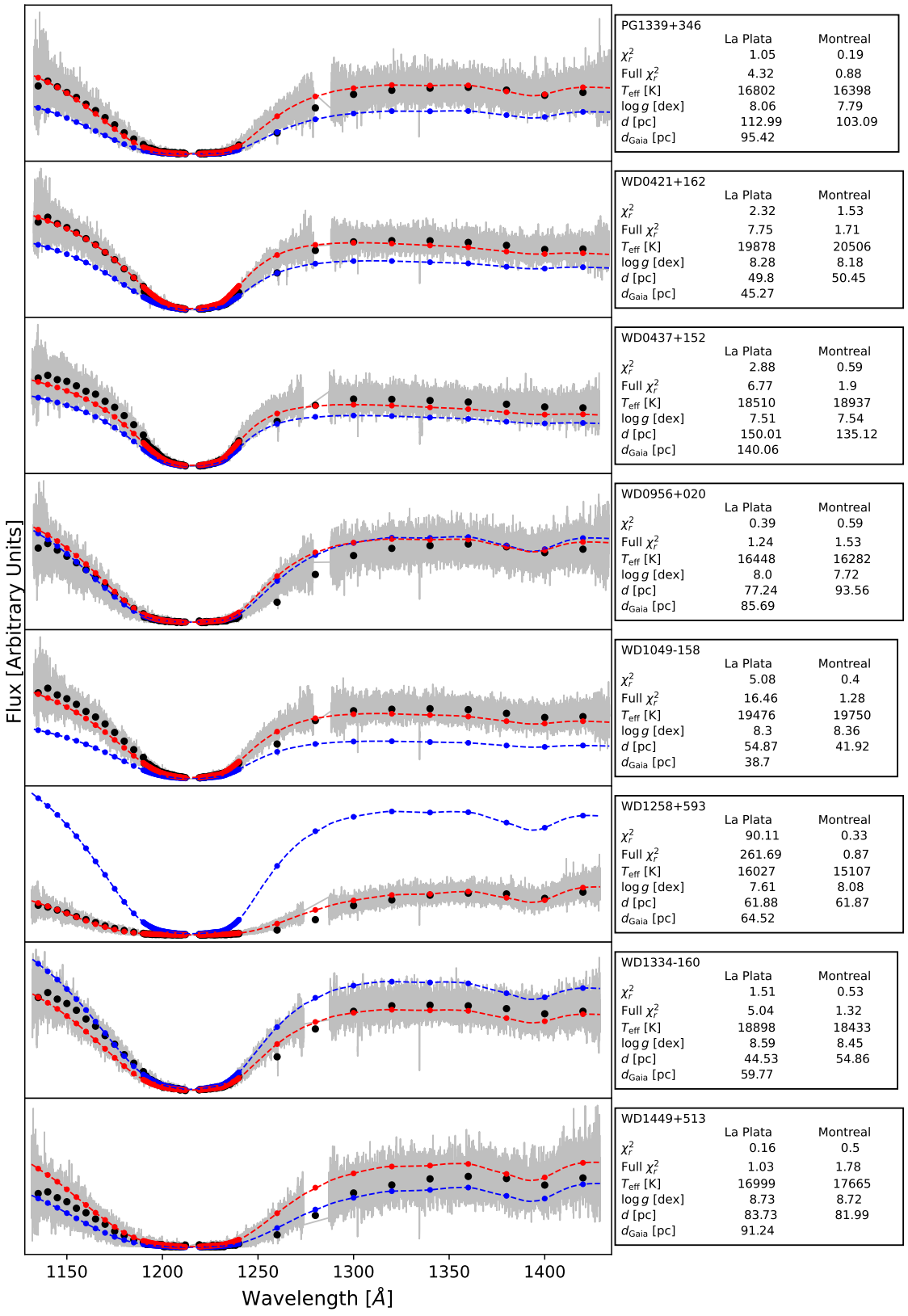


Figure B.3: Same as Fig. B.1.

Appendix C

Observed Fluxes: La Plata Golden Sample

In Figures C.1 to C.3 we show the golden sample observed spectrum according to the La Plata predictions, with their corresponding synthetic spectra for each M-R relation.

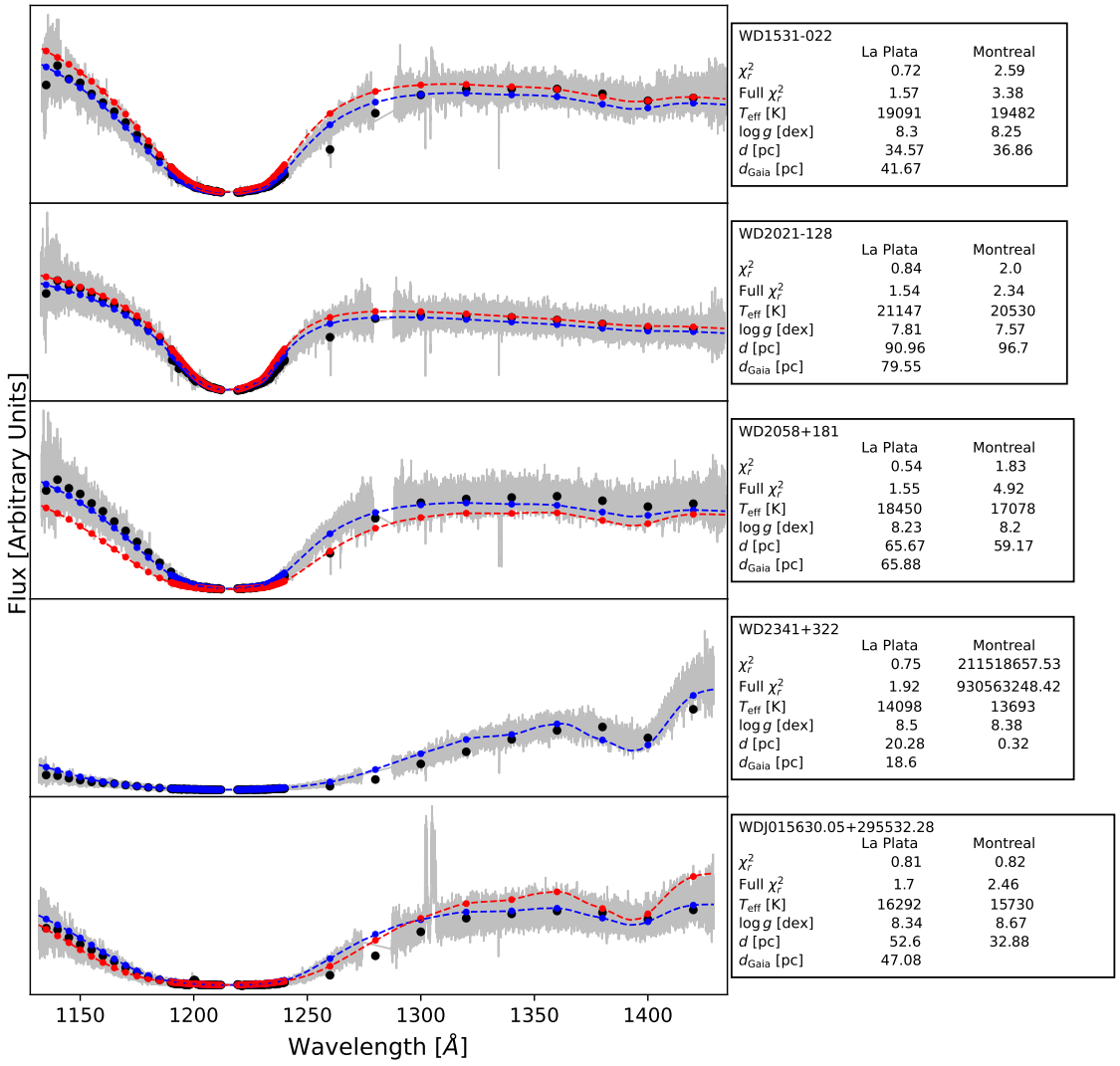


Figure C.1: Observed spectrum (gray) with its corresponding sampling (black dots). The fits obtained with Montreal (red dots and dashed line) and La Plata (blue dots and dashed line) are also included. The name of the object, the χ^2 value for the sampling and the full spectrum, the predicted $T_{\rm eff}$, $\log(g)$, and distance, and the corresponding*Gaia* distance are included in a text box at the right of each golden (according to the La Plata distance prediction and full χ^2 value) spectrum.

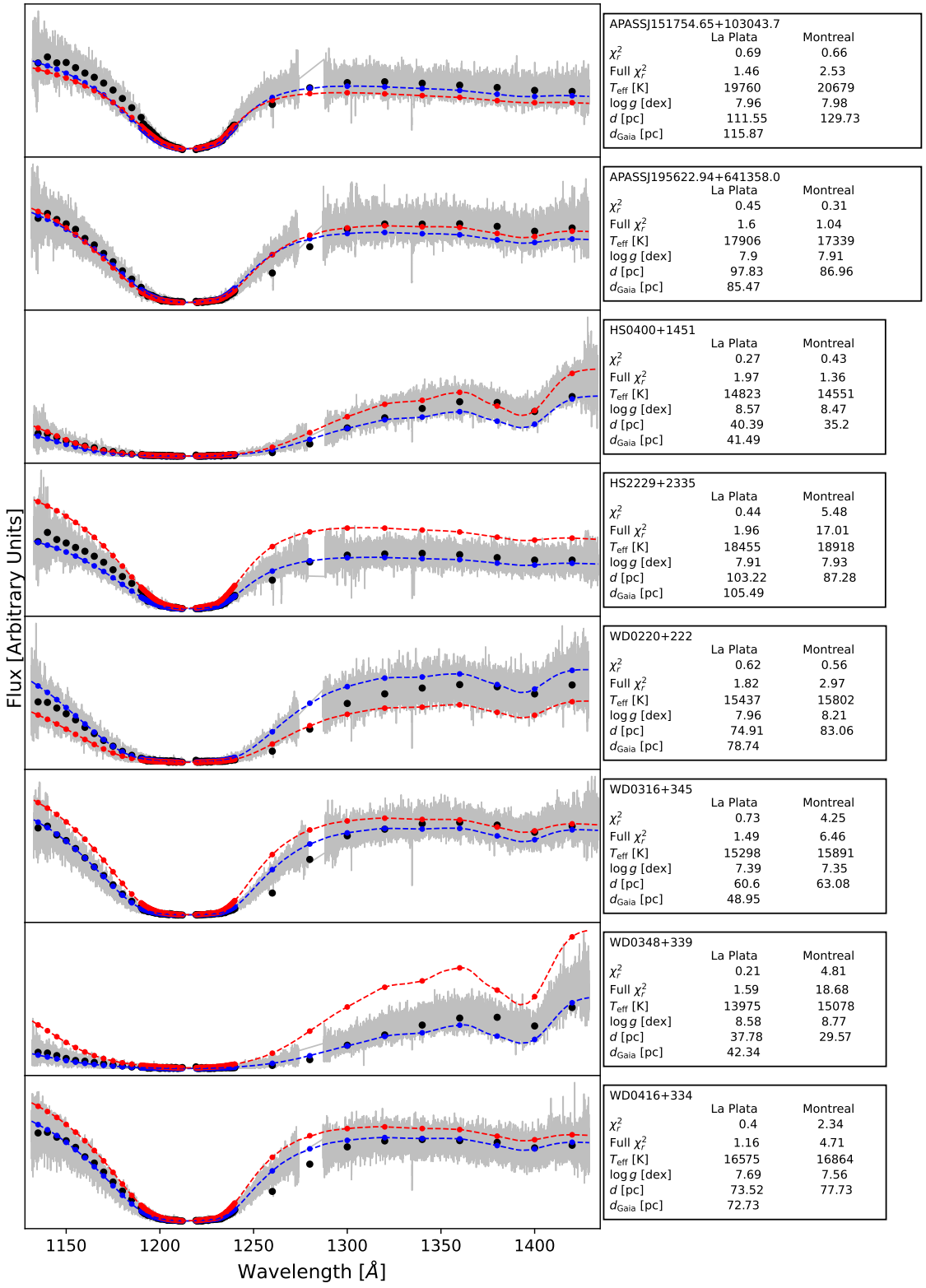
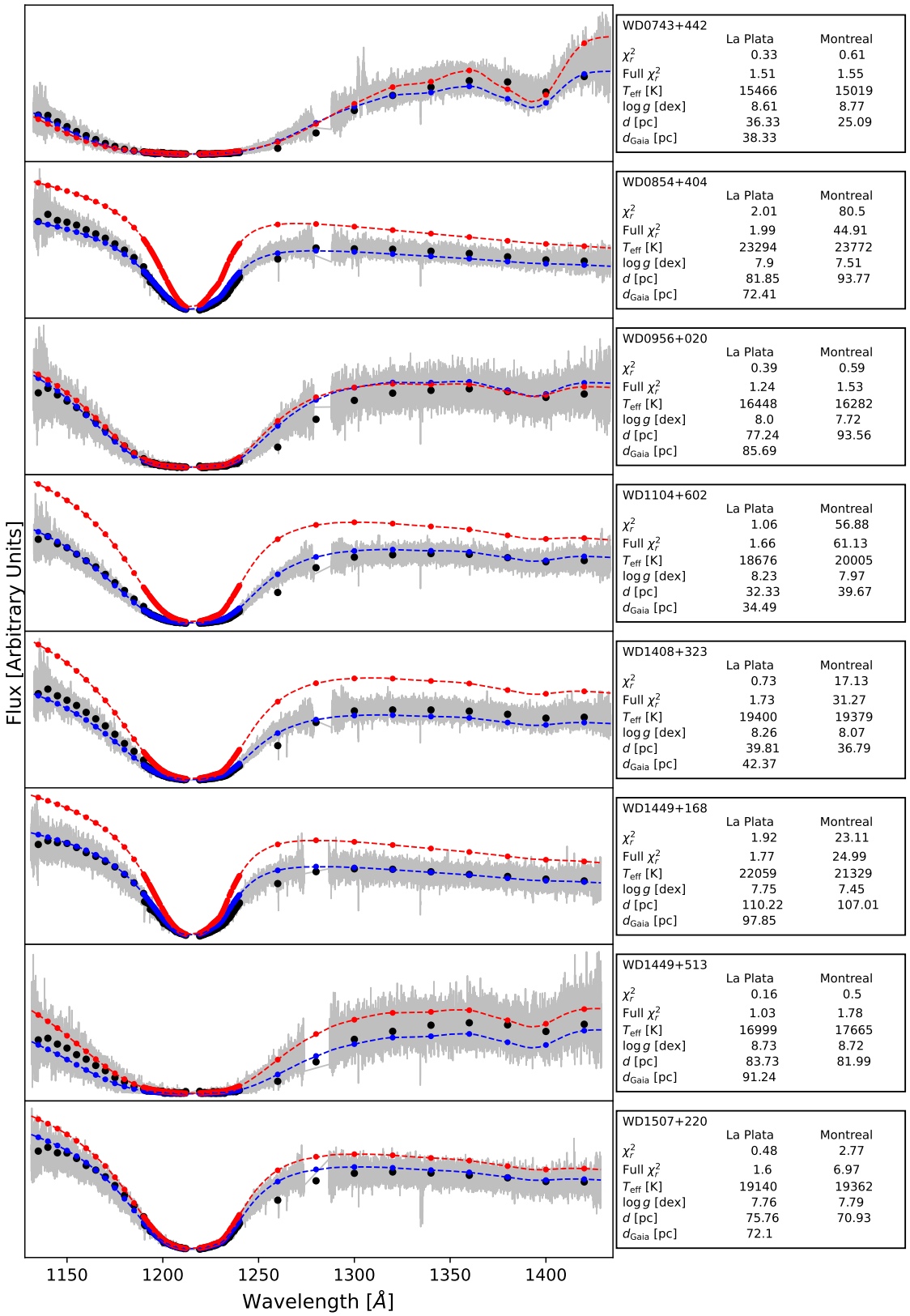


Figure C.2: Same as Fig. C.1.



WD0743+442	La Plata	Montreal
χ^2_r	0.33	0.61
Full χ^2_r	1.51	1.55
$T_{\rm eff}$ [K]	15466	15019
log g [dex]	8.61	8.77
d [pc]	36.33	25.09
$d_{\rm Gaia}$ [pc]	38.33	

WD0854+404	La Plata	Montreal
χ^2_r	2.01	80.5
Full χ^2_r	1.99	44.91
$T_{\rm eff}$ [K]	23294	23772
log g [dex]	7.9	7.51
d [pc]	81.85	93.77
$d_{\rm Gaia}$ [pc]	72.41	

WD0956+020	La Plata	Montreal
χ^2_r	0.39	0.59
Full χ^2_r	1.24	1.53
$T_{\rm eff}$ [K]	16448	16282
log g [dex]	8.0	7.72
d [pc]	77.24	93.56
$d_{\rm Gaia}$ [pc]	85.69	

WD1104+602	La Plata	Montreal
χ^2_r	1.06	56.88
Full χ^2_r	1.66	61.13
$T_{\rm eff}$ [K]	18676	20005
log g [dex]	8.23	7.97
d [pc]	32.33	39.67
$d_{\rm Gaia}$ [pc]	34.49	

WD1408+323	La Plata	Montreal
χ^2_r	0.73	17.13
Full χ^2_r	1.73	31.27
$T_{\rm eff}$ [K]	19400	19379
log g [dex]	8.26	8.07
d [pc]	39.81	36.79
$d_{\rm Gaia}$ [pc]	42.37	

WD1449+168	La Plata	Montreal
χ^2_r	1.92	23.11
Full χ^2_r	1.77	24.99
$T_{\rm eff}$ [K]	22059	21329
log g [dex]	7.75	7.45
d [pc]	110.22	107.01
$d_{\rm Gaia}$ [pc]	97.85	

WD1449+513	La Plata	Montreal
χ^2_r	0.16	0.5
Full χ^2_r	1.03	1.78
$T_{\rm eff}$ [K]	16999	17665
log g [dex]	8.73	8.72
d [pc]	83.73	81.99
$d_{\rm Gaia}$ [pc]	91.24	

WD1507+220	La Plata	Montreal
χ^2_r	0.48	2.77
Full χ^2_r	1.6	6.97
$T_{\rm eff}$ [K]	19140	19362
log g [dex]	7.76	7.79
d [pc]	75.76	70.93
$d_{\rm Gaia}$ [pc]	72.1	

Figure C.3: Same as Fig. C.1.

Appendix D

Observed Fluxes: Non-Golden Objects

In Figures Figures D.1 to D.28 we include the predicted parameters and the corresponding spectra for the Montreal and La Plata models, where the objects where not in the golden sample shown above, and described in Section 3.2.2.

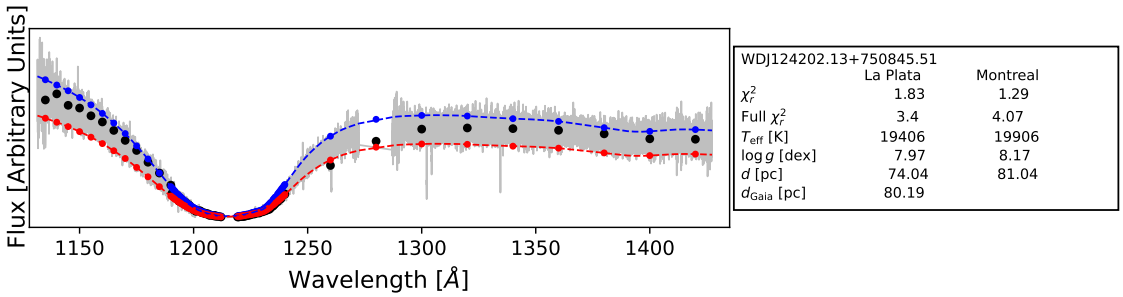


Figure D.1: Observed spectrum (gray) with its corresponding sampling (black dots). The fits obtained with Montreal (red dots and dashed line) and La Plata (blue dots and dashed line) are also included. The name of the object, the χ^2 value for the sampling and the full spectrum, the predicted $T_{\rm eff}$, $\log(g)$, and distance, and the corresponding *Gaia* distance are included in a text box at the right of each non-golden spectrum.

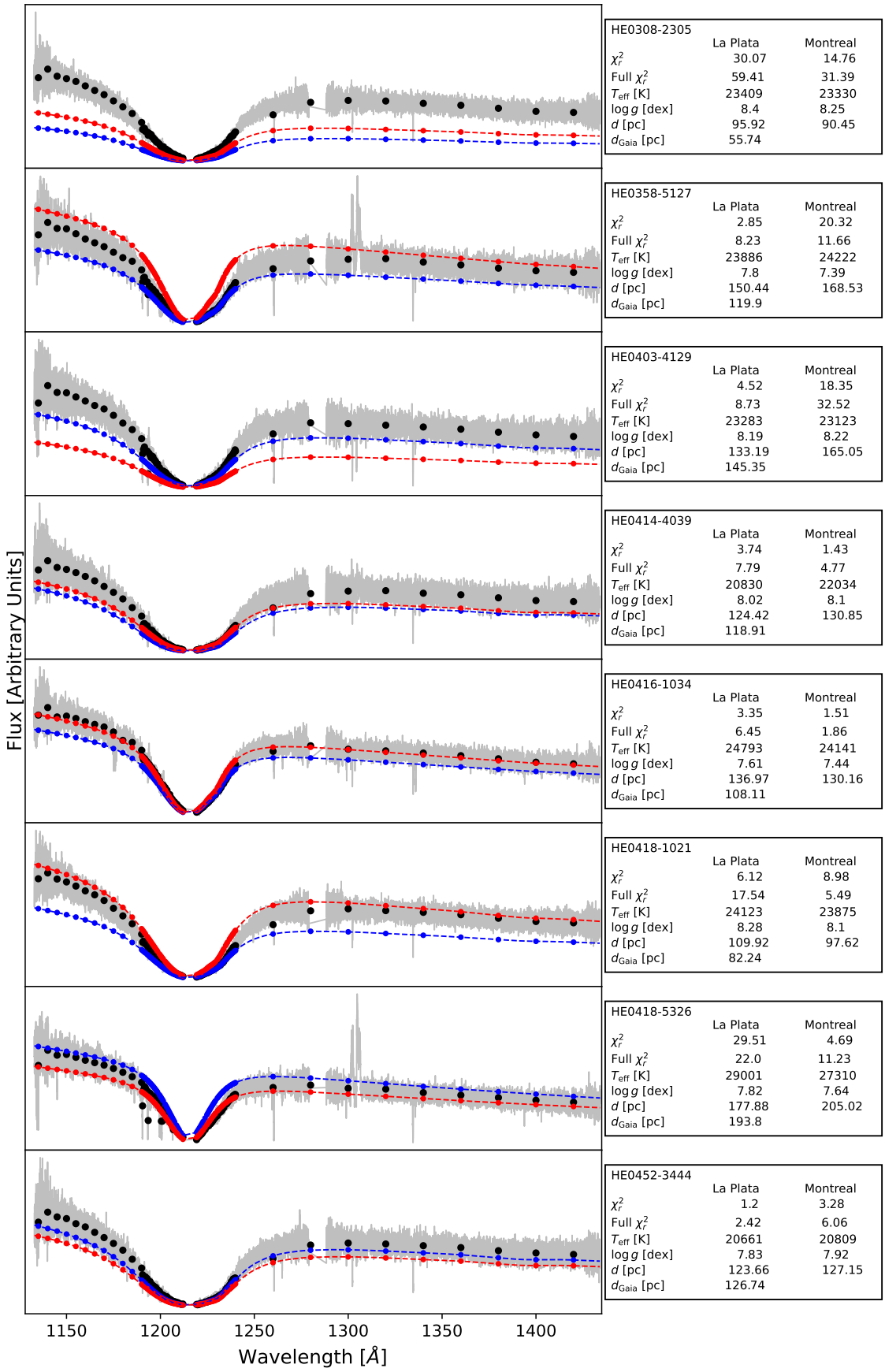


Figure D.2: Same as Fig. D.1.

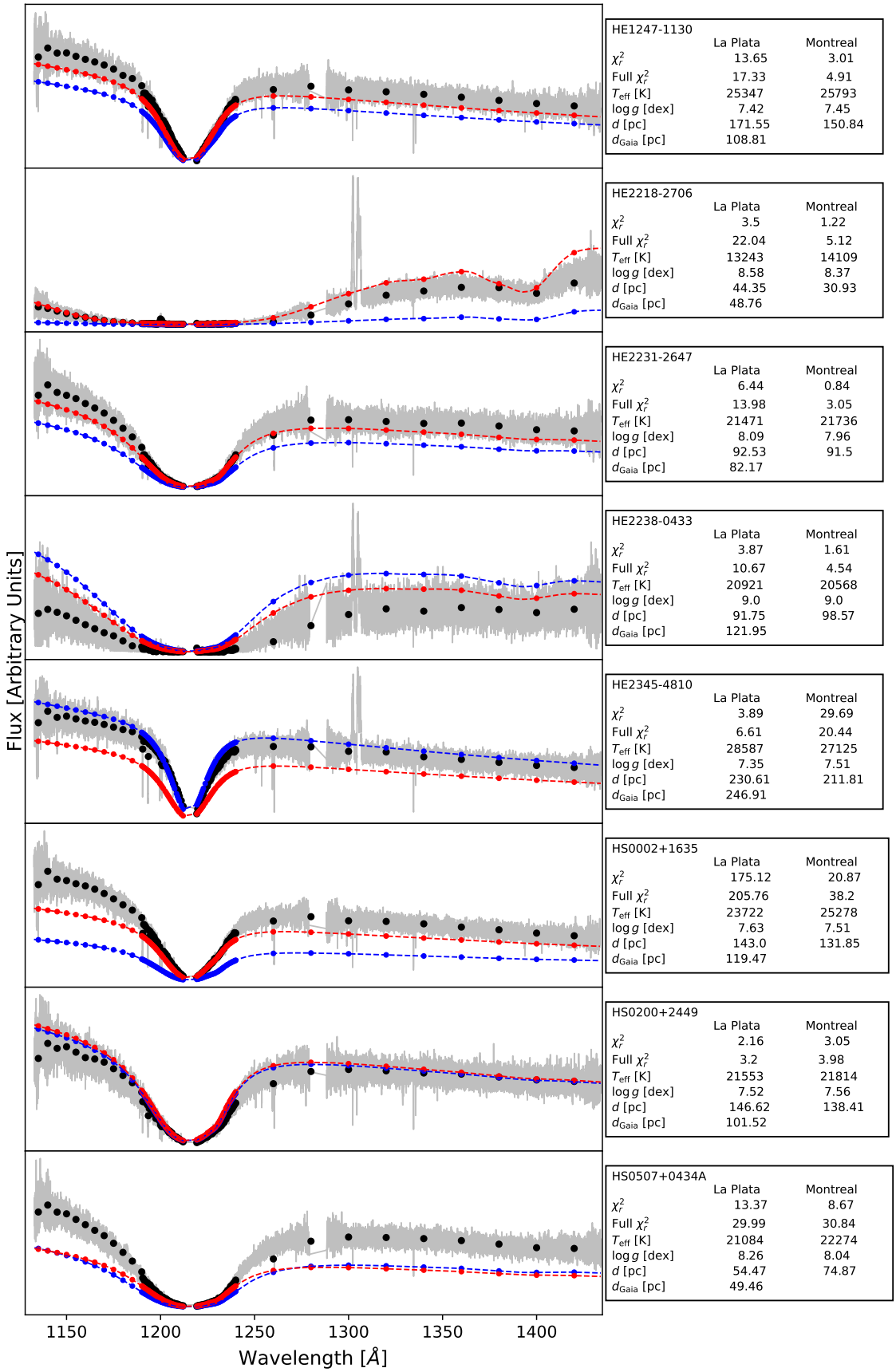


Figure D.3: Same as Fig. D.1.

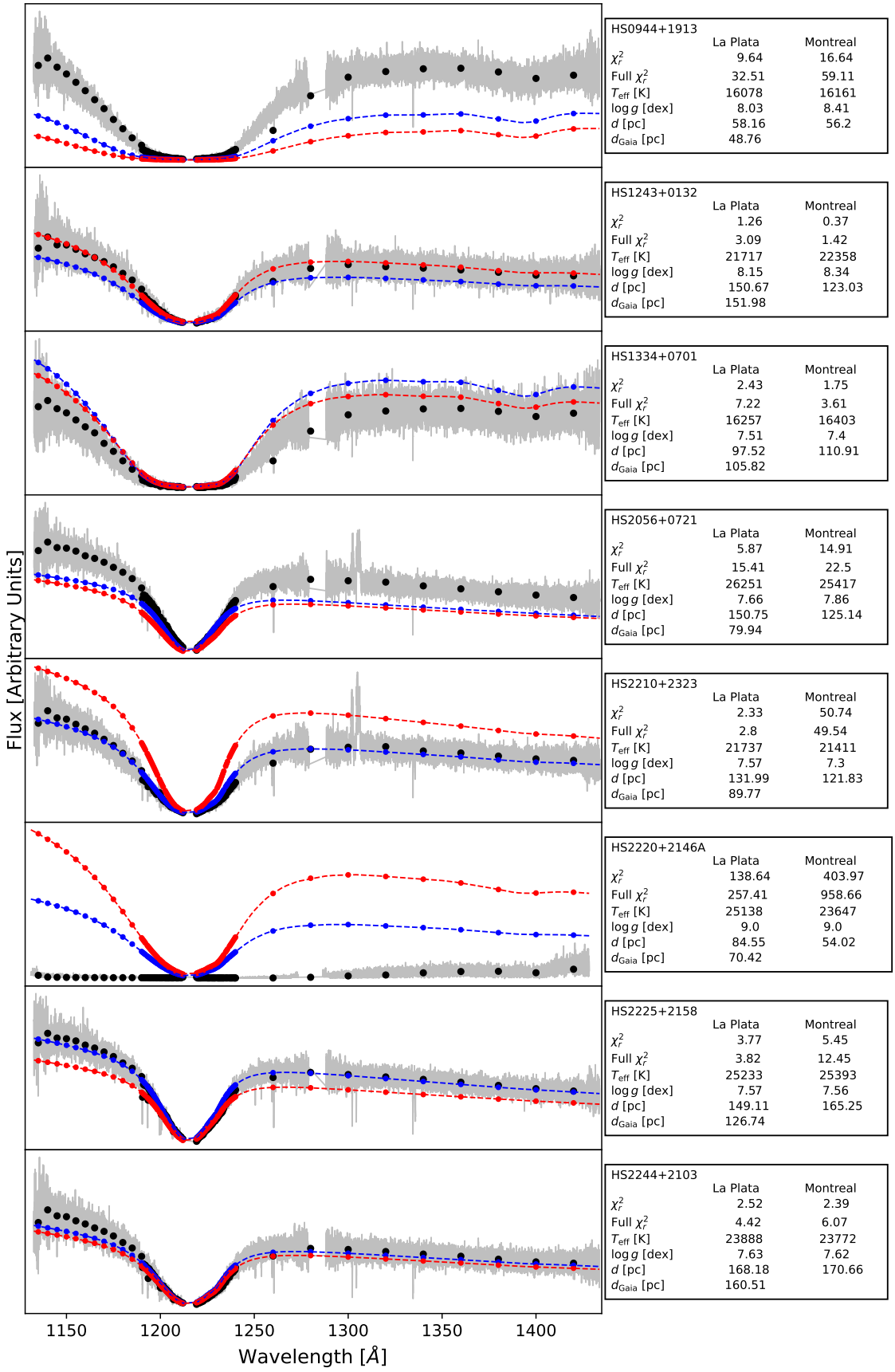


Figure D.4: Same as Fig. D.1.

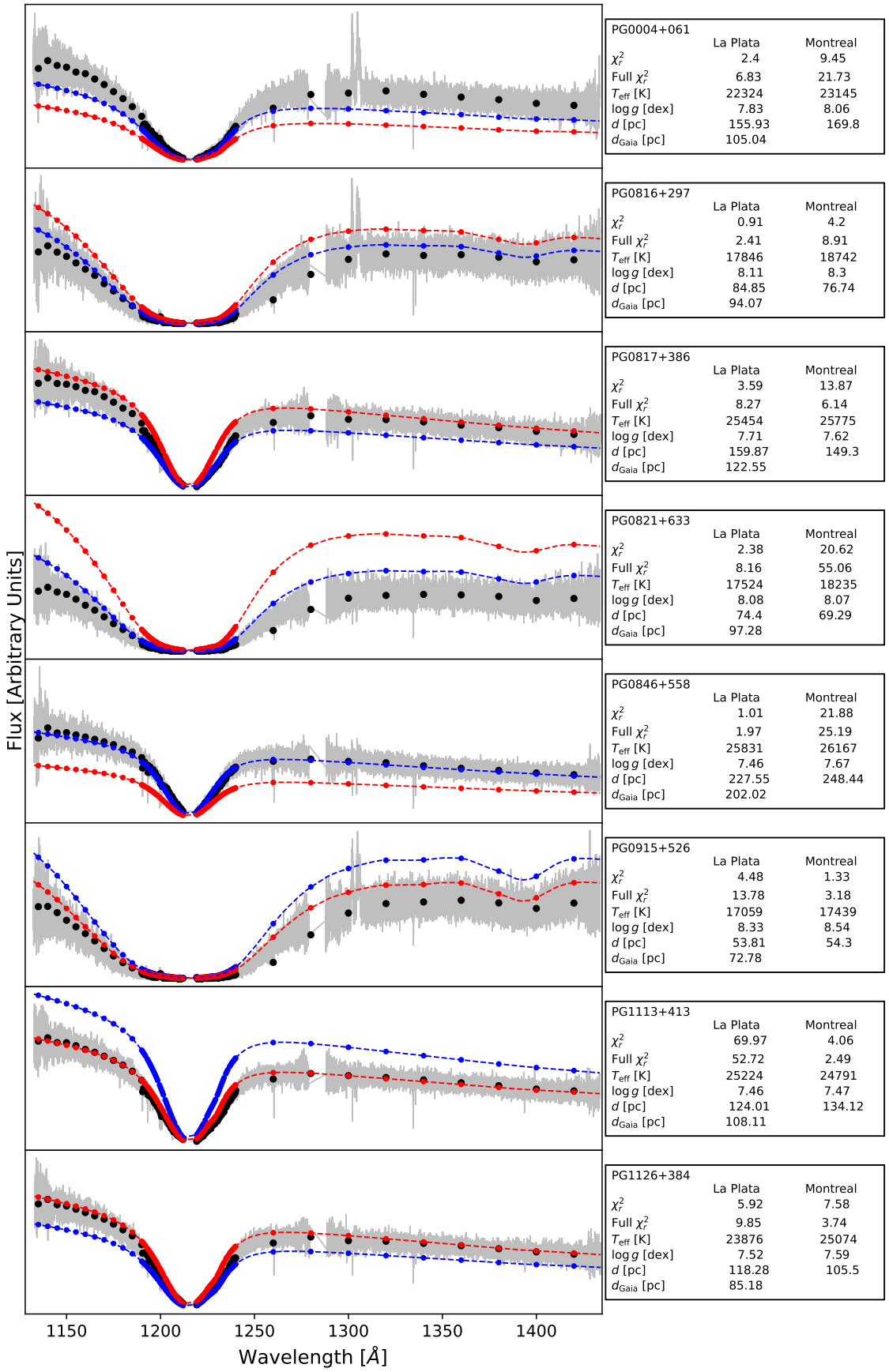


Figure D.5: Same as Fig. D.1.

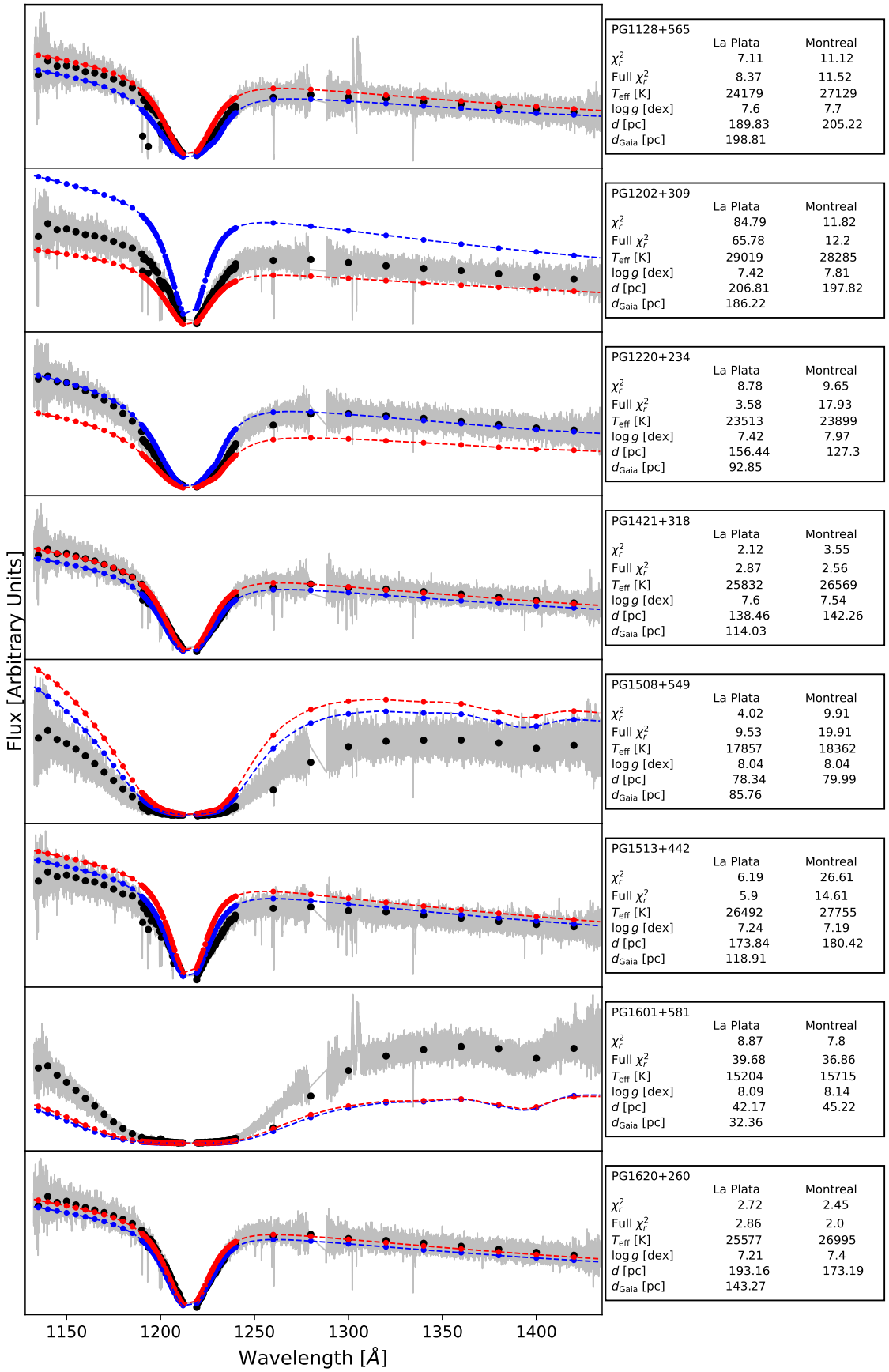


Figure D.6: Same as Fig. D.1.

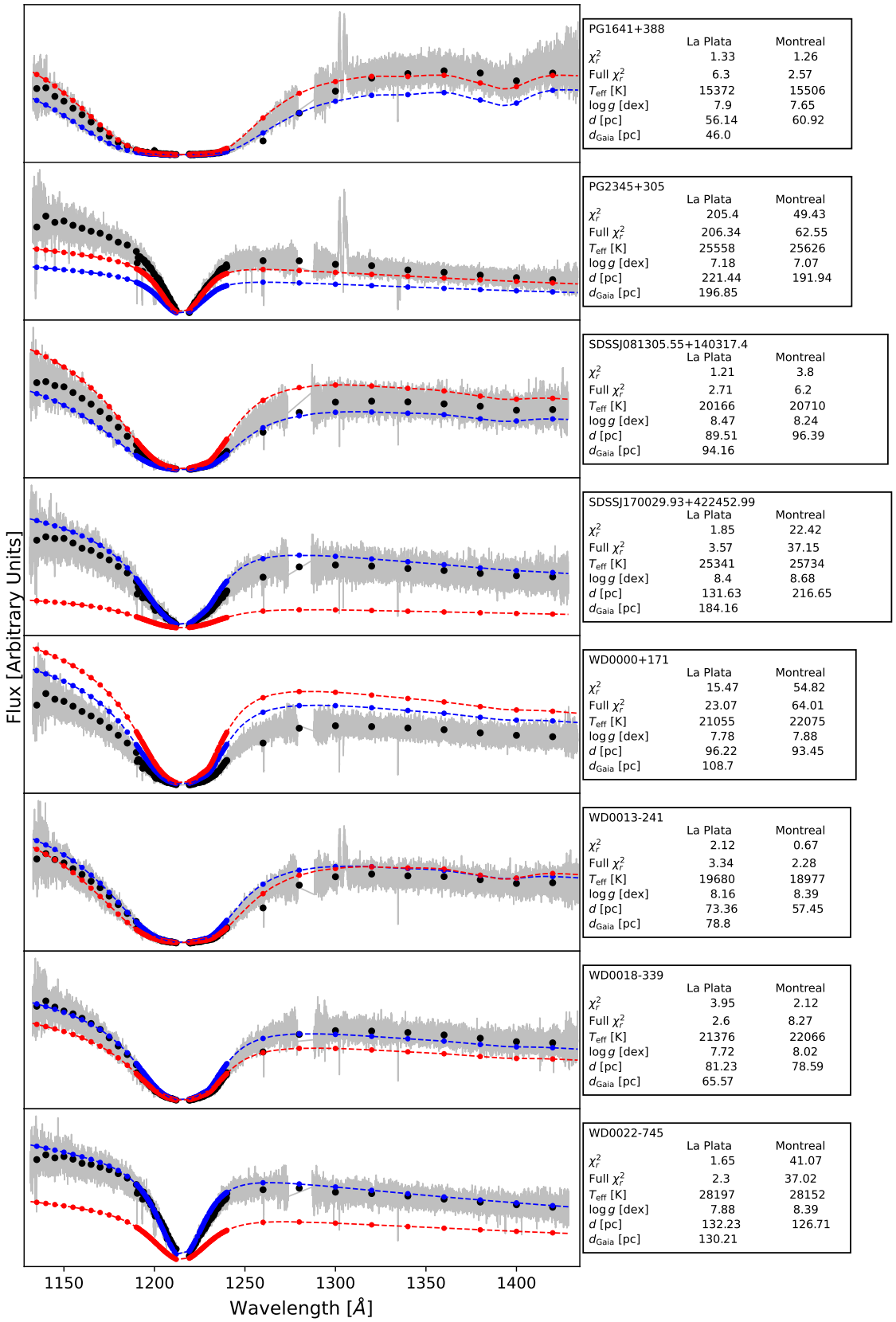


Figure D.7: Same as Fig. D.1.

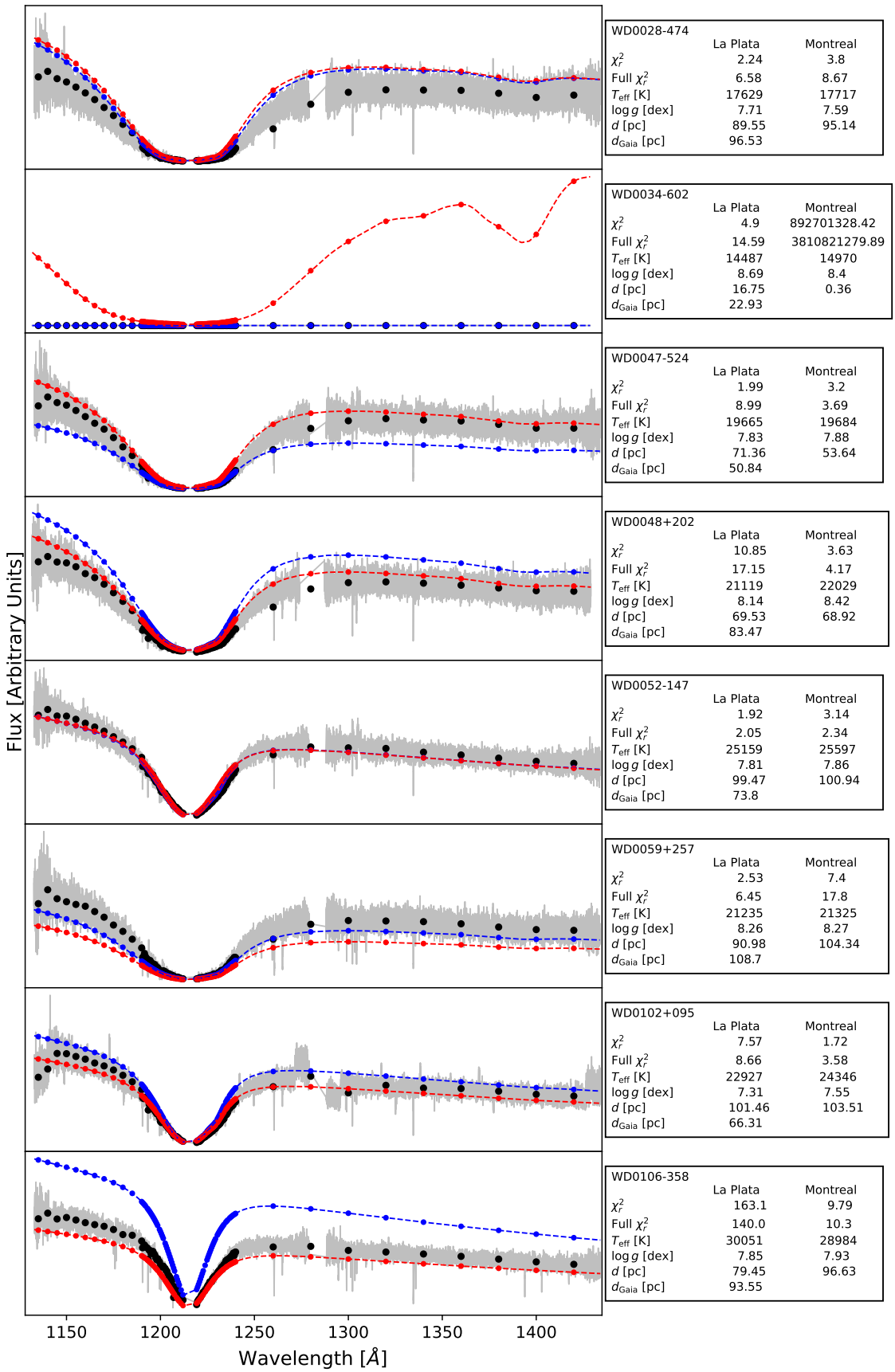


Figure D.8: Same as Fig. D.1.

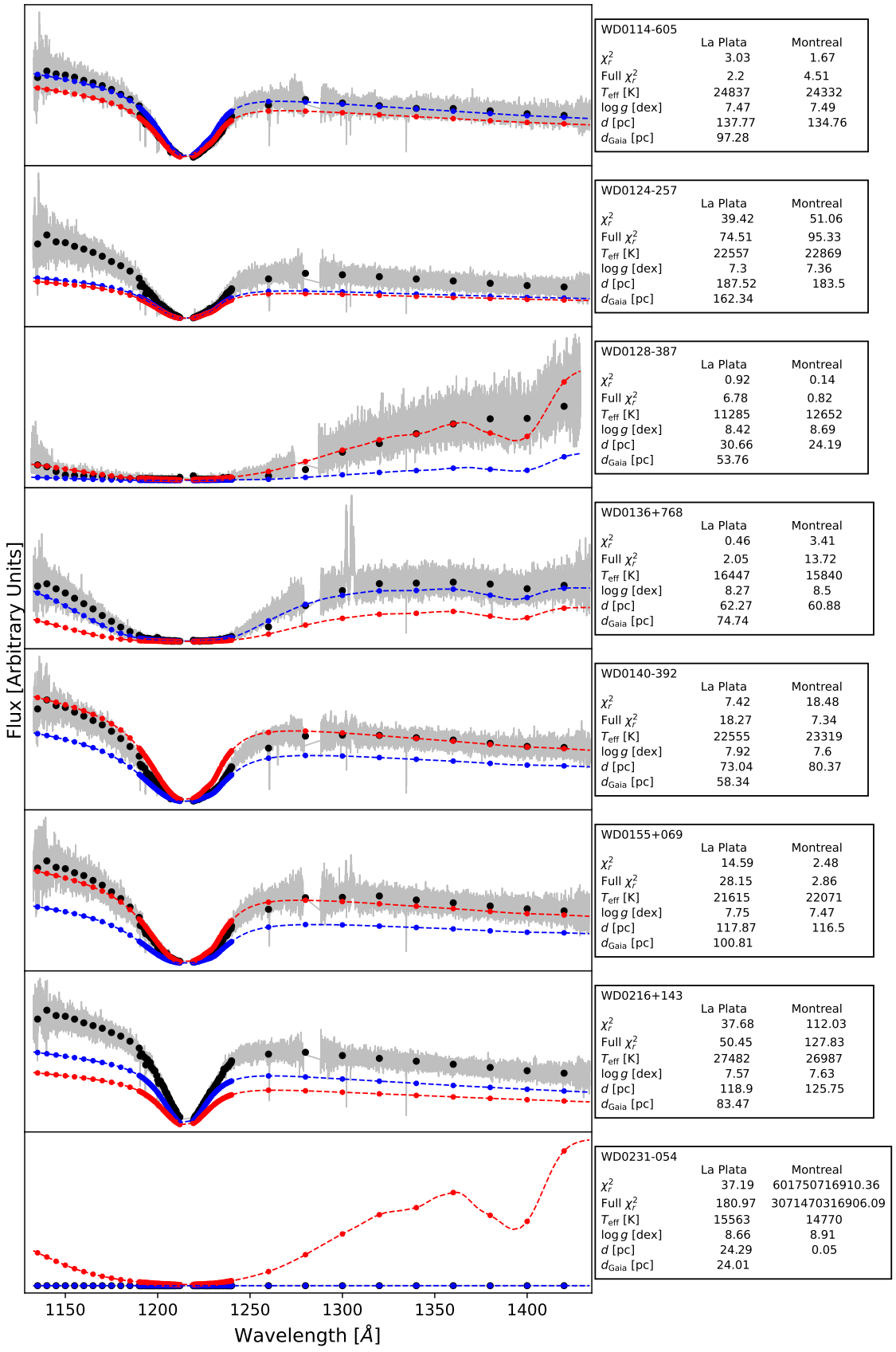


Figure D.9: Same as Fig. D.1.

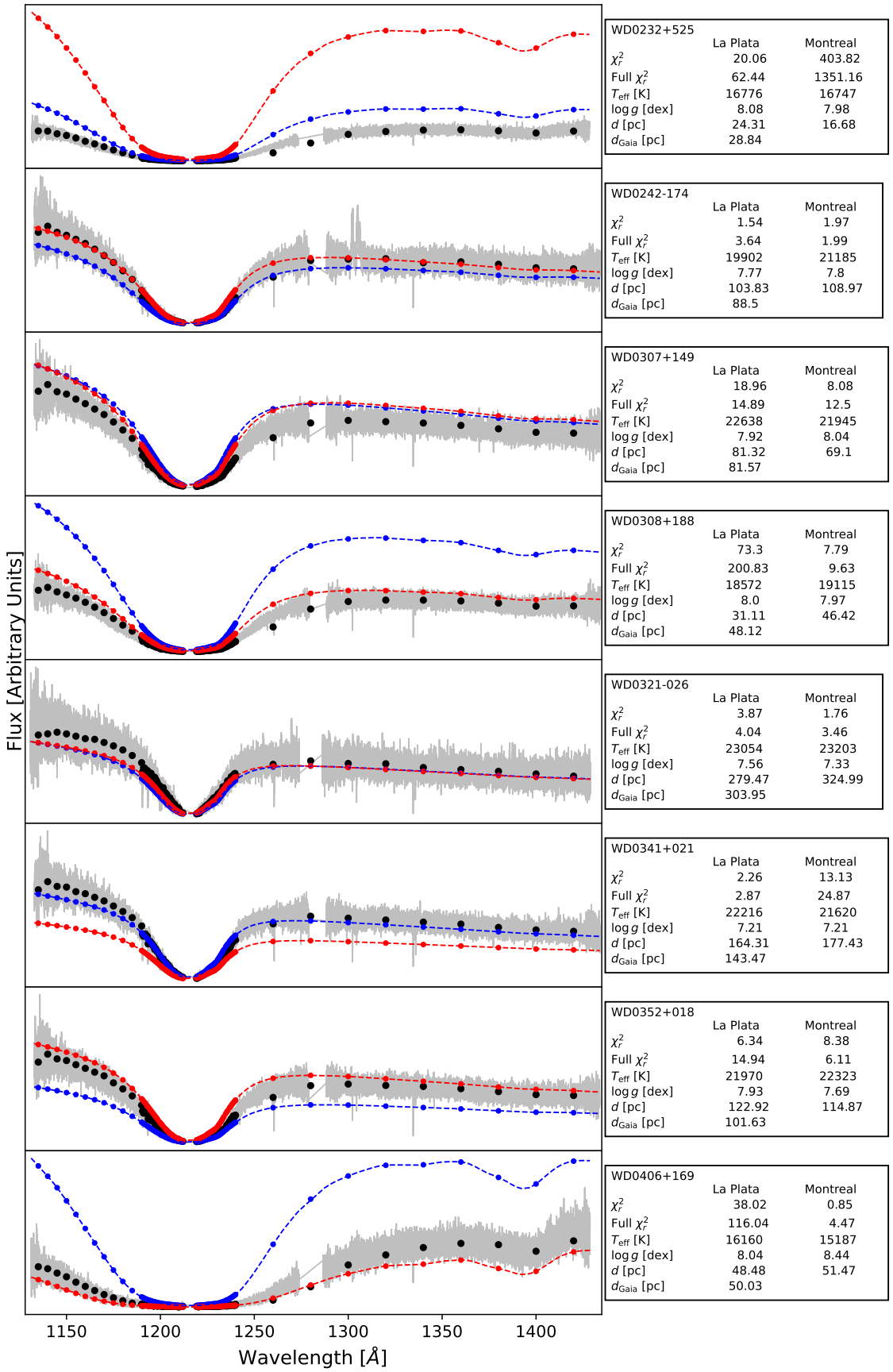


Figure D.10: Same as Fig. D.1.

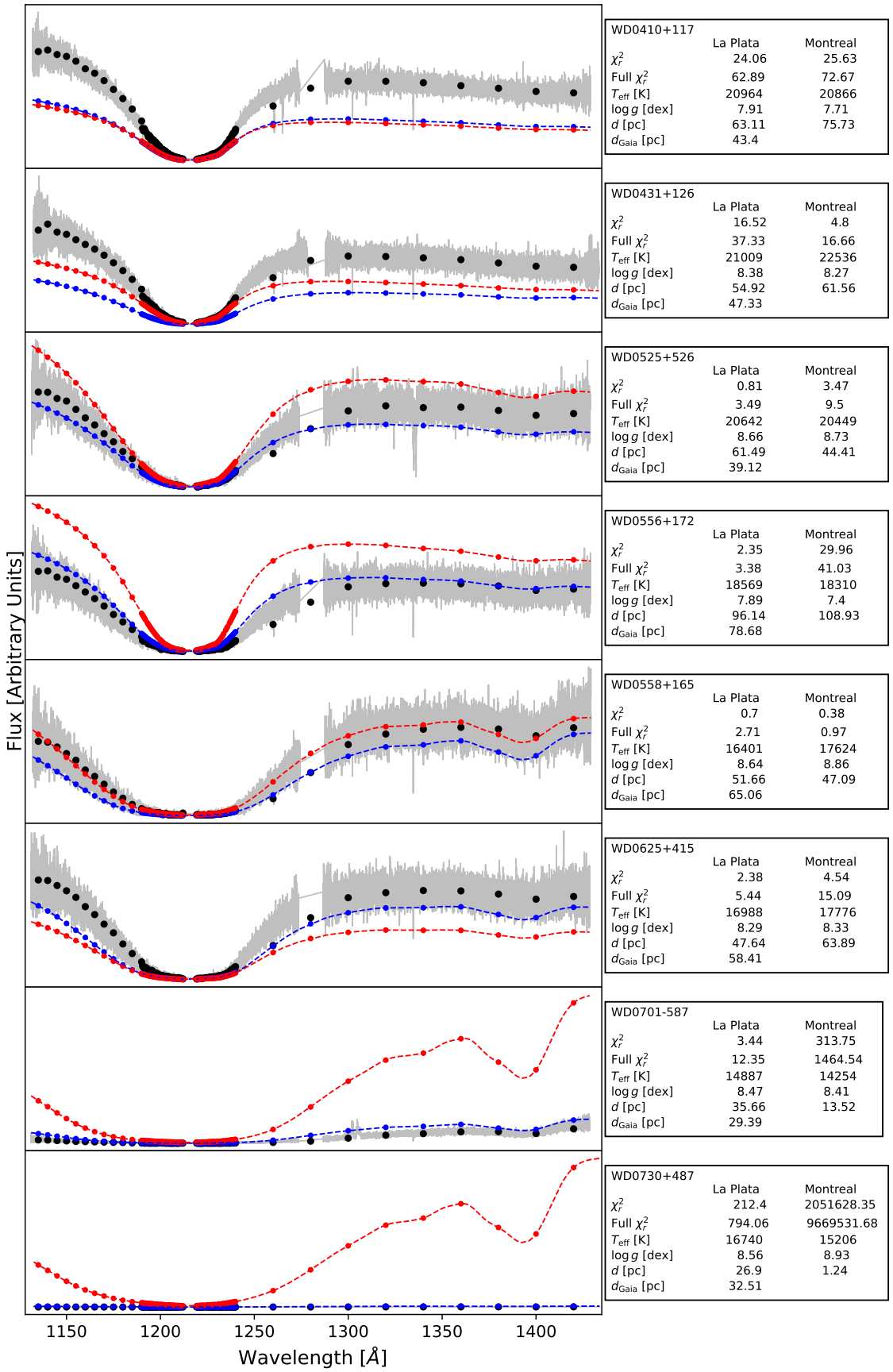
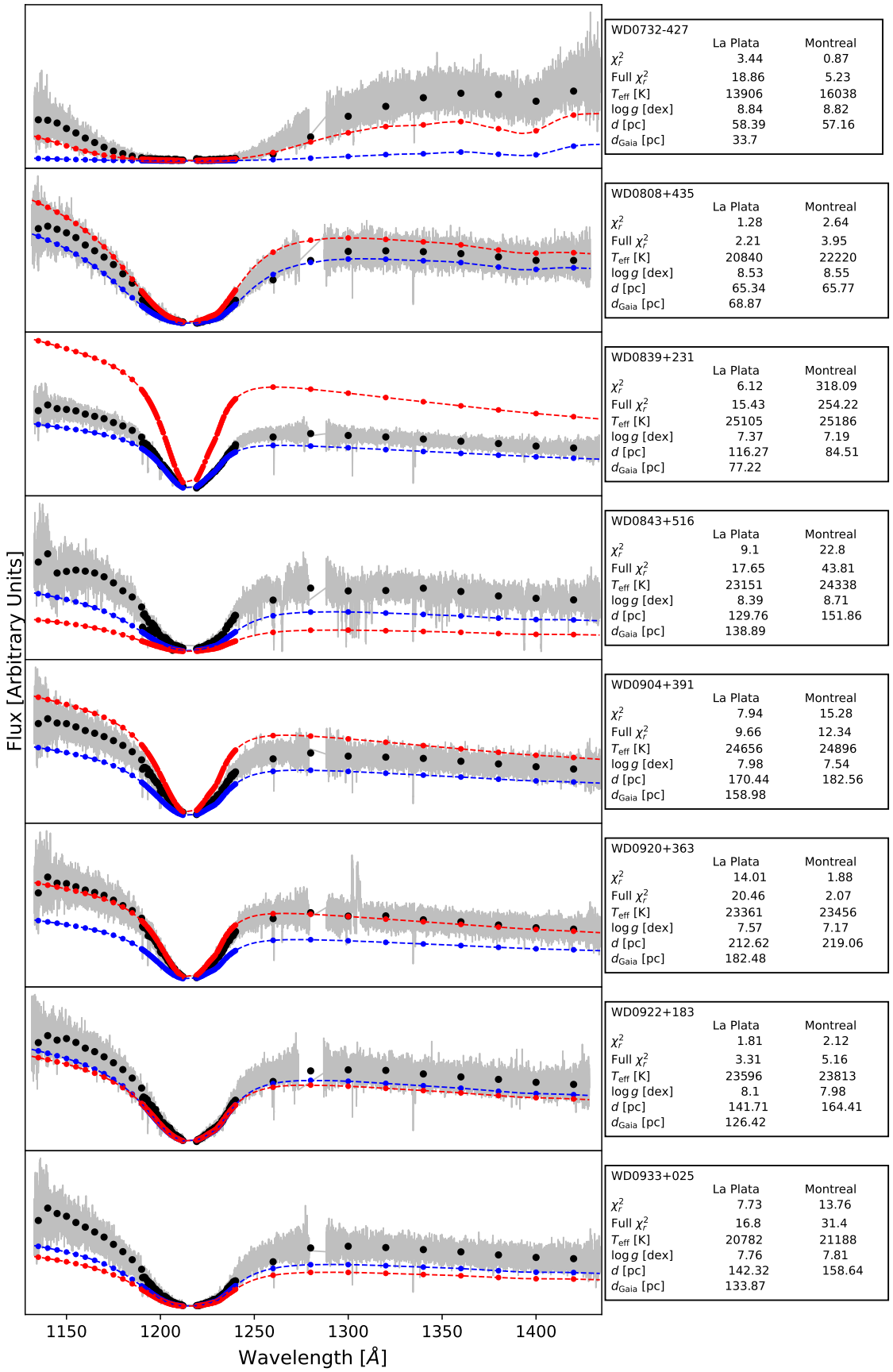


Figure D.11: Same as Fig. D.1.



WD0732-427	La Plata	Montreal
χ^2_r	3.44	0.87
Full χ^2_r	18.86	5.23
$T_{\rm eff}$ [K]	13906	16038
log g [dex]	8.84	8.82
d [pc]	58.39	57.16
$d_{\rm Gaia}$ [pc]	33.7	

WD0808+435	La Plata	Montreal
χ^2_r	1.28	2.64
Full χ^2_r	2.21	3.95
$T_{\rm eff}$ [K]	20840	22220
log g [dex]	8.53	8.55
d [pc]	65.34	65.77
$d_{\rm Gaia}$ [pc]	68.87	

WD0839+231	La Plata	Montreal
χ^2_r	6.12	318.09
Full χ^2_r	15.43	254.22
$T_{\rm eff}$ [K]	25105	25186
log g [dex]	7.37	7.19
d [pc]	116.27	84.51
$d_{\rm Gaia}$ [pc]	77.22	

WD0843+516	La Plata	Montreal
χ^2_r	9.1	22.8
Full χ^2_r	17.65	43.81
$T_{\rm eff}$ [K]	23151	24338
log g [dex]	8.39	8.71
d [pc]	129.76	151.86
$d_{\rm Gaia}$ [pc]	138.89	

WD0904+391	La Plata	Montreal
χ^2_r	7.94	15.28
Full χ^2_r	9.66	12.34
$T_{\rm eff}$ [K]	24656	24896
log g [dex]	7.98	7.54
d [pc]	170.44	182.56
$d_{\rm Gaia}$ [pc]	158.98	

WD0920+363	La Plata	Montreal
χ^2_r	14.01	1.88
Full χ^2_r	20.46	2.07
$T_{\rm eff}$ [K]	23361	23456
log g [dex]	7.57	7.17
d [pc]	212.62	219.06
$d_{\rm Gaia}$ [pc]	182.48	

WD0922+183	La Plata	Montreal
χ^2_r	1.81	2.12
Full χ^2_r	3.31	5.16
$T_{\rm eff}$ [K]	23596	23813
log g [dex]	8.1	7.98
d [pc]	141.71	164.41
$d_{\rm Gaia}$ [pc]	126.42	

WD0933+025	La Plata	Montreal
χ^2_r	7.73	13.76
Full χ^2_r	16.8	31.4
$T_{\rm eff}$ [K]	20782	21188
log g [dex]	7.76	7.81
d [pc]	142.32	158.64
$d_{\rm Gaia}$ [pc]	133.87	

Figure D.12: Same as Fig. D.1.

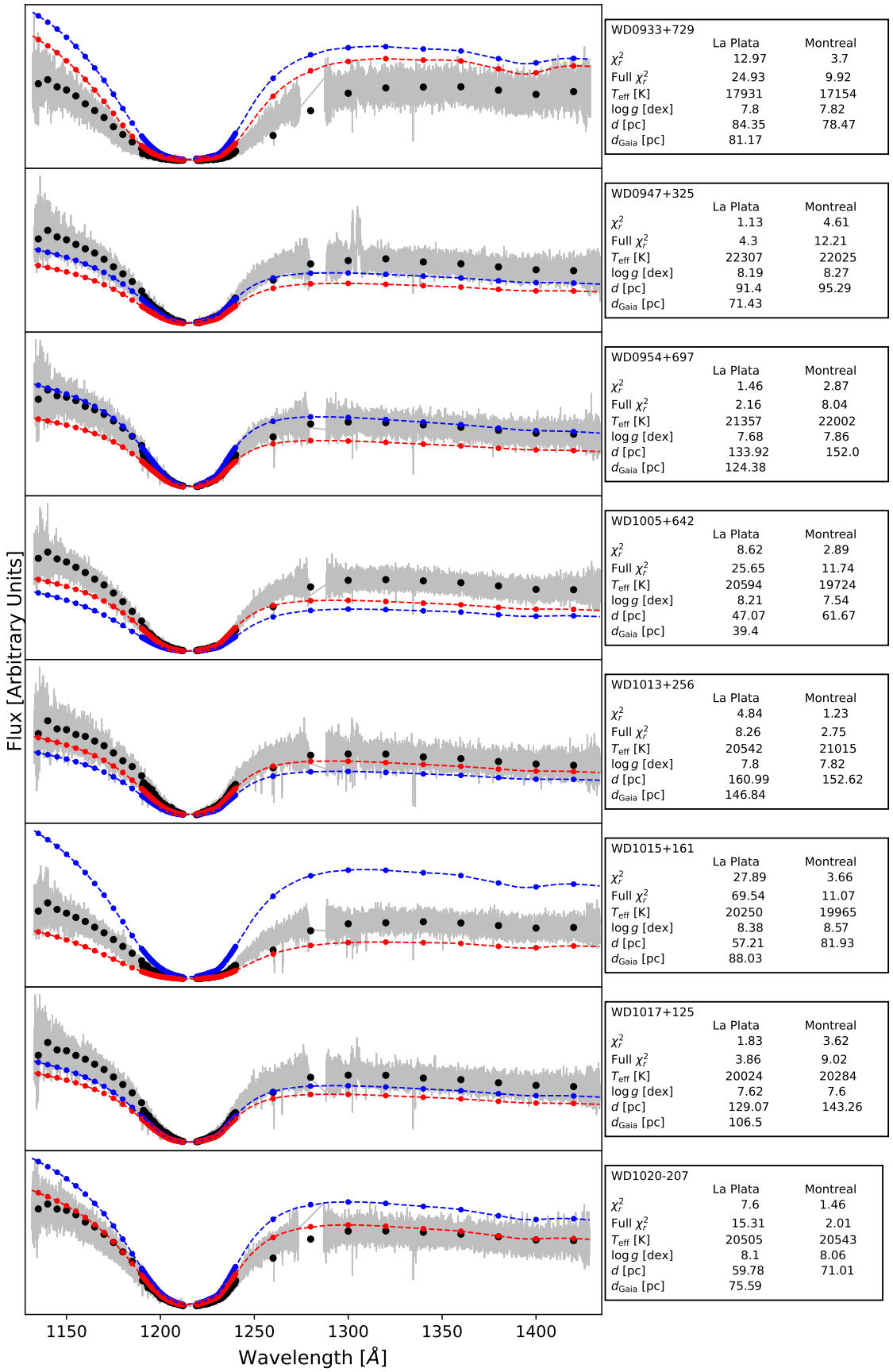


Figure D.13: Same as Fig. D.1.

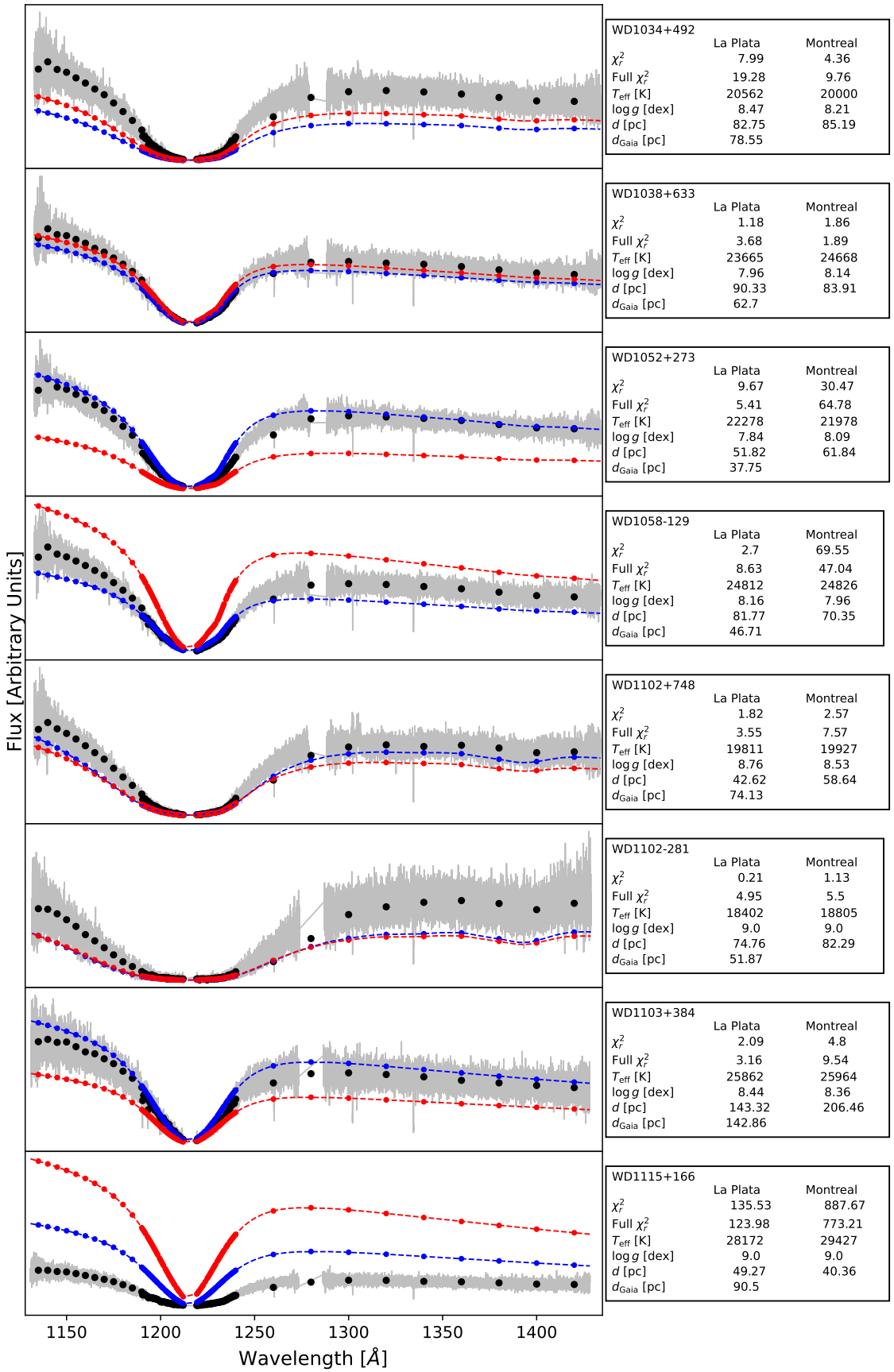


Figure D.14: Same as Fig. D.1.

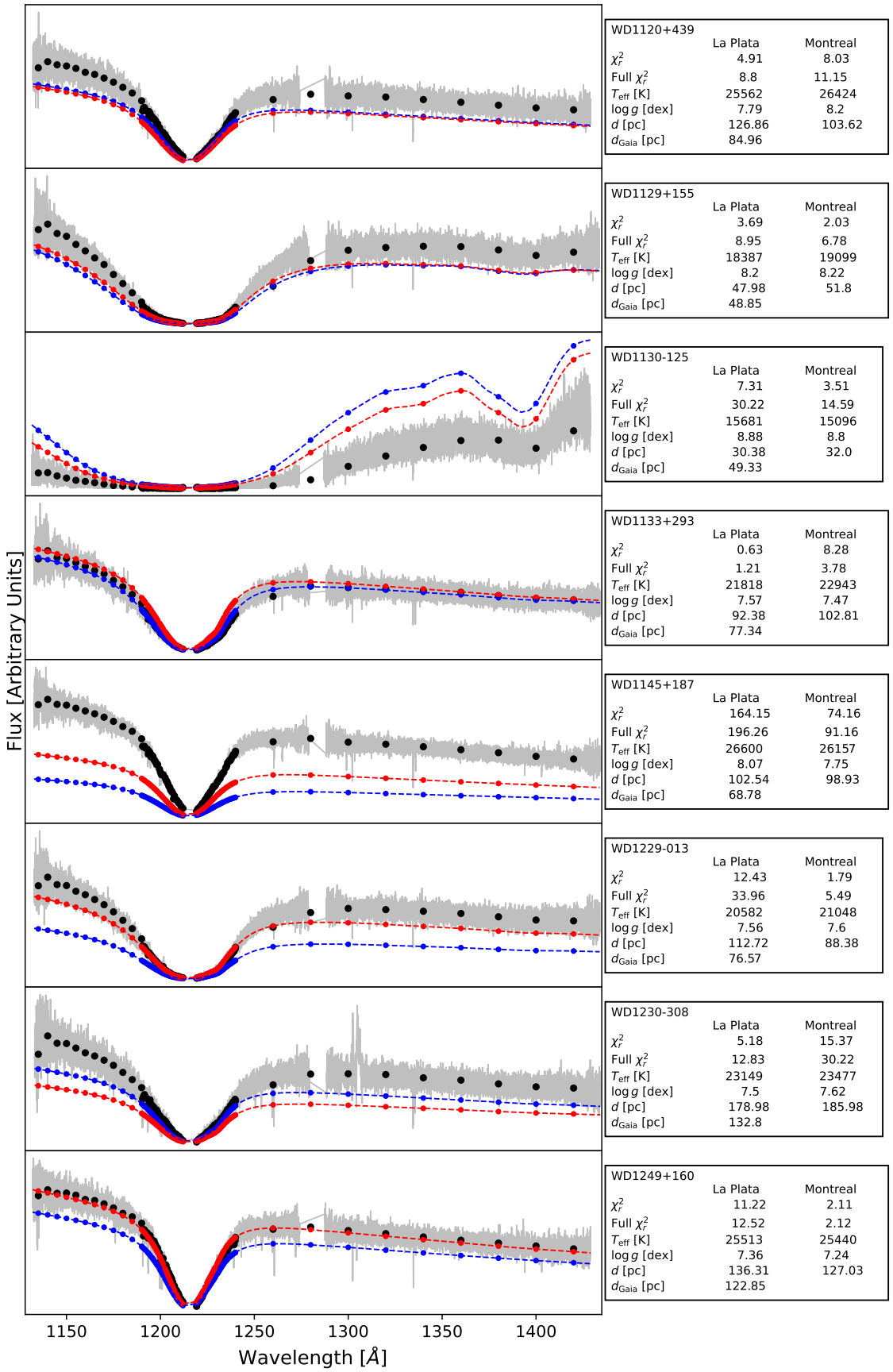


Figure D.15: Same as Fig. D.1.

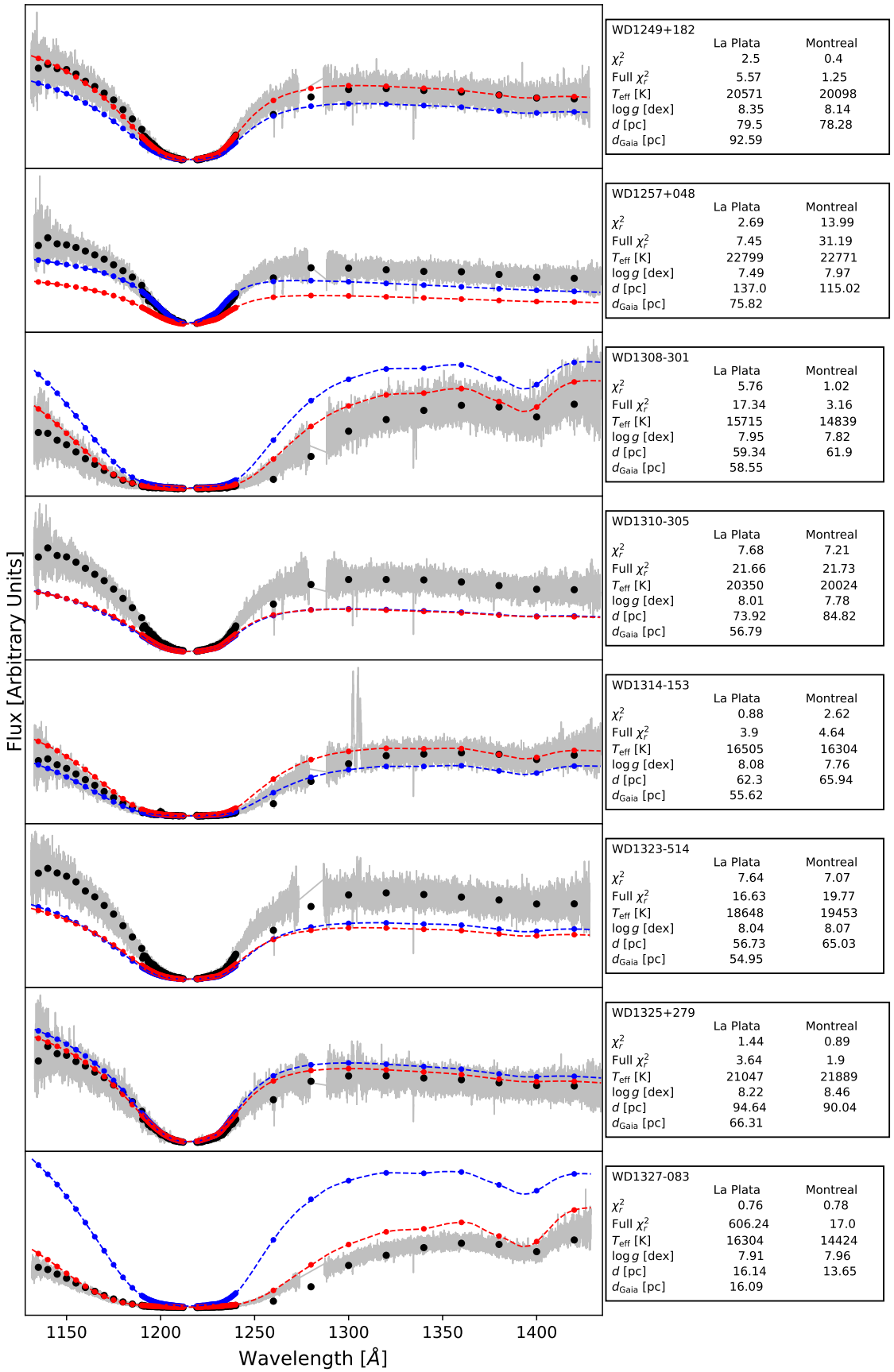


Figure D.16: Same as Fig. D.1.

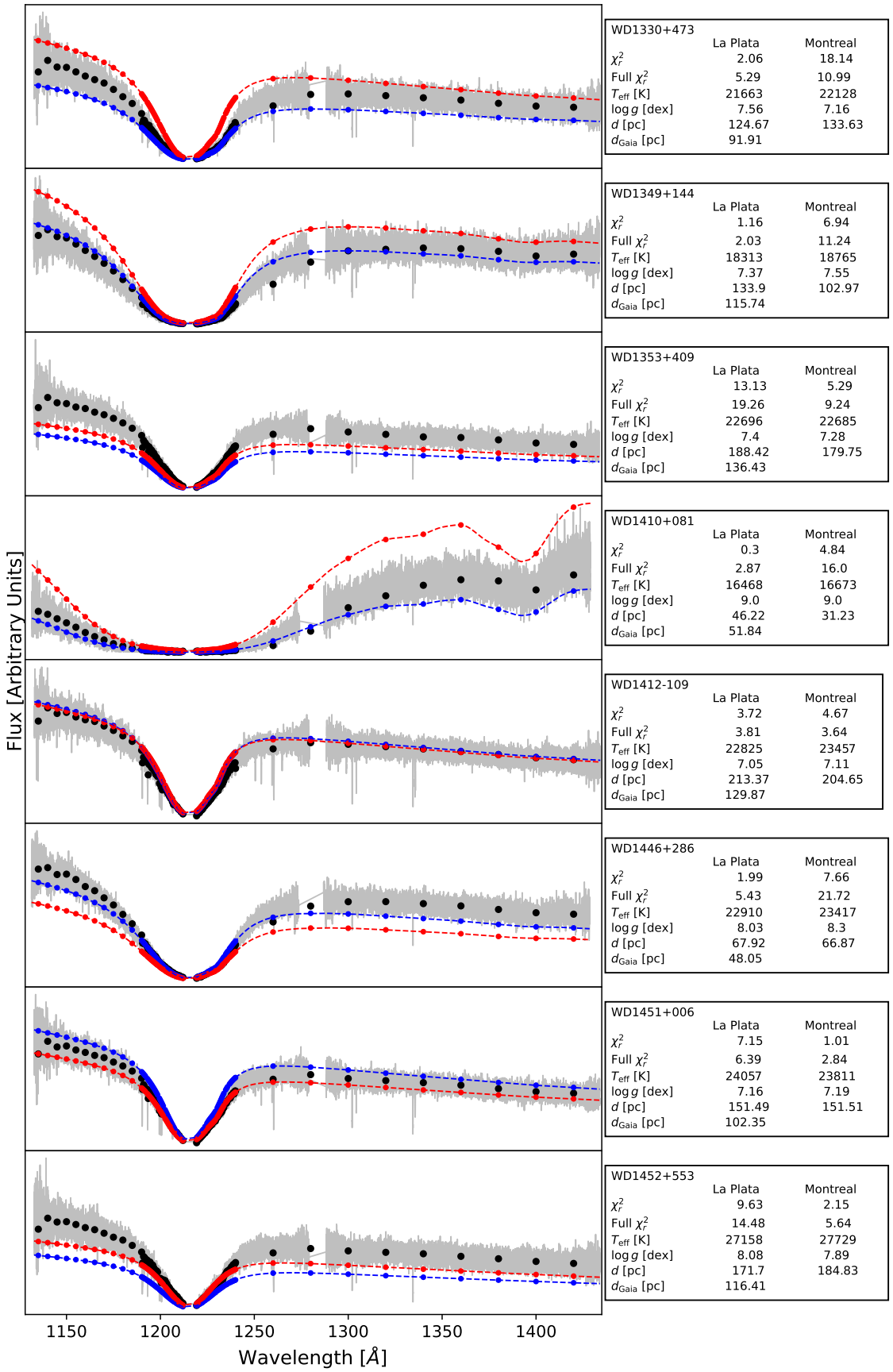
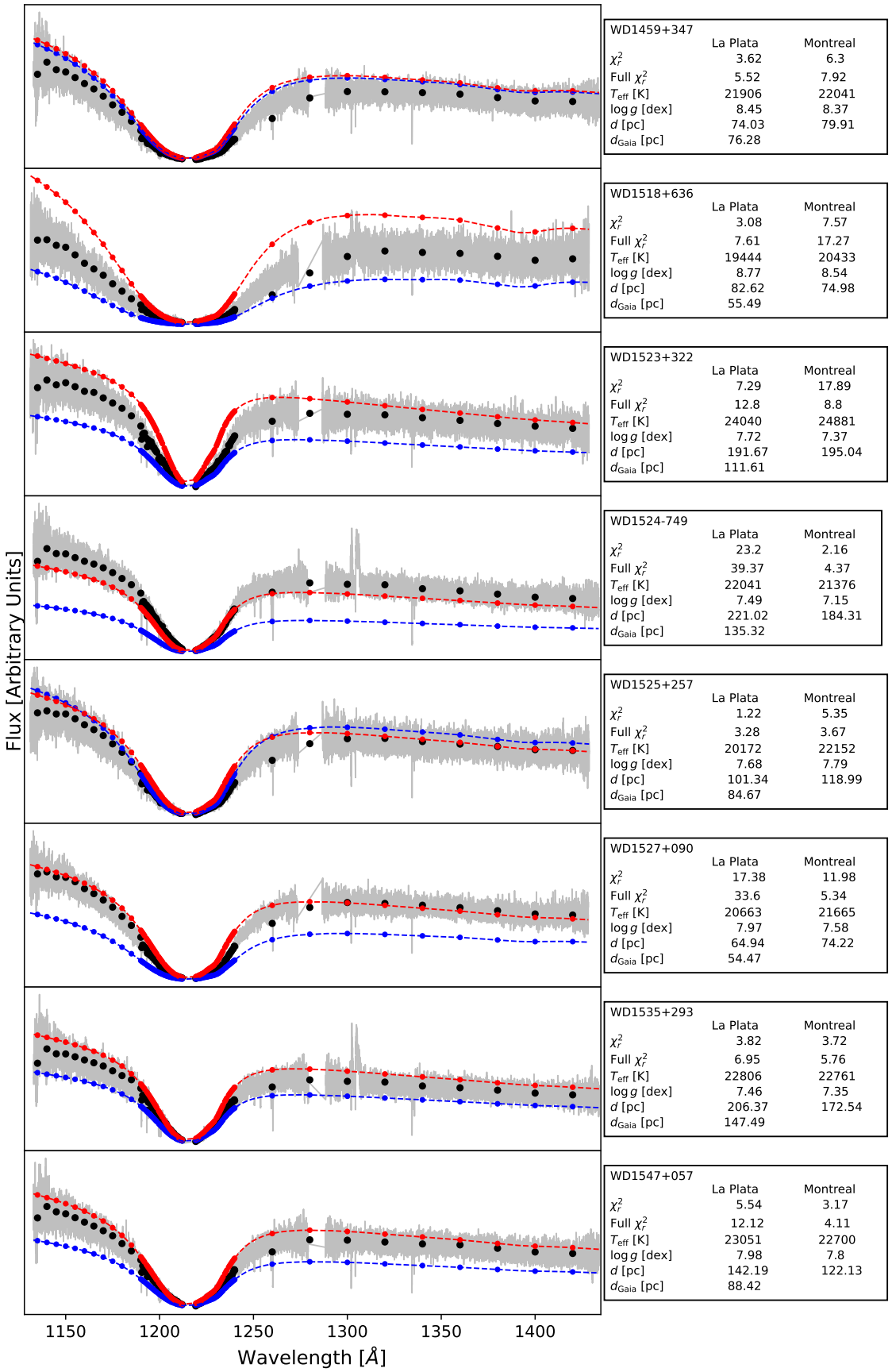


Figure D.17: Same as Fig. D.1.



WD1459+347	La Plata	Montreal
χ^2_r	3.62	6.3
Full χ^2_r	5.52	7.92
$T_{\rm eff}$ [K]	21906	22041
log g [dex]	8.45	8.37
d [pc]	74.03	79.91
$d_{\rm Gaia}$ [pc]	76.28	

WD1518+636	La Plata	Montreal
χ^2_r	3.08	7.57
Full χ^2_r	7.61	17.27
$T_{\rm eff}$ [K]	19444	20433
log g [dex]	8.77	8.54
d [pc]	82.62	74.98
$d_{\rm Gaia}$ [pc]	55.49	

WD1523+322	La Plata	Montreal
χ^2_r	7.29	17.89
Full χ^2_r	12.8	8.8
$T_{\rm eff}$ [K]	24040	24881
log g [dex]	7.72	7.37
d [pc]	191.67	195.04
$d_{\rm Gaia}$ [pc]	111.61	

WD1524-749	La Plata	Montreal
χ^2_r	23.2	2.16
Full χ^2_r	39.37	4.37
$T_{\rm eff}$ [K]	22041	21376
log g [dex]	7.49	7.15
d [pc]	221.02	184.31
$d_{\rm Gaia}$ [pc]	135.32	

WD1525+257	La Plata	Montreal
χ^2_r	1.22	5.35
Full χ^2_r	3.28	3.67
$T_{\rm eff}$ [K]	20172	22152
log g [dex]	7.68	7.79
d [pc]	101.34	118.99
$d_{\rm Gaia}$ [pc]	84.67	

WD1527+090	La Plata	Montreal
χ^2_r	17.38	11.98
Full χ^2_r	33.6	5.34
$T_{\rm eff}$ [K]	20663	21665
log g [dex]	7.97	7.58
d [pc]	64.94	74.22
$d_{\rm Gaia}$ [pc]	54.47	

WD1535+293	La Plata	Montreal
χ^2_r	3.82	3.72
Full χ^2_r	6.95	5.76
$T_{\rm eff}$ [K]	22806	22761
log g [dex]	7.46	7.35
d [pc]	206.37	172.54
$d_{\rm Gaia}$ [pc]	147.49	

WD1547+057	La Plata	Montreal
χ^2_r	5.54	3.17
Full χ^2_r	12.12	4.11
$T_{\rm eff}$ [K]	23051	22700
log g [dex]	7.98	7.8
d [pc]	142.19	122.13
$d_{\rm Gaia}$ [pc]	88.42	

Figure D.18: Same as Fig. D.1.

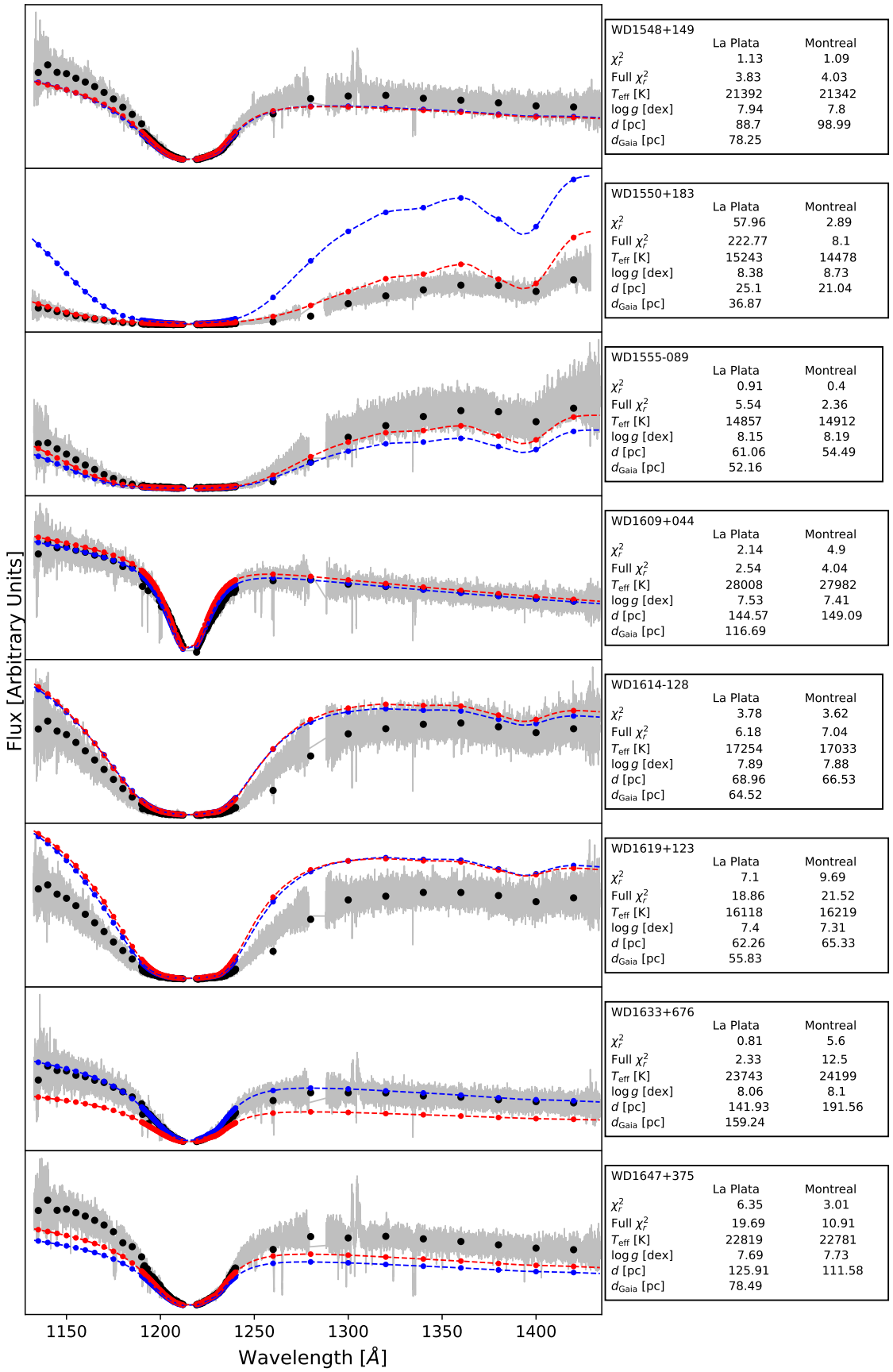
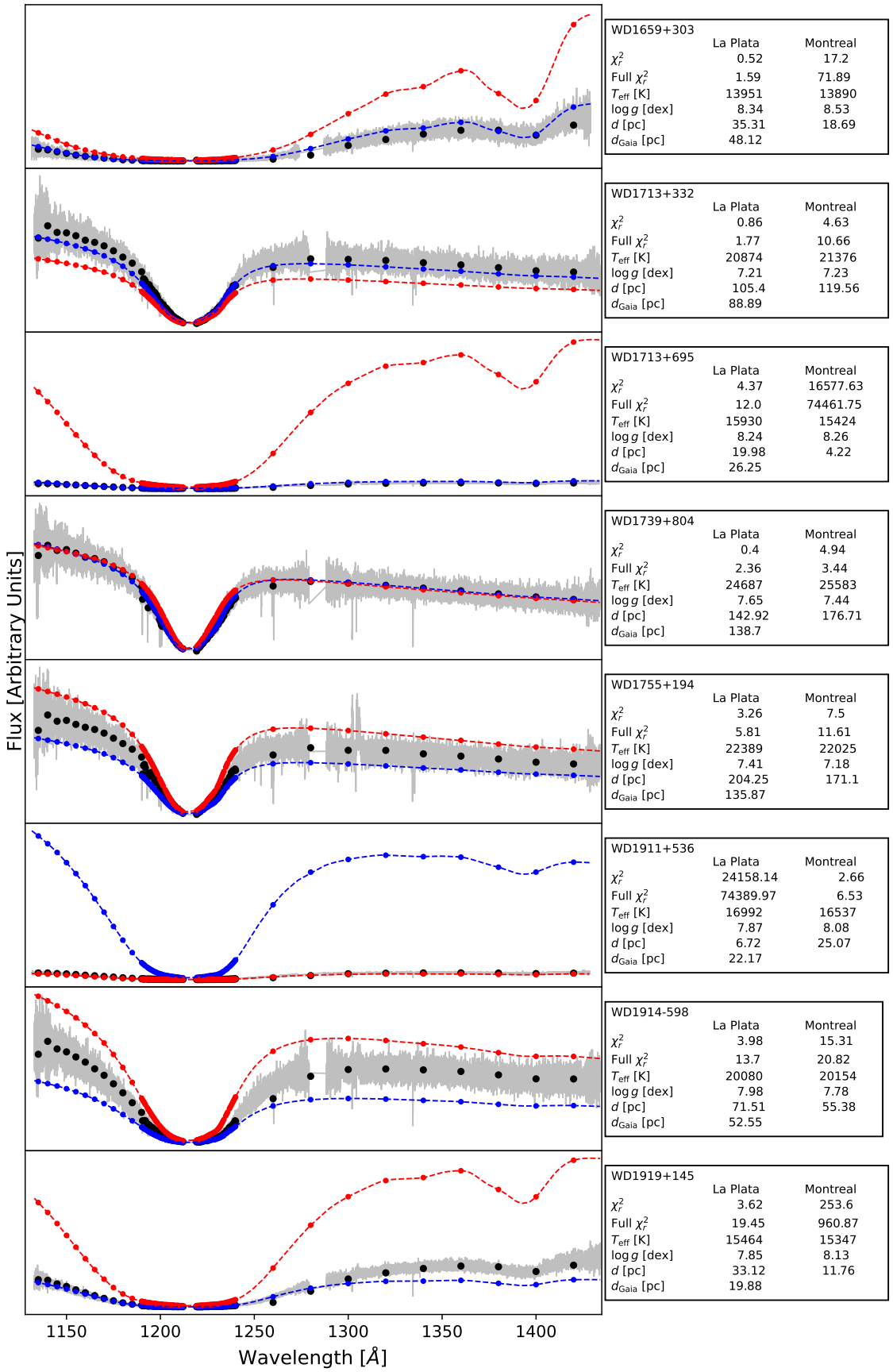


Figure D.19: Same as Fig. D.1.



WD1659+303	La Plata	Montreal
χ^2_r	0.52	17.2
Full χ^2_r	1.59	71.89
$T_{\rm eff}$ [K]	13951	13890
log g [dex]	8.34	8.53
d [pc]	35.31	18.69
$d_{\rm Gaia}$ [pc]	48.12	

WD1713+332	La Plata	Montreal
χ^2_r	0.86	4.63
Full χ^2_r	1.77	10.66
$T_{\rm eff}$ [K]	20874	21376
log g [dex]	7.21	7.23
d [pc]	105.4	119.56
$d_{\rm Gaia}$ [pc]	88.89	

WD1713+695	La Plata	Montreal
χ^2_r	4.37	16577.63
Full χ^2_r	12.0	74461.75
$T_{\rm eff}$ [K]	15930	15424
log g [dex]	8.24	8.26
d [pc]	19.98	4.22
$d_{\rm Gaia}$ [pc]	26.25	

WD1739+804	La Plata	Montreal
χ^2_r	0.4	4.94
Full χ^2_r	2.36	3.44
$T_{\rm eff}$ [K]	24687	25583
log g [dex]	7.65	7.44
d [pc]	142.92	176.71
$d_{\rm Gaia}$ [pc]	138.7	

WD1755+194	La Plata	Montreal
χ^2_r	3.26	7.5
Full χ^2_r	5.81	11.61
$T_{\rm eff}$ [K]	22389	22025
log g [dex]	7.41	7.18
d [pc]	204.25	171.1
$d_{\rm Gaia}$ [pc]	135.87	

WD1911+536	La Plata	Montreal
χ^2_r	24158.14	2.66
Full χ^2_r	74389.97	6.53
$T_{\rm eff}$ [K]	16992	16537
log g [dex]	7.87	8.08
d [pc]	6.72	25.07
$d_{\rm Gaia}$ [pc]	22.17	

WD1914-598	La Plata	Montreal
χ^2_r	3.98	15.31
Full χ^2_r	13.7	20.82
$T_{\rm eff}$ [K]	20080	20154
log g [dex]	7.98	7.78
d [pc]	71.51	55.38
$d_{\rm Gaia}$ [pc]	52.55	

WD1919+145	La Plata	Montreal
χ^2_r	3.62	253.6
Full χ^2_r	19.45	960.87
$T_{\rm eff}$ [K]	15464	15347
log g [dex]	7.85	8.13
d [pc]	33.12	11.76
$d_{\rm Gaia}$ [pc]	19.88	

Figure D.20: Same as Fig. D.1.

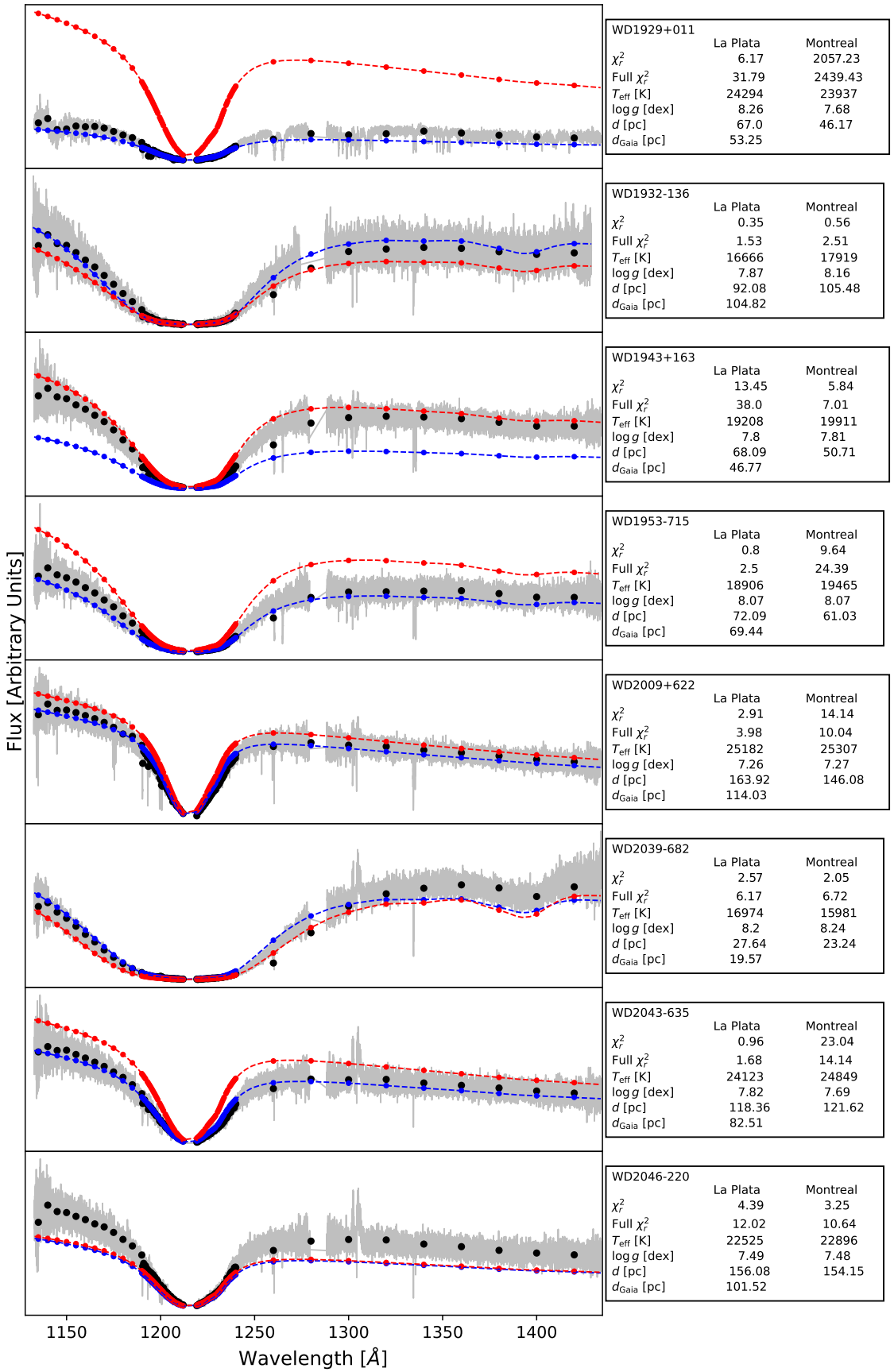


Figure D.21: Same as Fig. D.1.

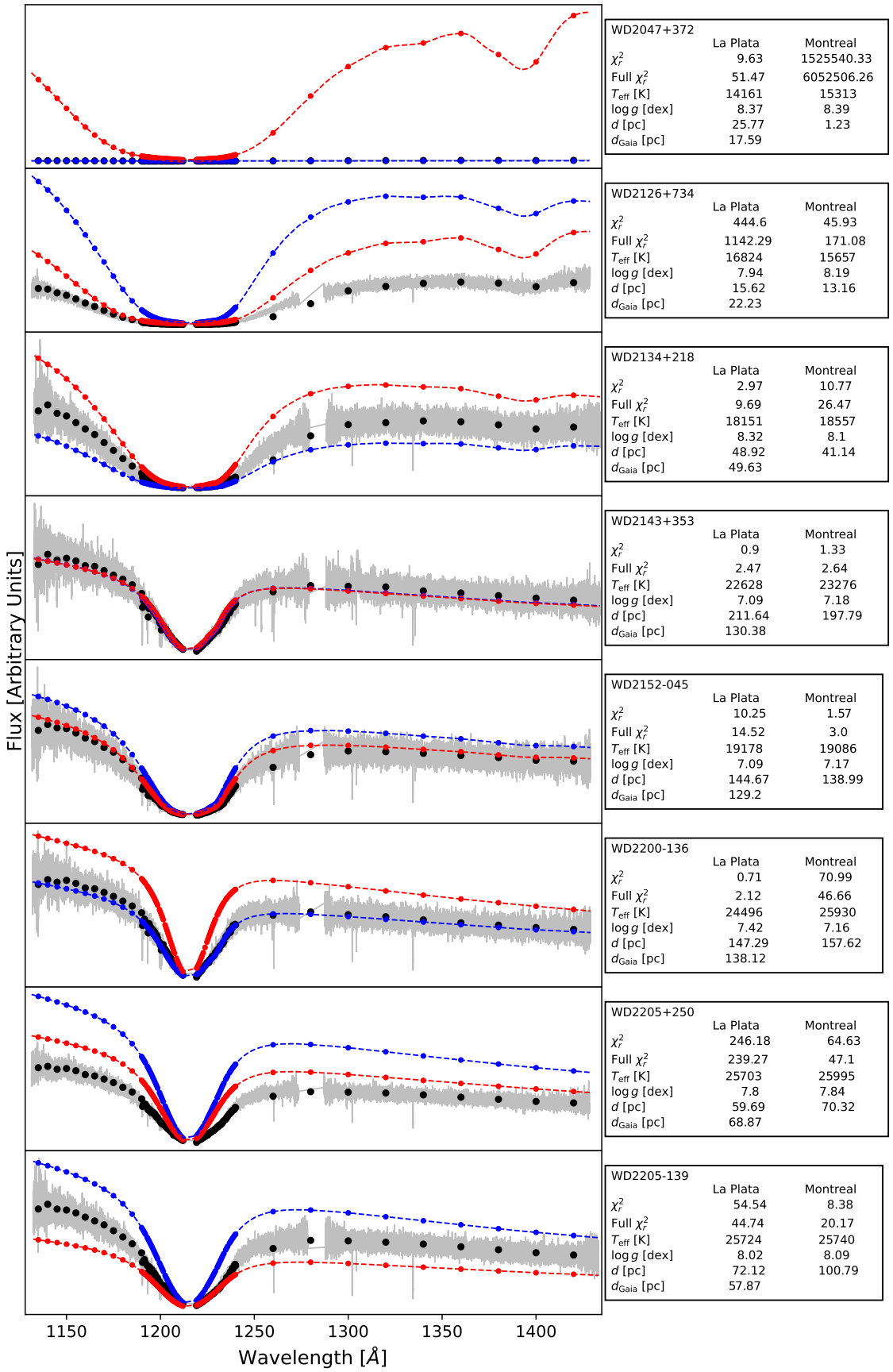


Figure D.22: Same as Fig. D.1.

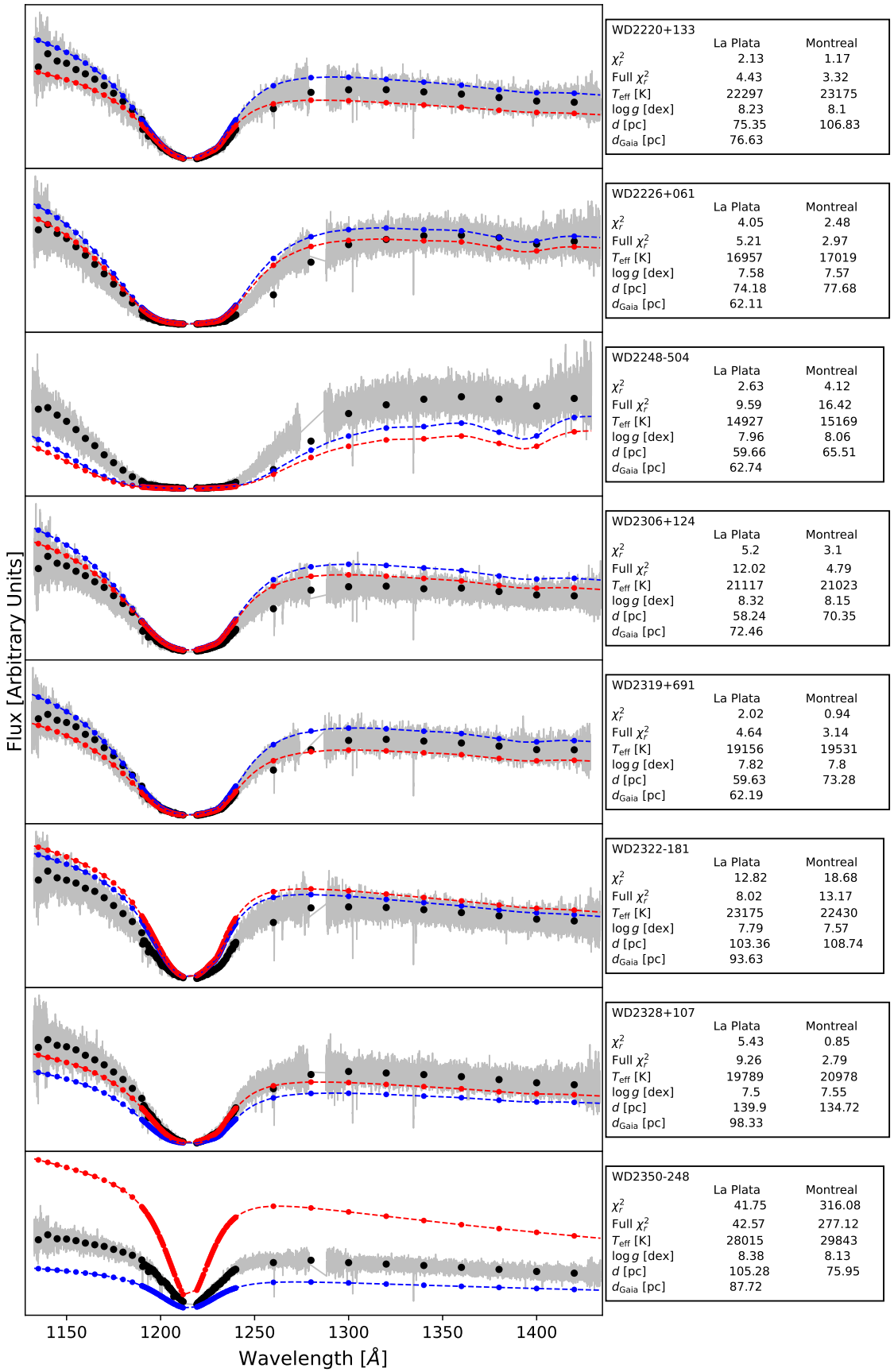


Figure D.23: Same as Fig. D.1.

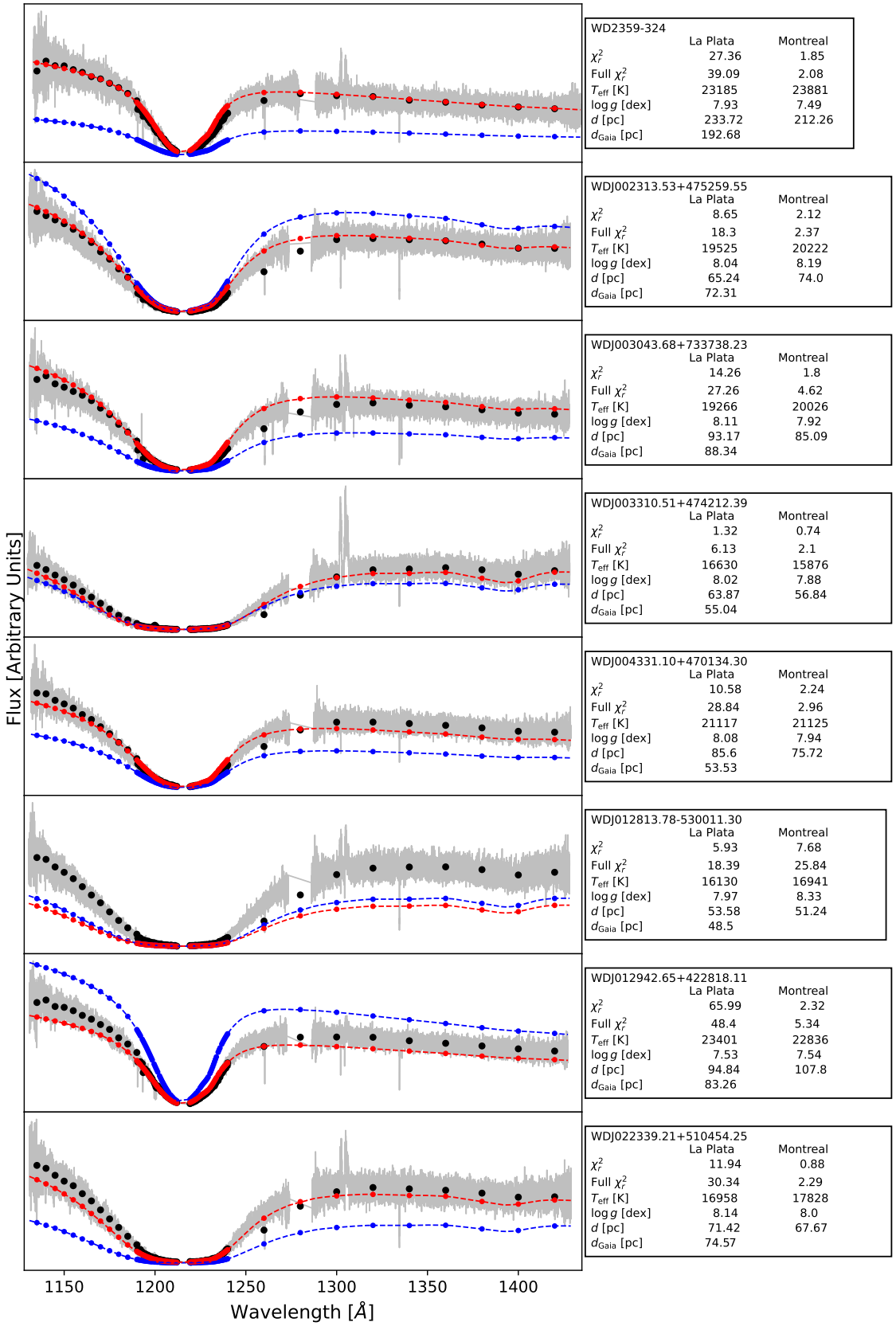


Figure D.24: Same as Fig. D.1.

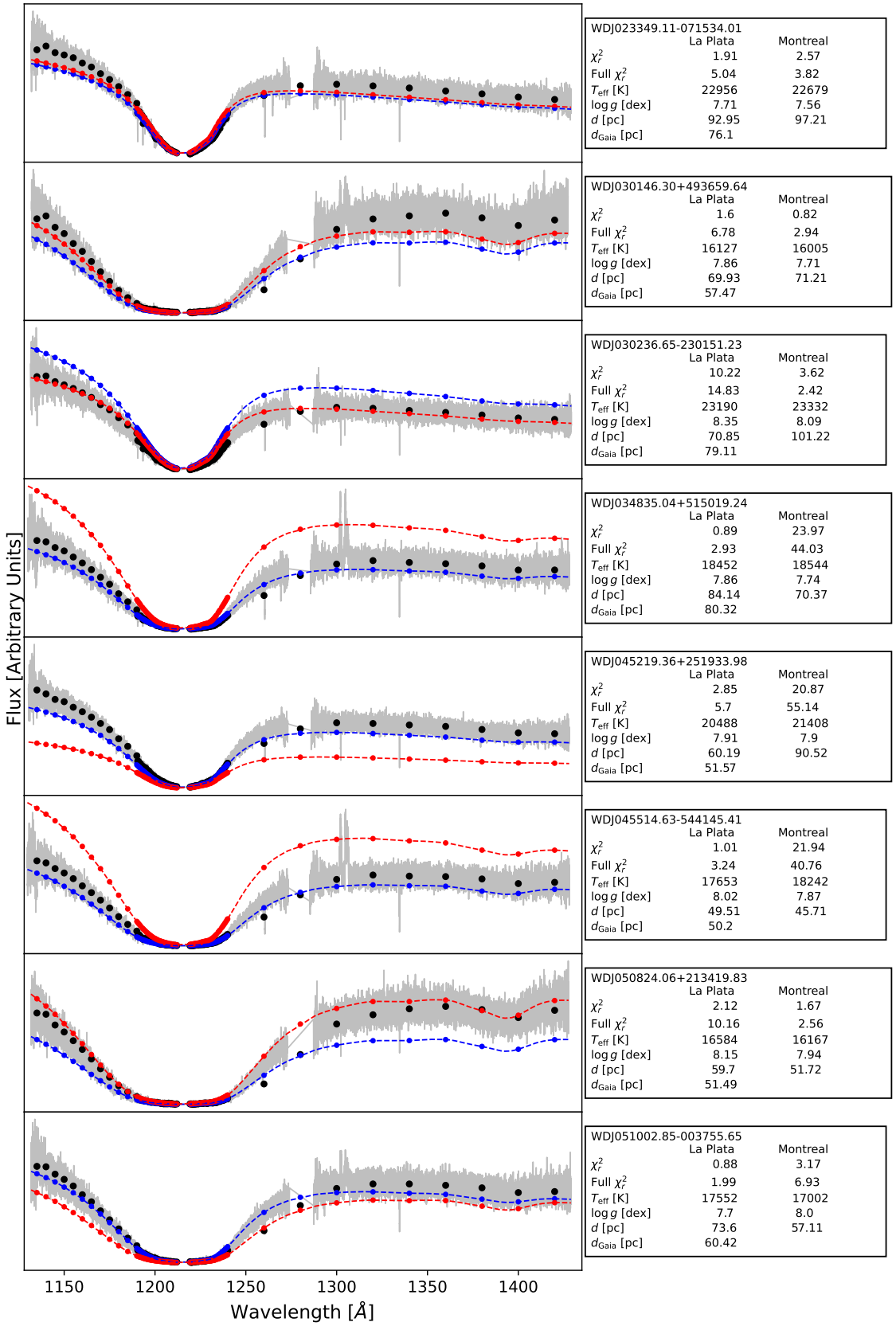


Figure D.25: Same as Fig. D.1.

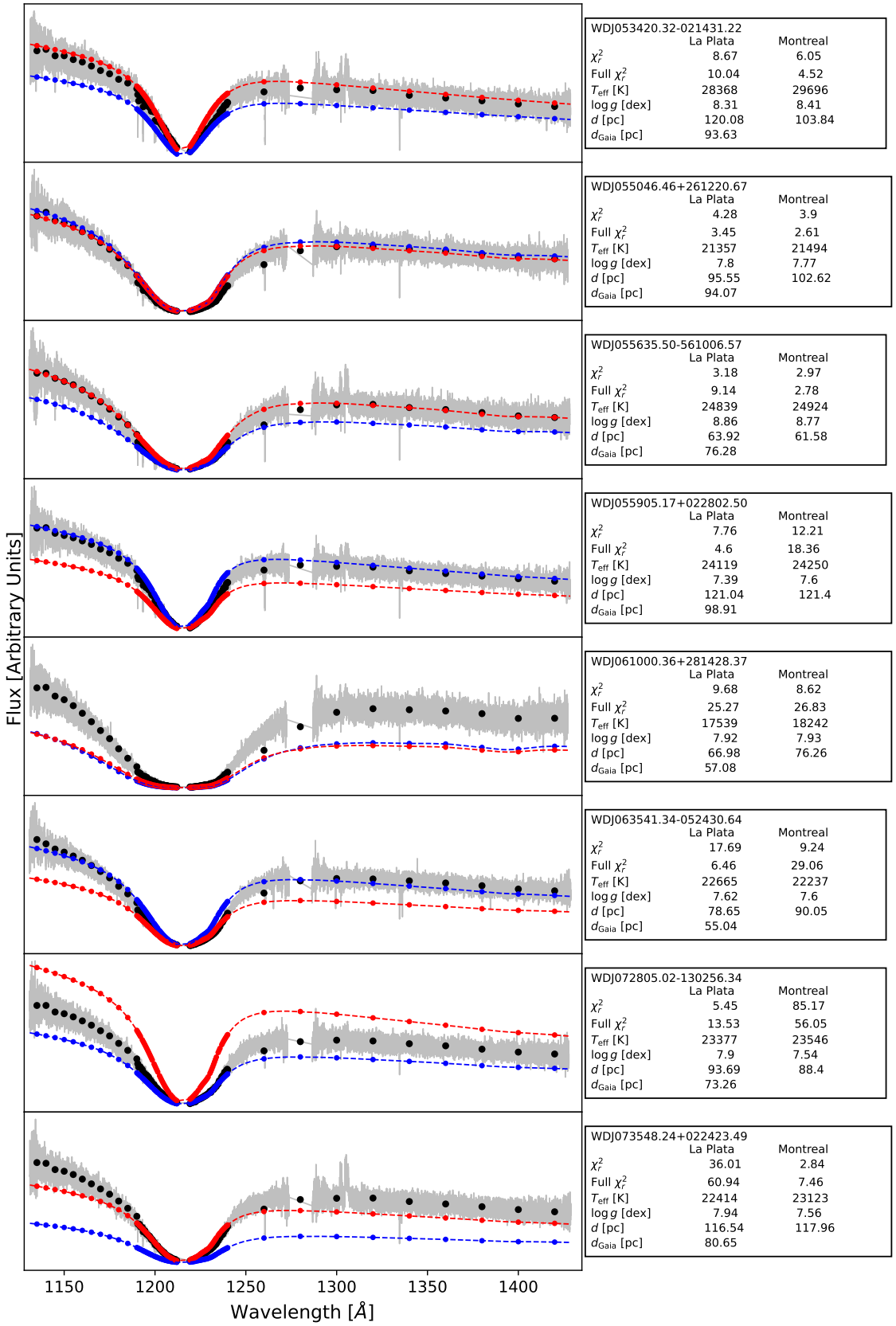


Figure D.26: Same as Fig. D.1.

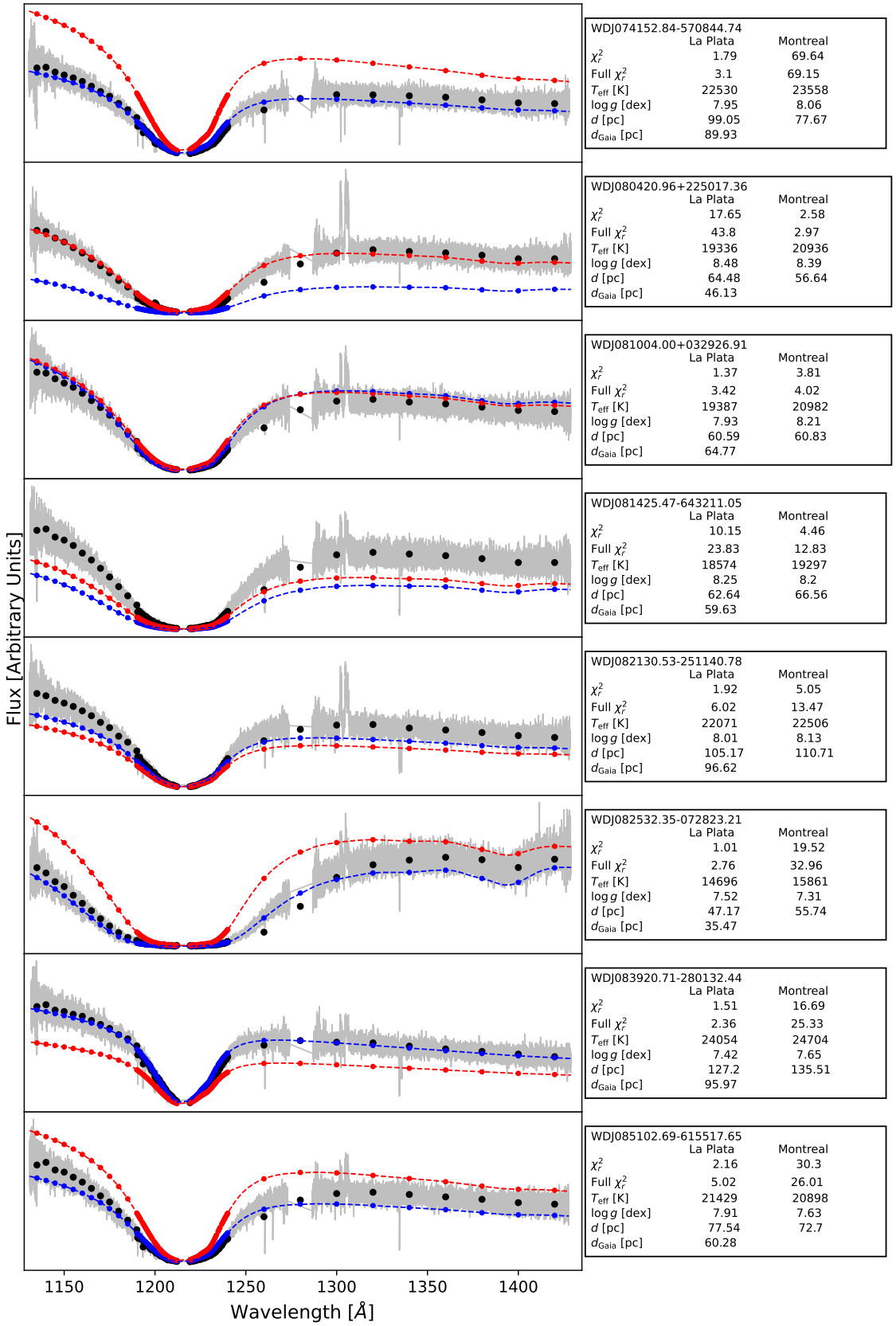


Figure D.27: Same as Fig. D.1.

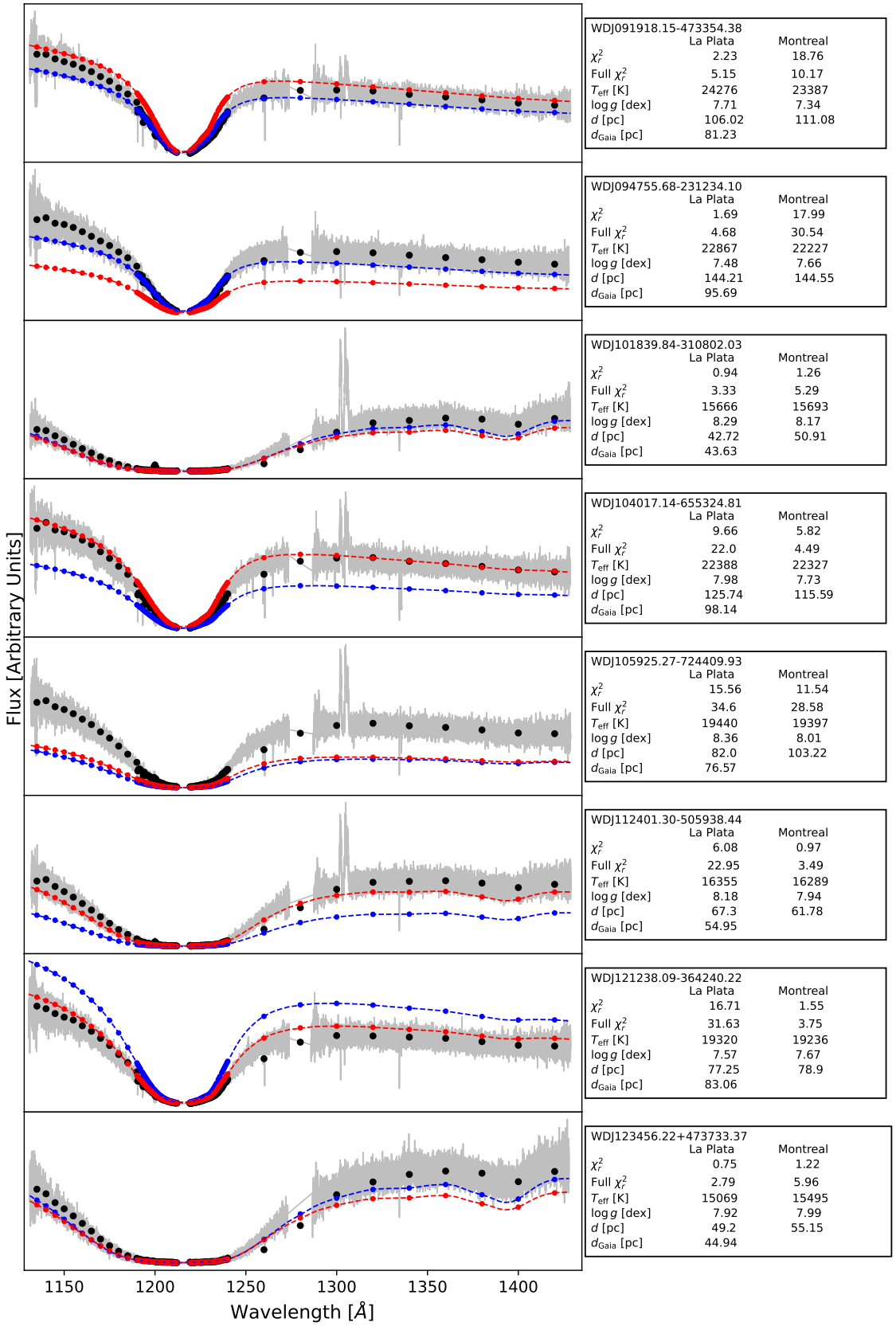


Figure D.28: Same as Fig. D.1.

Appendix E

Targets

Object	$T_{\mathrm{eff,M}}$ [K]	$\log(g)_{\mathrm{M}}$ [dex]	$\mathrm{d_M}$ [pc	$T_{\mathrm{eff,LP}}$ [K]	$\log(g)_{\mathrm{LP}}$ [dex]	$\mathrm{d_{LP}}$ [pc]
APASSJ013001.36+263857.4	14659.0	8.84	18.13	15137.0	9.03	32.79
APASSJ081237.87+173700.3	15308.0	7.85	4.0	15955.0	7.66	32.88
APASSJ083857.48-214611.0	20678.0	7.67	96.75	21139.0	7.73	104.51
APASSJ085913.51-312416.3	27486.0	7.2	236.43	26943.0	7.18	249.57
APASSJ090028.59-090923.2	21116.0	7.56	81.5	20721.0	7.71	81.5
APASSJ145521.26+565544.3	15615.0	8.14	45.11	16131.0	8.08	44.77
APASSJ151754.65+103043.7	20680.0	7.98	129.73	19761.0	7.96	111.55
APASSJ152827.83-251503.0	16480.0	8.85	40.54	16527.0	8.43	44.56
APASSJ195622.94+641358.0	17340.0	7.91	86.96	17907.0	7.9	97.83
APASSJ202336.88-111551.3	16620.0	8.22	45.85	17413.0	8.24	41.6
APASSJ225612.94-131939.3	23343.0	8.87	173.04	24067.0	9.2	189.4
HE0131+0149	14329.0	7.97	50.7	13637.0	7.96	47.29
HE0305-1145	25618.0	7.57	173.59	28663.0	7.53	178.08
HE0308-2305	23330.0	8.25	90.45	23410.0	8.4	95.92
HE0358-5127	24223.0	7.39	168.53	23886.0	7.8	150.44
HE0403-4129	23123.0	8.22	165.05	23284.0	8.19	133.19
HE0414-4039	22034.0	8.1	130.85	20830.0	8.02	124.42
HE0416-1034	24142.0	7.44	130.16	24793.0	7.61	136.97
HE0418-1021	23876.0	8.1	97.62	24123.0	8.28	109.92
HE0418-5326	27311.0	7.64	205.02	29001.0	7.82	177.88
HE0452-3444	20810.0	7.92	127.15	20662.0	7.83	123.66

HE1247-1130	25793.0	7.45	150.84	25347.0	7.42	171.55
HE1518-0020	15278.0	7.78	74.96	15510.0	7.81	72.03
HE2218-2706	14109.0	8.37	30.93	13243.0	8.58	44.35
HE2231-2647	21736.0	7.96	91.5	21472.0	8.09	92.53
HE2238-0433	20568.0	9.0	98.57	20922.0	9.1	91.75
HE2345-4810	27125.0	7.51	211.81	28588.0	7.35	230.61
HS0002+1635	25278.0	7.51	131.85	23722.0	7.63	143.0
HS0200+2449	21815.0	7.56	138.41	21554.0	7.52	146.62
HS0400+1451	14551.0	8.47	35.2	14824.0	8.57	40.39
HS0507+0434A	22275.0	8.04	74.87	21084.0	8.26	54.47
HS0944+1913	16162.0	8.41	56.2	16079.0	8.03	58.16
HS1243+0132	22358.0	8.34	123.03	21718.0	8.15	150.67
HS1334+0701	16404.0	7.4	110.91	16257.0	7.51	97.52
HS2056+0721	25417.0	7.86	125.14	26251.0	7.66	150.75
HS2210+2323	21411.0	7.3	121.83	21737.0	7.57	131.99
HS2220+2146A	23648.0	9.22	54.02	25138.0	9.35	84.55
HS2220+2146B	19876.0	8.66	68.41	19171.0	8.33	69.61
HS2225+2158	25393.0	7.56	165.25	25233.0	7.57	149.11
HS2229+2335	18919.0	7.93	87.28	18456.0	7.91	103.22
HS2244+2103	23772.0	7.62	170.66	23889.0	7.63	168.18
PG0004+061	23145.0	8.06	169.8	22325.0	7.83	155.93
PG0816+297	18743.0	8.3	76.74	17847.0	8.11	84.85
PG0817+386	25775.0	7.62	149.3	25454.0	7.71	159.87
PG0821+633	18235.0	8.07	69.29	17524.0	8.08	74.4
PG0846+558	26168.0	7.67	248.44	25832.0	7.46	227.55
PG0915+526	17439.0	8.54	54.3	17059.0	8.33	53.81
PG1113+413	24791.0	7.47	134.12	25224.0	7.46	124.01
PG1126+384	25075.0	7.59	105.5	23877.0	7.52	118.28
PG1128+565	27130.0	7.7	205.22	24180.0	7.6	189.83
PG1143+321	16223.0	8.11	32.14	16419.0	7.82	45.37
PG1202+309	28285.0	7.81	197.82	29019.0	7.42	206.81
PG1220+234	23899.0	7.97	127.3	23514.0	7.42	156.44
PG1339+346	16399.0	7.79	103.09	16802.0	8.06	112.99
PG1421+318	26569.0	7.54	142.26	25833.0	7.6	138.46
PG1508+549	18363.0	8.04	79.99	17857.0	8.04	78.34
PG1513+442	27756.0	7.19	180.42	26493.0	7.24	173.84

PG1601+581	15716.0	8.14	45.22	15204.0	8.09	42.17
PG1620+260	26996.0	7.4	173.19	25577.0	7.21	193.16
PG1641+388	15507.0	7.65	60.92	15373.0	7.9	56.14
PG2345+305	25627.0	7.07	191.94	25559.0	7.18	221.44
SDSSJ081305.55+140317.4	20710.0	8.24	96.39	20167.0	8.47	89.51
SDSSJ170029.93+422452.99	25735.0	8.68	216.65	25342.0	8.4	131.63
WD0000+171	22076.0	7.88	93.45	21056.0	7.78	96.22
WD0013-241	18978.0	8.39	57.45	19680.0	8.16	73.36
WD0018-339	22066.0	8.02	78.59	21376.0	7.72	81.23
WD0022-745	28153.0	8.39	126.71	28198.0	7.88	132.23
WD0028-474	17717.0	7.59	95.14	17630.0	7.71	89.55
WD0034-602	14970.0	8.4	0.36	14488.0	8.69	16.75
WD0047-524	19684.0	7.88	53.64	19666.0	7.83	71.36
WD0048+202	22030.0	8.42	68.92	21120.0	8.14	69.53
WD0052-147	25597.0	7.86	100.94	25159.0	7.81	99.47
WD0059+257	21325.0	8.27	104.34	21235.0	8.26	90.98
WD0102+095	24346.0	7.55	103.51	22927.0	7.31	101.46
WD0106-358	28985.0	7.93	96.63	30052.0	7.85	79.45
WD0114-605	24332.0	7.49	134.76	24838.0	7.47	137.77
WD0124-257	22870.0	7.36	183.5	22558.0	7.3	187.52
WD0128-387	12652.0	8.69	24.19	11286.0	8.42	30.66
WD0136+768	15841.0	8.5	60.88	16448.0	8.27	62.27
WD0140-392	23319.0	7.6	80.37	22555.0	7.92	73.04
WD0155+069	22071.0	7.47	116.5	21616.0	7.75	117.87
WD0216+143	26988.0	7.63	125.75	27482.0	7.57	118.9
WD0220+222	15803.0	8.21	83.06	15438.0	7.96	74.91
WD0231-054	14770.0	8.91	0.05	15564.0	8.66	24.29
WD0232+525	16748.0	7.98	16.68	16776.0	8.08	24.31
WD0242-174	21185.0	7.8	108.97	19902.0	7.77	103.83
WD0307+149	21946.0	8.04	69.1	22638.0	7.92	81.32
WD0308+188	19115.0	7.97	46.42	18572.0	8.0	31.11
WD0316+345	15892.0	7.35	63.08	15298.0	7.39	60.6
WD0321-026	23203.0	7.33	324.99	23054.0	7.56	279.47
WD0341+021	21620.0	7.21	177.43	22216.0	7.21	164.31
WD0348+339	15078.0	8.77	29.57	13976.0	8.58	37.78
WD0352+018	22323.0	7.69	114.87	21971.0	7.93	122.92

WD0406+169	15188.0	8.44	51.47	16160.0	8.04	48.48
WD0410+117	20867.0	7.71	75.73	20964.0	7.91	63.11
WD0416+334	16864.0	7.56	77.73	16575.0	7.69	73.52
WD0421+162	20506.0	8.18	50.45	19879.0	8.28	49.8
WD0431+126	22536.0	8.27	61.56	21009.0	8.38	54.92
WD0437+152	18937.0	7.54	135.12	18511.0	7.51	150.01
WD0525+526	20449.0	8.73	44.41	20642.0	8.66	61.49
WD0556+172	18311.0	7.4	108.93	18569.0	7.89	96.14
WD0558+165	17625.0	8.86	47.09	16402.0	8.64	51.66
WD0625+415	17777.0	8.33	63.89	16989.0	8.29	47.64
WD0701-587	14255.0	8.41	13.52	14887.0	8.47	35.66
WD0730+487	15206.0	8.93	1.24	16740.0	8.56	26.9
WD0732-427	16039.0	8.82	57.16	13906.0	8.84	58.39
WD0743+442	15019.0	8.77	25.09	15467.0	8.61	36.33
WD0808+435	22221.0	8.55	65.77	20841.0	8.53	65.34
WD0839+231	25186.0	7.19	84.51	25106.0	7.37	116.27
WD0843+516	24339.0	8.71	151.86	23152.0	8.39	129.76
WD0854+404	23773.0	7.51	93.77	23295.0	7.9	81.85
WD0904+391	24896.0	7.54	182.56	24657.0	7.98	170.44
WD0920+363	23457.0	7.17	219.06	23362.0	7.57	212.62
WD0922+183	23813.0	7.98	164.41	23596.0	8.1	141.71
WD0933+025	21189.0	7.81	158.64	20782.0	7.76	142.32
WD0933+729	17155.0	7.82	78.47	17931.0	7.8	84.35
WD0947+325	22025.0	8.27	95.29	22307.0	8.19	91.4
WD0954+697	22002.0	7.86	152.0	21357.0	7.68	133.92
WD0956+020	16282.0	7.72	93.56	16449.0	8.0	77.24
WD1005+642	19725.0	7.54	61.67	20595.0	8.21	47.07
WD1013+256	21015.0	7.82	152.62	20543.0	7.8	160.99
WD1015+161	19966.0	8.57	81.93	20251.0	8.38	57.21
WD1017+125	20285.0	7.6	143.26	20025.0	7.62	129.07
WD1020-207	20544.0	8.06	71.01	20505.0	8.1	59.78
WD1034+492	20000.0	8.21	85.19	20563.0	8.47	82.75
WD1038+633	24668.0	8.14	83.91	23666.0	7.96	90.33
WD1049-158	19751.0	8.36	41.92	19476.0	8.3	54.87
WD1052+273	21979.0	8.09	61.84	22278.0	7.84	51.82
WD1058-129	24826.0	7.96	70.35	24812.0	8.16	81.77

WD1102-281	19927.0	8.53	58.64	19811.0	8.76	42.62
WD1102+748	18806.0	9.12	82.29	18402.0	8.77	74.76
WD1103+384	25964.0	8.36	206.46	25863.0	8.44	143.32
WD1104+602	20005.0	7.97	39.67	18677.0	8.23	32.33
WD1115+166	29428.0	9.03	40.36	28173.0	9.19	49.27
WD1120+439	26424.0	8.2	103.62	25562.0	7.79	126.86
WD1129+155	19099.0	8.22	51.8	18388.0	8.2	47.98
WD1130-125	15097.0	8.8	32.0	15682.0	8.88	30.38
WD1133+293	22944.0	7.47	102.81	21818.0	7.57	92.38
WD1145+187	26157.0	7.75	98.93	26601.0	8.07	102.54
WD1229-013	21048.0	7.6	88.38	20583.0	7.56	112.72
WD1230-308	23477.0	7.62	185.98	23149.0	7.5	178.98
WD1249+160	25441.0	7.24	127.03	25513.0	7.36	136.31
WD1249+182	20098.0	8.14	78.28	20572.0	8.35	79.5
WD1257+048	22771.0	7.97	115.02	22800.0	7.49	137.0
WD1258+593	15108.0	8.08	61.87	16028.0	7.61	61.88
WD1308-301	14840.0	7.82	61.9	15715.0	7.95	59.34
WD1310-305	20025.0	7.78	84.82	20351.0	8.01	73.92
WD1314-153	16305.0	7.76	65.94	16505.0	8.08	62.3
WD1323-514	19454.0	8.07	65.03	18648.0	8.04	56.73
WD1325-089	21890.0	8.46	90.04	21048.0	8.22	94.64
WD1325+279	17464.0	8.1	56.95	18368.0	8.12	71.55
WD1327-083	14424.0	7.96	13.65	16305.0	7.91	16.14
WD1330+473	22129.0	7.16	133.63	21663.0	7.56	124.67
WD1334-160	18433.0	8.45	54.86	18898.0	8.59	44.53
WD1349+144	18765.0	7.55	102.97	18313.0	7.37	133.9
WD1353+409	22685.0	7.28	179.75	22696.0	7.4	188.42
WD1408+323	19379.0	8.07	36.79	19400.0	8.26	39.81
WD1410+081	16673.0	9.05	31.23	16469.0	8.95	46.22
WD1412-109	23457.0	7.11	204.65	22826.0	7.05	213.37
WD1446+286	23417.0	8.3	66.87	22910.0	8.03	67.92
WD1449+168	21329.0	7.45	107.01	22060.0	7.75	110.22
WD1449+513	17665.0	8.72	81.99	17000.0	8.73	83.73
WD1451+006	23811.0	7.19	151.51	24058.0	7.16	151.49
WD1452+553	27730.0	7.89	184.83	27158.0	8.08	171.7
WD1459+347	22041.0	8.37	79.91	21906.0	8.45	74.03

WD1507+220	19362.0	7.79	70.93	19140.0	7.76	75.76
WD1518+636	20434.0	8.54	74.98	19444.0	8.77	82.62
WD1523+322	24881.0	7.37	195.04	24040.0	7.72	191.67
WD1524-749	21377.0	7.15	184.31	22041.0	7.49	221.02
WD1525+257	22152.0	7.79	118.99	20173.0	7.68	101.34
WD1527+090	21665.0	7.58	74.22	20664.0	7.97	64.94
WD1531-022	19482.0	8.25	36.86	19092.0	8.3	34.57
WD1535+293	22761.0	7.35	172.54	22807.0	7.46	206.37
WD1547+057	22700.0	7.8	122.13	23052.0	7.98	142.19
WD1548+149	21342.0	7.8	98.99	21392.0	7.94	88.7
WD1550+183	14478.0	8.73	21.04	15243.0	8.38	25.1
WD1555-089	14913.0	8.19	54.49	14857.0	8.15	61.06
WD1609+044	27983.0	7.41	149.09	28008.0	7.53	144.57
WD1614-128	17034.0	7.88	66.53	17255.0	7.89	68.96
WD1619+123	16220.0	7.31	65.33	16118.0	7.4	62.26
WD1633+676	24199.0	8.1	191.56	23744.0	8.06	141.93
WD1647+375	22782.0	7.73	111.58	22820.0	7.69	125.91
WD1659+303	13890.0	8.53	18.69	13951.0	8.34	35.31
WD1713+332	21376.0	7.23	119.56	20875.0	7.21	105.4
WD1713+695	15425.0	8.26	4.22	15931.0	8.24	19.98
WD1739+804	25584.0	7.44	176.71	24687.0	7.65	142.92
WD1755+194	22025.0	7.18	171.1	22390.0	7.41	204.25
WD1911+536	16538.0	8.08	25.07	16992.0	7.87	6.72
WD1914-598	20154.0	7.78	55.38	20081.0	7.98	71.51
WD1919+145	15347.0	8.13	11.76	15464.0	7.85	33.12
WD1929+011	23937.0	7.68	46.17	24294.0	8.26	67.0
WD1932-136	17919.0	8.16	105.48	16667.0	7.87	92.08
WD1943+163	19912.0	7.81	50.71	19208.0	7.8	68.09
WD1953-715	19466.0	8.07	61.03	18907.0	8.07	72.09
WD2009+622	25308.0	7.27	146.08	25182.0	7.26	163.92
WD2018-233	16273.0	8.42	47.05	16323.0	8.4	56.4
WD2021-128	20530.0	7.57	96.7	21147.0	7.81	90.96
WD2032+188	19473.0	7.85	100.36	19400.0	7.87	71.12
WD2039-682	15981.0	8.24	23.24	16975.0	8.2	27.64
WD2043-635	24849.0	7.69	121.62	24123.0	7.82	118.36
WD2046-220	22897.0	7.48	154.15	22526.0	7.49	156.08

WD2047+372	15313.0	8.39	1.23	14161.0	8.37	25.77
WD2058+181	17079.0	8.2	59.17	18450.0	8.23	65.67
WD2126+734	15657.0	8.19	13.16	16824.0	7.94	15.62
WD2134+218	18558.0	8.1	41.14	18151.0	8.32	48.92
WD2143+353	23276.0	7.18	197.79	22629.0	7.09	211.64
WD2152-045	19086.0	7.17	138.99	19179.0	7.09	144.67
WD2200-136	25931.0	7.16	157.62	24497.0	7.42	147.29
WD2205-139	25995.0	7.84	70.32	25703.0	7.8	59.69
WD2205+250	25740.0	8.09	100.79	25725.0	8.02	72.12
WD2220+133	23176.0	8.1	106.83	22298.0	8.23	75.35
WD2226+061	17020.0	7.57	77.68	16958.0	7.58	74.18
WD2248-504	15169.0	8.06	65.51	14927.0	7.96	59.66
WD2306+124	21024.0	8.15	70.35	21117.0	8.32	58.24
WD2319+691	19531.0	7.8	73.28	19156.0	7.82	59.63
WD2322-181	22430.0	7.57	108.74	23175.0	7.79	103.36
WD2328+107	20979.0	7.55	134.72	19790.0	7.5	139.9
WD2341+322	13694.0	8.38	0.32	14099.0	8.5	20.28
WD2350-248	29843.0	8.13	75.95	28015.0	8.38	105.28
WD2359-324	23882.0	7.49	212.26	23186.0	7.93	233.72
WDJ002313.53+475259.55	20222.0	8.19	74.0	19525.0	8.04	65.24
WDJ003043.68+733738.23	20026.0	7.92	85.09	19266.0	8.11	93.17
WDJ003310.51+474212.39	15877.0	7.88	56.84	16630.0	8.02	63.87
WDJ004331.10+470134.30	21126.0	7.94	75.72	21118.0	8.08	85.6
WDJ012813.78-530011.30	16941.0	8.33	51.24	16130.0	7.97	53.58
WDJ012942.65+422818.11	22837.0	7.54	107.8	23402.0	7.53	94.84
WDJ015630.05+295532.28	15731.0	8.67	32.88	16292.0	8.34	52.6
WDJ022339.21+510454.25	17829.0	8.0	67.67	16958.0	8.14	71.42
WDJ023349.11-071534.01	22679.0	7.56	97.21	22957.0	7.71	92.95
WDJ030146.30+493659.64	16006.0	7.71	71.21	16127.0	7.86	69.93
WDJ030236.65-230151.23	23333.0	8.09	101.22	23190.0	8.35	70.85
WDJ034835.04+515019.24	18544.0	7.74	70.37	18453.0	7.86	84.14
WDJ045219.36+251933.98	21408.0	7.9	90.52	20488.0	7.91	60.19
WDJ045514.63-544145.41	18243.0	7.87	45.71	17653.0	8.02	49.51
WDJ050824.06+213419.83	16167.0	7.94	51.72	16584.0	8.15	59.7
WDJ051002.85-003755.65	17003.0	8.0	57.11	17552.0	7.7	73.6
WDJ053420.32-021431.22	29697.0	8.41	103.84	28369.0	8.31	120.08

WDJ055046.46+261220.67	21495.0	7.77	102.62	21357.0	7.8	95.55
WDJ055635.50-561006.57	24924.0	8.77	61.58	24839.0	8.86	63.92
WDJ055905.17+022802.50	24251.0	7.6	121.4	24120.0	7.39	121.04
WDJ061000.36+281428.37	18243.0	7.93	76.26	17539.0	7.92	66.98
WDJ063541.34-052430.64	22237.0	7.6	90.05	22665.0	7.62	78.65
WDJ072805.02-130256.34	23546.0	7.54	88.4	23377.0	7.9	93.69
WDJ073548.24+022423.49	23124.0	7.56	117.96	22414.0	7.94	116.54
WDJ074152.84-570844.74	23558.0	8.06	77.67	22530.0	7.95	99.05
WDJ080420.96+225017.36	20937.0	8.39	56.64	19337.0	8.48	64.48
WDJ081004.00+032926.91	20983.0	8.21	60.83	19388.0	7.93	60.59
WDJ081425.47-643211.05	19297.0	8.2	66.56	18574.0	8.25	62.64
WDJ082130.53-251140.78	22507.0	8.13	110.71	22071.0	8.01	105.17
WDJ082532.35-072823.21	15862.0	7.31	55.74	14696.0	7.52	47.17
WDJ083920.71-280132.44	24704.0	7.65	135.51	24054.0	7.42	127.2
WDJ085102.69-615517.65	20899.0	7.63	72.7	21430.0	7.91	77.54
WDJ091918.15-473354.38	23388.0	7.34	111.08	24277.0	7.71	106.02
WDJ094755.68-231234.10	22227.0	7.66	144.55	22868.0	7.48	144.21
WDJ101839.84-310802.03	15693.0	8.17	50.91	15666.0	8.29	42.72
WDJ104017.14-655324.81	22327.0	7.73	115.59	22389.0	7.98	125.74
WDJ105925.27-724409.93	19398.0	8.01	103.22	19441.0	8.36	82.0
WDJ112401.30-505938.44	16290.0	7.94	61.78	16356.0	8.18	67.3
WDJ121238.09-364240.22	19236.0	7.67	78.9	19321.0	7.57	77.25
WDJ123456.22+473733.37	15496.0	7.99	55.15	15069.0	7.92	49.2
WDJ124202.13+750845.51	19906.0	8.17	81.04	19407.0	7.97	74.04
WDJ133752.09+363733.44	21250.0	7.59	81.52	20983.0	7.85	90.29

Table E.1: Predicted $T_{\rm eff}$, $\log(g)$ and distance for the objects in this study, for the Montreal (M) and La Plata (LP) M-R relations.