universe

MDPI

Article

Non-Parametric Reconstruction of Cosmological Observables Using Gaussian Processes Regression

José de Jesús Velázquez [1], Luis A. Escamilla [2,3], Purba Mukherjee [4,*] and J. Alberto Vázquez [2,*]

[1] Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito de la Investigación Científica Ciudad Universitaria, Mexico City 04510, Mexico; jess_2497@ciencias.unam.mx
[2] Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca 62210, Mexico; luis.escamilla@icf.unam.mx
[3] School of Mathematics and Statistics, University of Sheffield, Hounsfield Road, Sheffield S3 7RH, UK
[4] Centre for Theoretical Physics, Jamia Millia Islamia, New Delhi 110025, India
* Correspondence: pdf.pmukherjee@jmi.ac.in (P.M.); javazquez@icf.unam.mx (J.A.V.)

Abstract: The current accelerated expansion of the Universe remains one of the most intriguing topics in modern cosmology, driving the search for innovative statistical techniques. Recent advancements in machine learning have significantly enhanced its application across various scientific fields, including physics, and particularly cosmology, where data analysis plays a crucial role in problem-solving. In this work, a non-parametric regression method with Gaussian processes is presented along with several applications to reconstruct some cosmological observables, such as the deceleration parameter and the dark energy equation of state, in order to contribute some information that helps to clarify the behavior of the Universe. It was found that the results are consistent with λCDM and the predicted value of the Hubble parameter at redshift zero is $H_0 = 68.798 \pm 6.340(1\sigma)$ km s^{-1} Mpc^{-1}.

Keywords: cosmology; dark energy; Hubble parameter; deceleration parameter; linear regression; gaussian process

Citation: Velázquez, J.d.J.; Escamilla, L.A.; Mukherjee, P.; Vázquez, J.A. Non-Parametric Reconstruction of Cosmological Observables Using Gaussian Processes Regression. *Universe* **2024**, *10*, 464. https://doi.org/10.3390/universe10120464

Academic Editor: Jean-Pierre Gazeau

Received: 3 October 2024
Revised: 7 December 2024
Accepted: 16 December 2024
Published: 20 December 2024

1. Introduction

Nowadays, it is becoming more common to hear that we are currently living in the Golden Age of Cosmology, whose origin goes back to the early 1990s when the Cosmic Background Explorer (COBE) satellite was launched in order to provide information about the Cosmic Microwave Background [1]. This event marked the beginning of a series of outstanding discoveries such as the necessity to incorporate the Dark Matter (DM) and Dark Energy (DE) components to account for the structure formation and the current accelerated expansion of the Universe, which later on gave rise to the standard cosmological model or Lambda Cold Dark Matter (LCDM) (more on this model later). This Golden Age is also characterized by the huge amount of observations and data obtained as a result of several worldwide collaborations, such as Planck [2], SDSS [3], SNLS [4], DESI [5], JWST [6] and Euclid [7], to mention a few. This was definitely a remarkable achievement since it provided the community with valuable information to work with, but it also came with a set of obstacles, such as how to process and analyze the avalanche of new data. Fortunately, around the same time, a new field of mathematics was starting to grow in strength: machine learning.

Machine Learning (ML) is the subfield of artificial intelligence dedicated to the mathematical modeling of data. It is a method for finding solutions to problems by using computers, which differs from regular programming since the latter uses data and rules to return results. In contrast, ML uses data and results to deduce the rules that relate them. An ML system is said to be trained rather than programmed [8]. ML can be broadly categorized into three types: supervised, unsupervised, and reinforcement learning [9]. It can handle a wide variety of problems, but the main goal is to learn the process of mapping inputs

into outputs, which can then be used to predict the outputs for new, unseen, inputs. These algorithms have been widely compared against traditional techniques in related fields, obtaining promising results in terms of efficiency and performance in favor of ML [10–14]. The main advantage of ML algorithms is that they can automate repetitive tasks such as data cleaning and pattern recognition that might require direct human intervention with traditional methods.

ML also contains valuable tools for the process known as reconstructions. In the absence of a fundamental and/or well-defined theory, a reconstruction is able to analyze a physical quantity and provide some insights into its general behavior; it can be broadly categorized into parametric and non-parametric [15]. In parametric reconstructions, the target quantity is studied by proposing a particular function with free parameters that should be inferred using observations. These functions are commonly phenomenological parameterizations to model, for instance, the DE equation of state $w(z)$, which could bring about some information about the DE's fundamental nature. Non-parametric reconstructions, on the other hand, apply numerical or statistical tools directly to the data instead of focusing on a particular function. This approach allows to decrease the bias towards a particular behavior; however, they are prone to overfitting or may produce inaccurate results on extrapolations outside the range spanned by the data. In this work, the main focus is on non-parametric reconstructions, where ML algorithms shine through. That is, ML algorithms allow for predicting the behavior of some observable quantities, even when an exact model of them is not fully specified [16]. Some useful and popular supervised learning methods that have been applied to cosmology are:

Artificial Neural Networks (ANNs): Named so because of their analogy to the behavior of the human brain. ANNs are made up of layers of sets of units called neurons that individually process data inputs. Each neuron is connected to the others through links with weights that are evaluated by an activation function, discarding the worst options and prioritizing the best ones. ANN are commonly used to solve classification and pattern recognition problems in images, speech, or signals. ANNs also have predictive applications in the financial [17] or atmospheric [18] sector. The field of cosmology has widely adopted neural networks across various applications. For example, CosmicNet I [19] and CosmicNet II [20], which are used to accelerate Einstein–Boltzmann solvers; physically-informed neural networks, which serve as a replacement for numerical solvers in cosmological codes for differential equations [21,22]; ANNs are applied directly to data to non-parametrically reconstruct certain cosmological quantities, including the Hubble parameter and structure formation through $f\sigma_8$ [23], the deceleration parameter [24], rotation curves [25], on scalar-tensor theories [26], and testing the cosmic distance ladder [27,28]. Neural networks also emulate functions such as the power spectrum [29–31] and speed up the computational processes [32–35]. Genetic algorithms [36] have also been employed in related tasks. For an introduction to ANN in cosmology, see [37].

Decision Trees and Random Forests (RF): Essentially, decision trees learn a hierarchy of if/else questions and reach an appropriate decision. Decision trees can be used in marketing campaigns [38] or diagnoses of diseases [39], to mention a few examples. Random forests are based on a set of decision trees that are uncorrelated and merged to create more accurate data predictions. These types of algorithms are often used to solve classification problems [40], which can be of great use in the field of cosmology. Some examples are gravitational waves' classification [41,42], joint redshift-stellar mass probability distribution functions [43] and N-body simulations [44–46].

k-Nearest Neighbors (k-NN): This algorithm consists of storing the training dataset and formulating a method that finds the closest data values to make predictions for a new test data point. It is possibly the simplest ML method and has a wide spectrum of applications, such as the creation of customized recommended systems [47]. Given the ease with which the k-NN finds groups/agglomerations, its use in cosmology has focused on topics related to structure formation such as galaxy-clustering [48–51].

There is another ML technique that works as the basis for this paper, and it is known as **Gaussian Process Regression** (GPR). Over the last decade, GPR has become particularly popular in cosmology for testing the concordance model [52–55], cosmographic studies [56–60], the reconstruction of parameters that characterize the cosmic expansion [56,61–64], reconstructing dark energy [65–70], constraining spatial curvature [71–75], exploring the interaction between dark matter and dark energy [76–80], testing modified theories of gravity [26,81–86], testing consistency among datasets [87], emulating the matter power spectrum [88], thermodynamic viability analysis [89,90], probing the cosmic reionization history [91–93] and the classification and identification of blended galaxies [94]; among many other research fields that take advantage of the ML capabilities for analyzing and classifying images, videos, and numerical data. For a pedagogical introduction to GPR, one can refer to the Gaussian process website (http://gaussianprocess.org/). Over the course of this work, a GPR will be defined and then tested by applying it to the prediction of observable quantities in cosmology. Therefore, the main objective of this work is to provide a basic introduction to Gaussian Processes (GPs) and a presentation of some applications of this method through examples.

The paper is structured as follows: in Section 2, a general description of the GPR is given; in Section 3, we explain the types of existing kernels in the context of a GPR; then, in Section 4, we give a brief review of cosmology that yields to the standard cosmological model; in Section 5, we make use of the GPR methodology on some cosmological quantities, and finally, in Section 6, we discuss our results and present our conclusions.

2. Gaussian Processes

In this section, we present some of the relevant concepts before delving into the GPR:

- Random Variable: a function that assigns a value to each event in the sample space of a random experiment, which could be either discrete or continuous. For example, when rolling two six-sided dice, the result will be two outcomes n_1 and n_2. In this case, a discrete random variable X can be the sum of the result of rolling both dice, i.e., $X(n_1, n_2) = n_1 + n_2$. In contrast, a continuous random variable could be the weight or height of a population. These quantities, once measured, comprise an interval on the number line, making it continuous when taking an infinite number of possible values.
- Correlation: also called "dependence", it is a statistical relationship between two random variables. For example, when comparing the height of a person with that of their parents, in general, it will be observed that the descendants have heights similar to the progenitors. This means that there is a connection or positive correlation between both heights. In general, the presence of consequences does not imply causality.
- Probability distribution: a function that assigns to each event, defined on the random variable, the probability that said event occurs. They can be discrete or continuous. A widely used one is the binomial probability distribution (where there are two possible mutually exclusive events):

$$P(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^k, \tag{1}$$

  where k is the number of times an event has occurred, p the probability that said event occurs, and n the number of total events.
- Normal distribution: also called Gaussian distribution, it is a type of continuous probability distribution with the form:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \tag{2}$$

  where x is a random variable, μ is the mean and σ the standard deviation.
- Random process: also called stochastic process, which is an object made up of several random variables. An example of a stochastic process is the random walker since each

step the walker takes is a random variable. The random variables are not necessarily independent of each other since there may be correlations as in the Markov chains where the next step in the chain depends on the immediate preceding one.

Let $x = X(\omega)$ be the value of a random variable X at ω and $f(x)$ its probability distribution. We say that X is normally distributed if f has the form of Equation (2), which is defined only by the mean μ and variance σ^2; hence, it can be denoted as

$$f(x) \sim N(\mu, \sigma^2). \tag{3}$$

Therefore, $N(\mu, \sigma^2)$ is said to be the normal (Gaussian) distribution. If we now have an arbitrary number of random variables $x_1, \ldots, x_n$, then the distribution becomes a multivariate normal distribution, which can be denoted as:

$$\bar{f} = [f(x_1), \ldots, f(x_n)] \sim \bar{N}(\bar{\mu}, K(x, x')), \tag{4}$$

where $\bar{\mu} = (\mu(x_1), \mu(x_2), \ldots, \mu(x_n))$ is the vector that contains the means of the random variables and

$$K(x, x') = \begin{pmatrix} K(x_1, x_1) & K(x_1, x_2) & \cdots & K(x_1, x_n) \\ K(x_2, x_1) & K(x_2, x_2) & \cdots & K(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_n, x_1) & K(x_n, x_2) & \cdots & K(x_n, x_n) \end{pmatrix}, \tag{5}$$

is a matrix with the covariances among the variables. Note that each diagonal element is the covariance of a random variable with itself, which equals its variance.

This reasoning can be extended to the case of a continuous random variable x where each value of x is a random variable. In this case, the mean vector becomes a function that returns the mean of the Gaussian distribution that defines x, and the covariance matrix has to be a function that gives the covariance between two continuous random variables x and x'. This generalization of a normal distribution for continuous random variables is known as a *Gaussian Process*. Therefore, a GP is an infinite collection of random variables that is defined by a mean function $\mu(x)$ and a covariance function $k(x, x')$, also known as the kernel of the process. Usually, the mean $\mu(x)$ is taken to be zero for simplicity, but it can be different with analogous calculations.

There are several types of kernels such as the rational quadratic, exponential or Matern (which will be further explained in later sections). For example, one of the most commonly used covariance functions due to its simplicity and infinite differentiability is the squared exponential kernel, which can be written as:

$$k(x_i, x_j) = e^{-\theta(x_i - x_j)^2}, \tag{6}$$

where the parameter θ indicates how the correlation is spread, as shown in Figure 1. The larger the value of θ, the stronger the correlation between variables.

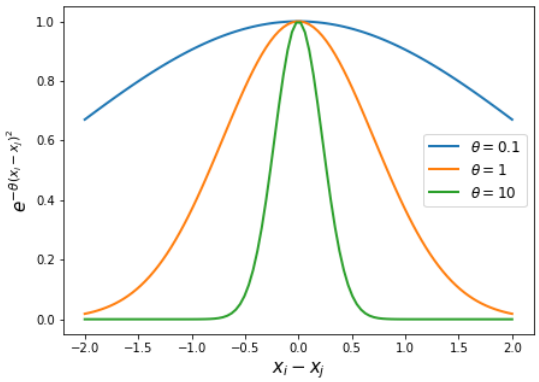

Figure 1. Changes in the correlation when varying the θ value.

2.1. Gaussian Process Regression

In order to train a Gaussian Process Regression (GPR) model, a dataset of n points $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ is needed. Let us define the vectors $\vec{x} = (x_1, x_2, \ldots, x_n)$ and $\vec{y} = (y_1, y_2, \ldots, y_n)$. The aim of a GPR is to find the posterior probability distribution for the values of the independent variable $P(\vec{w}|\vec{y}, \vec{x})$, where $\vec{w}$ is a vector of weights that defines the model. The posterior is computed by the Bayes' Rule:

$$P(\vec{w}|\vec{y}, \vec{x}) = \frac{P(\vec{y}|\vec{x}, \vec{w})P(\vec{w})}{P(\vec{y}|\vec{x})}. \tag{7}$$

Here, $P(\vec{w})$ is referred to as the *prior*, which is a probability distribution that contains information about $\vec{w}$ before the observed data; $P(\vec{y}|\vec{x}, \vec{w})$ is named the *likelihood* and it relates information about the prior distribution with the data; the marginal likelihood $P(\vec{y}|\vec{x})$ is a constant of normalization that guarantees the posterior is a probability ($0 \leq P(\vec{w}|\vec{y}, \vec{x}) \leq 1$), and it is given by the integral of the numerator over all possible values of $\vec{w}$:

$$P(\vec{y}|\vec{x}) = \int P(\vec{y}|\vec{x}, \vec{w})P(\vec{w})d\vec{w}. \tag{8}$$

Note that Bayes' Rule is not restricted to Gaussian distributions; however, in the context of GPR, the prior and posterior are both a GP and the data are Gaussian (each value is determined by a mean and a standard deviation). For this particular case, the prior and posterior are called conjugate distributions with respect to the likelihood function.

The GPR consists of making predictions based on the training data set (also called observables), assuming the observations are distributed around a model f with an additive noise ε, which is assumed to be Gaussian with zero mean and variance σ_n^2:

$$\begin{aligned} \vec{y} &= f(\vec{x}) + \varepsilon, \\ \mathrm{cov}(\vec{y}) &= K(\vec{x}, \vec{x}) + \sigma_n^2 I, \end{aligned} \tag{9}$$

where I is the identity matrix and $K(\vec{x}, \vec{x})$ is the covariance matrix obtained when evaluating the kernel in the corresponding training points, that is, $[K(\vec{x}, \vec{x})]_{ij} = k(x_i, x_j)$.

Therefore, it is required to find the test outputs $\vec{f}_*$, which are the values of the model at the test points $\vec{x}_* \equiv (x_{1*}, x_{2*}, \cdots, x_{n*})$. The posterior distribution of Equation (7) can be derived by conditioning the prior on the training observations, such that the conditional distribution of $\vec{f}_*$ only contains those functions from the prior that are consistent with the data set. Using the conditioning and marginalizing properties of the Gaussian distribution on the joint distribution for $\vec{f}_*$ and $\vec{y}$, it can be proven [95] that the mean and covariance of the predictions for the test set $\vec{x}_*$ are:

$$\begin{aligned} \bar{\vec{f}}_* &= \vec{K}_*^\top (\vec{K} + \sigma_n^2 \vec{I})^{-1} \vec{y}, \\ \mathrm{cov}(\vec{f}_*) &= \vec{K}_{**} - \vec{K}_*^\top \left(\vec{K} + \sigma_n^2 \vec{I}\right)^{-1} \vec{K}_*. \end{aligned} \tag{10}$$

The notation $\vec{K} = K(\vec{x}, \vec{x})$, $\vec{K}_* = K(\vec{x}, \vec{x}_*)$ and $\vec{K}_{**} = K(\vec{x}_*, \vec{x}_*)$ is introduced to simplify the calculations.

2.2. Maximum Likelihood Estimation

Assuming the cases in the training set are independent of each other, the probability density of the observations given a set of parameters $\vec{w}$, which is the likelihood from Equation (7), can be expressed as a product of individual densities

$$P(\vec{y}|\vec{x}, \vec{w}) = \prod_{i=1}^{n} p(y_i|x_i, \vec{w}), \tag{11}$$

where n is the number of input training points. Therefore, using the fact that the product of Gaussian distributions is also Gaussian, the marginal likelihood from Equation (8), in logarithmic form, becomes the *log marginal likelihood*

$$\log P(\vec{y}|\vec{x}) = -\frac{1}{2}\vec{y}^T(\vec{K}+\sigma_n^2\vec{I})^{-1}\vec{y} - \frac{1}{2}\log|\vec{K}+\sigma_n^2\vec{I}| - \frac{n}{2}\log 2\pi. \tag{12}$$

Optimal values of the parameters can be estimated by maximizing the log marginal likelihood. This training method used in GPR is known as the *maximum likelihood estimation* [95]. The maximizing can be performed by any optimizing algorithm, such as gradient descent or Markov Chain Monte Carlo.

3. GP Kernel

As seen so far, a fundamental feature of GPR that plays an important role in the fitting of a model is the kernel. A kernel (or covariance function) describes the covariance (correlation) of the random variables of the GP. Together with the mean function, the kernel completely defines a GP. In principle, any function that relates two points based on the distances between them can be a kernel, but it must satisfy certain conditions in order to represent a covariance function. For a function to be a valid kernel, the associated resulting matrix in Equation (5) must be positive definite, which implies that it has to be symmetric and invertible.

The covariance function of the variables x and x' is said to be *stationary* if it is a function only of $x - x'$ since it is invariant under translations, and *non-stationary* otherwise. Moreover, if it is a function only of $|x - x'|$ it is *isotropic* since it is invariant under rigid transformations.

As mentioned previously, it is necessary to choose a suitable kernel type for each particular problem. The process of creating a kernel from scratch is not always trivial, so it is usual to invoke a predefined kernel in order to model a diversity of processes. Some of the most used kernels are [40]:

- **Radial Basis Function.**

$$k(x,x') = \exp\left(-\frac{d(x,x')^2}{2l^2}\right), \tag{13}$$

  where $d(x,x')$ represents the euclidean distance between x and x', and $l > 0$ is known as the length parameter. Sometimes it is written in terms of a value θ that depends on the length parameter, such as in Equation (6). It is known as the Radial Basis Function (RBF) because it only depends on the radial distance. Notably, this kernel is *infinitely differentiable*, making it ideal for modeling smooth functions where high regularity is expected.
- **Matern.**

$$k(x,x') = \frac{1}{\Gamma(\nu)2^{\nu-1}}\left(\frac{\sqrt{2\nu}}{l}d(x,x')\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}}{l}d(x,x')\right), \tag{14}$$

  where K_ν is the modified Bessel function of the second kind, $\Gamma(\nu)$ is the Gamma function, l is the characteristic length scale, and ν is the smoothness parameter that controls the degree of differentiability of the function. For $\nu = \frac{1}{2}$, the Matern kernel reduces to the *exponential kernel*, which models processes with rough, non-smooth behavior. Notably, for $\nu = 1.5$ and $\nu = 2.5$, the kernel corresponds to once and twice differentiable functions, respectively, allowing for more regular behavior. As $\nu \to \infty$, the kernel approaches the RBF kernel, which is infinitely differentiable. This flexibility and control over the smoothness makes the Matern kernel especially useful for modeling functions that exhibit varying degrees of smoothness, as often seen in real-world data.

- **Exponential Sine Squared (ESS).**

$$k(x, x') = \exp\left(-\frac{2\sin^2(\pi d(x, x')/p)}{l^2}\right), \tag{15}$$

where $p > 0$ is the periodicity parameter, controlling the periodicity of the kernel, and $l > 0$ is the length scale parameter. This kernel is often called a *periodic kernel* because it models periodic functions with a sinusoidal component, where the periodicity is governed by p. The exponential decay modulates the amplitude of the sine function, enabling the kernel to capture periodic behaviors with varying smoothness.

- **Dot Product.**

$$k(x, x') = \sigma_0^2 + x \cdot x', \tag{16}$$

where σ_0 is a parameter that controls the inhomogeneity or offset of the kernel, while $x \cdot x'$ represents the dot product between the vectors x and x'. The term σ_0^2 allows for a shift in the kernel's value, providing flexibility in modeling non-zero means or biases in the data.

- **Rational Quadratic (RQ).**

$$k(x, x') = \left(1 + \frac{d(x, x')^2}{2\alpha l^2}\right)^{-\alpha}, \tag{17}$$

where α is known as the scale mixture parameter, and l is the length scale. The RQ kernel is a scale mixture of *squared exponential kernels*, allowing it to model functions with varying smoothness over different scales. When $\alpha = 1$, the kernel behaves similarly to the RBF kernel. The kernel can model processes with longer-range dependencies for $0 < \alpha < 1$, while for $\alpha > 1$, it captures shorter-range correlations.

Each of the values that can be varied within the kernel, such as l, σ_0, etc. are called *hyperparameters*. It is said that GPR is a non-parametric technique because the number of hyperparameters is infinite. The reader might have noticed that all kernels described above are stationary (dependent on $|x - x'|$), except Dot Product. This dependence on distance alone makes stationary kernels more rigid, while also presenting poor predictive power when outside the scope of the used data when compared with their non-stationary counterparts. Non-stationary kernels are more flexible, which allows for a better estimate outside the scope covered by the observations. Nevertheless, they are rarely used given the high number of hyperparameters to optimize, the high complexity, the high computational costs and a greater risk of overfitting when compared to stationary ones [96–99]. In this work, we will exclusively use stationary kernels and Dot Product; although, we think that the idea of using non-stationary ones for cosmological observations might be worth visiting in future work.

Since the kernel is a key feature of GPR, modifying it might produce different models. Therefore, it is necessary to establish which kernel is the best option for a particular model. In a real problem, such as those presented in cosmology, the kind of relationship between two variables is not always previously known. In these cases, the kernel that results in the best fit after regression may be chosen from a set of default kernels.

3.1. Kernel Selection Through χ^2.

A robust statistical tool, known as the χ^2 test, could be employed to determine which model, derived from various kernels, fits best a specific dataset, thereby enhancing the regression analysis. This test evaluates the congruence between two datasets by assessing whether a significant discrepancy exists between the observed data values and the model's predictions.

The method consists of defining the objective function χ^2 as:

$$\chi^2 = \sum_{i,j} \big(y_i - f(x_i)\big) C_{ij}^{-1} \big(y_j - f(x_j)\big), \tag{18}$$

where (x_i, y_i) are the data points (or training set), C_{ij} is the covariance matrix and $f(x_i)$ are the values of the model at the independent variable of the data points. When the covariance matrix is diagonal, we obtain a simplified case for the χ^2 as:

$$\chi^2 = \sum_i \frac{[y_i - f(x_i)]^2}{\sigma_{y_i}^2}, \tag{19}$$

where $\sigma_{y_i}^2$ is the variance and the ith element in the diagonal of C_{ij}. The GPR produces a model data set that can be interpreted as a function f of the independent variable x. Given an observable (x_i, y_i), the numerator of Equation (19) represents the squared distance between the observable and the model for the same value of x_i. By computing this difference over all the available observations (and as such calculating the χ^2 function), we can generate an idea of how well model f fits the data.

If the value of χ^2 is obtained for models built with different kernels, the best fit will be the one that minimizes this objective function. Notice that this method is different from the maximum likelihood estimation explained in Section 2.2 since it is not used to determine the hyperparameters as in the training. In this case, the models of regression have been determined previously for different kernels and tested to find the best model in terms of the covariance function.

3.2. A Generic Example

In this section, regression models based on Gaussian Processes are constructed from a mock dataset exhibiting a straight-line behavior. Fortunately, nowadays there is a broad range of standard developed code and libraries that facilitate performing a Gaussian Process Regression, such as `GPy` [100], `GPflow` [101], `GPyTorch` [102], `PyMC` [103], `scikit-learn` [40] and `GaPP` [66]. The latter two are the ones used during the course of this example and the complete step-by-step procedure can be found in the public repository [104]. To further simplify, the construction of a GPR model consists of three steps: (1) specify the prior distribution via the kernel, (2) find the hyperparameters that maximize Equation (12) and (3) evaluate predictions with Equation (10) using the optimal hyperparameters and observables.

We will use the function `GaussianProcessRegressor()`, which initializes a GP prior to regression with a specified kernel and its parameters. The method `fit()` returns the same `GaussianProcessRegressor()` object fitted to the observables using the maximum likelihood estimation. This method takes two lists as parameters that correspond to the observational data variables $\vec{x}$ and $\vec{y}$. Finally, the `predict()` method returns the means and standard deviations of the predictions using Equation (10).

In the first of our examples of regression, the variances $\sigma_{y_i}^2$ or noises of the observational data are ignored. This approach assumes that the data measurements are exact, thus implying there are no uncertainties or error bars associated with them.

A mock data set scattered around a linear equation $Y = mX + b$ with $m = 3$ and $b = -4$ is created by adding a random value between -15 and 15 to 10 evaluations of the equation at different values of $X \in [0, 10]$. The aim of the GPR is to reproduce the graph of the line that originated the set.

In this case, the χ^2 test cannot be used to find an optimal kernel, since Equation (18) is undefined; thus, an alternative method to determine the kernel must be used. The predictions of the model using a specific kernel at different values of X will be compared via the sum of squared Euclidean distances to the points of the original linear relationship

at the corresponding X-values scaled by the number of data points, n. The result is called the Mean Squared Error (MSE) and can be written as:

$$\text{MSE} = \frac{1}{n}\sum_i [y_i - f(x_i)]^2. \tag{20}$$

The regression model that minimizes the MSE is the one that most resembles the desired line. At this point, without loss of generality, the kernel used in this GPR is the Matern (Equation (14)).

Figure 2 shows the observational mock dataset (black points), the model predictions (blue solid line), the line from which the data was obtained (red dash) and the confidence zones (in lilac colors) that correspond to 2σ and 3σ, respectively. These confidence intervals will be used for all the regression models throughout this work.

We tested different kernels, and the results achieved are quite similar; however, as can be seen in Figure 3, the one that minimizes the MSE is the Dot Product kernel. The Gaussian process regression for each kernel is shown in Figure A1. The Dot Product kernel produces a linear regression model, so it is usually the best choice when fitting a straight line. In contrast, for the rest of the kernels, the uncertainty reduces to zero when the model is evaluated at the observations. This can be interpreted as the model overfitting the data, which is expected given that the mock data presents no variances [105]. To mitigate overfitting, one approach is to introduce an additional hyperparameter, σ_n, for noise modeling. This hyperparameter accounts for the observational noise, preventing the model from fitting the data too precisely. However, adding σ_n increases the model complexity, requiring careful tuning to balance the bias and variance [95].

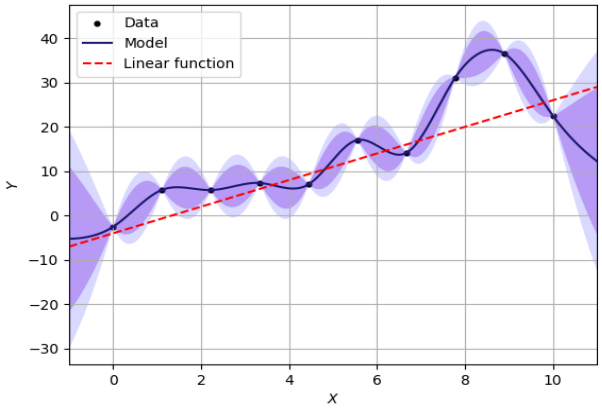

Figure 2. Linear model with null variances in the data. A Matern kernel was used for the reconstruction. It is evident that there is an overestimation of the confidence zone because our data lack errors.

On the other hand, when the observables do have uncertainties (which is the case that most closely resembles real data), the variances must be added to the diagonal of the kernel matrix, as shown in Equation (9). If these uncertainties come in the form of a non-diagonal covariance matrix, then it is also added to the kernel so that:

$$\text{cov}(\vec{y}) = \vec{K} + \vec{C}, \tag{21}$$

with $\vec{C}$ being the covariance matrix of the data. The `GaussianProcessRegressor()` function is able to obtain an array `alpha` as an input whose size is equal to the amount of data that corresponds to the variances associated with each observation. The outcome of this approach, illustrated in Figure 4, demonstrates that the prediction more closely resembles a straight line, especially when compared to the scenario with a mock dataset with null variances.

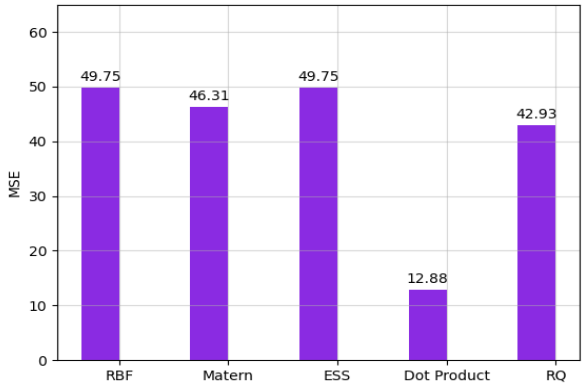

Figure 3. Comparison of the mean squared errors, calculated via Equation (20), for different kernels. As can be seen, the model that minimizes the MSE corresponds to the Dot Product kernel, which produces a linear regression.

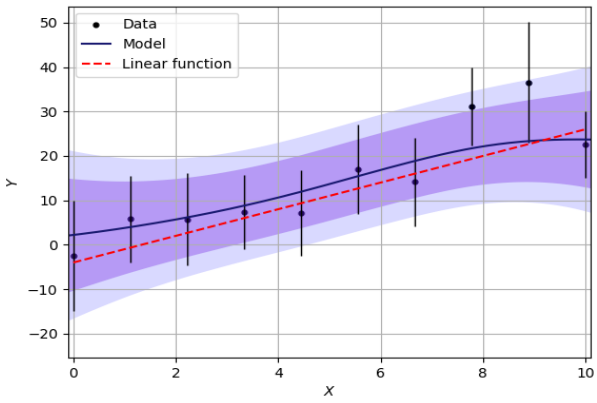

Figure 4. Linear model with variances in the data. An almost linear behavior is observed as we no longer have overfitting.

In this scenario, similar to the MSE test performed in Figure 3, a χ^2 test can be employed to determine the kernel that generates the optimal model. This involves creating a model for each kernel test, computing the χ^2 value for each model, and selecting the one with the lowest χ^2. In Figure 5, we plot the results of this test and, by analyzing it, it can be concluded that the model that yields the best fit to our data is the one utilizing a Matern kernel, as it produces the lowest value of the objective function. It is crucial to note that the model with the lowest χ^2 is not necessarily the best one, as excessively minimizing χ^2 can lead to overfitting. The linear regression models for each kernel are shown in Figure A2.

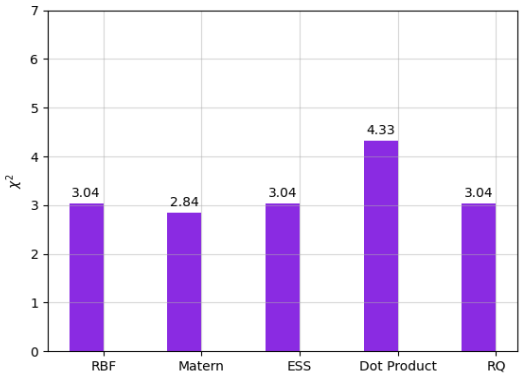

Figure 5. Different models with a specific kernel are produced to compute their corresponding χ^2 value. The kernel with the lowest χ^2 generates the optimal model—in this case, the Matern kernel.

3.3. Derivatives of a GP

The RBF kernel (Equation (13)) is infinitely differentiable, and the derivative of a GP is also a GP, which allows for reconstructing the derivatives of a function from data. In order to reconstruct the derivative, not only the covariance between the observational data is required but also the covariance between the function and its derivative and among the derivatives of the reconstruction. All of them can be calculated from the derivative of the kernel function, as described in [66].

As in Section 2.1, it can be proven that the mean and covariance of the prediction for the first derivative of this function at test points $\vec{x}_* \equiv (x_{1*}, x_{2*}, \cdots, x_{n*})$ using a differentiable kernel $k(x_i, x_j)$ are:

$$\begin{aligned} \vec{\bar{f}}'_* &= \vec{K}'^{\top}_* (\vec{K} + \sigma_n^2 \vec{I})^{-1} \vec{y}\,, \\ \operatorname{cov}(\vec{f}'_*) &= \vec{K}''_{**} - \vec{K}'^{\top}_* \left(\vec{K} + \sigma_n^2 \vec{I}\right)^{-1} \vec{K}'_*\,. \end{aligned} \tag{22}$$

Here, $\vec{K}'_* = K'(\vec{x}, \vec{x}_*) = \frac{\partial k(x_i, x_{j*})}{\partial x_{j*}}$ and $\vec{K}''_{**} = K''(\vec{x}_*, \vec{x}_*) = \frac{\partial^2 k(x_{i*}, x_{j*})}{\partial x_{i*} \partial x_{j*}}$ are introduced to simplify the notation.

As can be inferred from these equations, the derivative of the kernel must exist in order to compute the predictions of a derivative using GPR. Therefore, an infinitely differentiable covariance function is useful when reconstructing a derivative from data, which is why an RBF kernel (Equation (13)) is preferred among others in this type of problem. If an RBF kernel is used, the procedure can be generalized to any derivative of the model and, in particular, the package `GaPP` [66] allows to compute up to the third derivative of a function quickly.

In order to verify the reliability of the code, a mock data set of values scattered around a sinusoidal function $y(x) = \sin(x)$ was created by adding a random value between -0.15 and 0.15 for different values of x. Then the standard deviation (the error bar) of each data point was emulated by a random number between 0.1 and 0.3. The reconstructions of the function and its derivatives are shown in Figure 6. The red lines represent the analytical function (the sine function or its derivatives as appropriate) and the blue lines are the regression models. The confidence zones correspond to the intervals delimited by 2σ (95%) and 3σ (99%), where σ are the standard deviations of the predictions. Note that the analytical function is in the 2σ interval for all the cases, which indicates that the regression is considerably accurate.

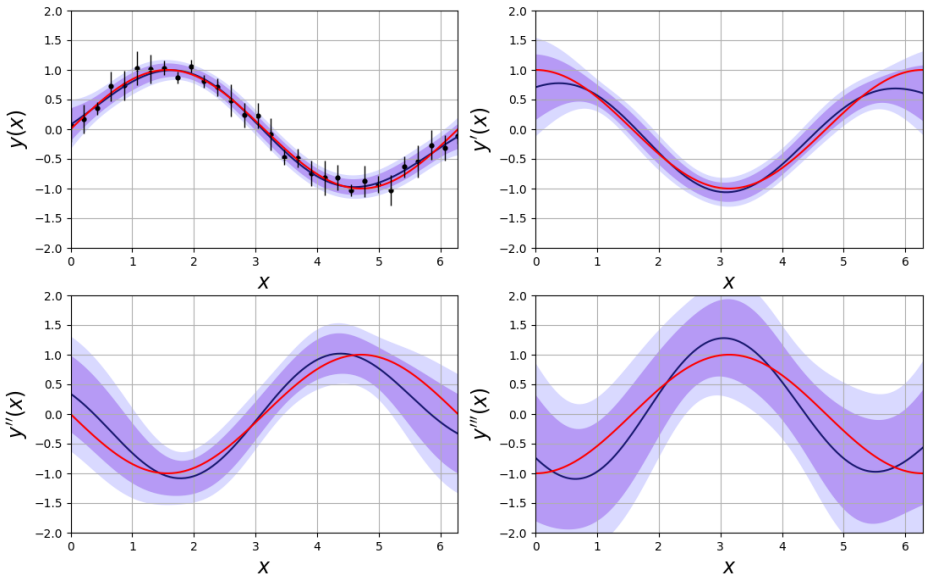

Figure 6. Reconstruction of an example test function $f(x) = \sin x$ and its derivatives on $[0, 2\pi]$ from a mock data set. The red lines represent the analytical function or derivative, and the blue lines are the predictions.

Figure 6 only shows the scatter plot of the mock data set because the observables for the derivatives do not exist. This is an advantage of the technique since it is possible to find the nth-derivative of a function only from data values of such function.

4. Cosmology

Let us start by considering the Universe as being homogeneous on scales larger than 150 Mpc, which means that the distribution of its components does not depend on the position of the observer, despite the fact that at short distances, the density of matter is perceived as random. Likewise, let us also assume the Universe is isotropic, which implies that its properties are the same regardless of the direction from which they are observed. The assumption of these two characteristics (homogeneity and isotropy at large scales) is known as the Cosmological Principle, and it has been adopted to set restrictions on a great variety of alternative cosmological theories.

It is firmly established by observations that our Universe expands [106]. The standard Big Bang model proposes that the Universe emerged about 15 billion years ago and it has been expanding and cooling since then. Measurements using Type IA supernovae as standard candles have proven that the expansion of the Universe is also accelerating [107,108] and such acceleration is only possible if a substantial fraction of the total energy density is a kind of energy with a negative pressure [109]. This energy component is referred to as Dark Energy (DE) given its unknown nature and origin. Furthermore, along with DE, another key component, known as Dark Matter (DM), is necessary to explain observations regarding structure formation. Given the enigmatic nature of both DE and DM, predicting the Universe's long-term future remains an elusive task. Consequently, DE and DM represent two of the most compelling and complex challenges in contemporary cosmology.

The expansion of the Universe is described by the Friedmann equations, obtained as solutions of the Einstein field equations for the Friedman–Lemaitre–Robertson–Walker (FLRW) metric and a perfect fluid with density ρ and pressure p. The equations in standard form are:

$$\begin{aligned} H^2 = \left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right), \end{aligned} \tag{23}$$

where a is known as the scale factor, which is a dimensionless function of time and is related to the size of the Universe; $\dot{a}$ and $\ddot{a}$ denote the first and second derivative of a with respect to the cosmic time; H is the Hubble parameter, which describes how fast the Universe is expanding; G is the gravitational constant; $c \approx 3 \times 10^5$ km/s is the speed of light in vacuum; and k is the curvature parameter, which determines the shape of the Universe [110].

One of the most favored models by evidence is the ΛCDM. This model proposes that the DM component of the Universe is non-relativistic (cold) and only interacts gravitationally, while the DE is due to an unknown component represented by the cosmological constant Λ. As mentioned previously, DE is an exotic component in the energy budget of the Universe, which is theorized to be responsible for its accelerated expansion. Most cosmological models consider DE to be a perfect fluid, which means that it is incompressible and has a viscosity of zero. Then, it follows that for a perfect fluid, its equation of state (EoS) is characterized by a dimensionless value w. In the case of barotropic fluids w given by the proportionality function between its pressure p and energy density ρ:

$$p = c^2 w \rho.$$

For perfect fluids, such as baryonic matter and relativistic matter (radiation), their EoS's are $w = 0$ and $w = 1/3$, respectively. Understanding the behavior of the Dark Energy's equation of state is a focal point of contemporary cosmology. It is established that the

pressure exerted by DE must be negative, given its role in driving cosmic expansion instead of contraction. Furthermore, accelerated expansion is predicted to occur when the equation of state parameter falls below $-1/3$. When working with the ΛCDM model, one assumes that $w = -1$ for the DE, giving it the characteristic behavior of a cosmological constant.

For the standard cosmological model, taking into consideration the equations of state for every component when solving Equation (23), the Hubble parameter obtained from the first Friedmann equation in terms of the present values of the density parameters Ω_i is:

$$H(z) = H_0\sqrt{\Omega_{r,0}a^{-4} + \Omega_{m,0}a^{-3} + \Omega_{k,0}a^{-2} + \Omega_{\Lambda,0}}, \tag{24}$$

where the density parameters are $\Omega_{r,0}$ for radiation, $\Omega_{m,0}$ for the matter sector, which includes DM and baryons, $\Omega_{k,0}$ to account for the spatial curvature, $\Omega_{\Lambda,0}$ to describe the vacuum density in the form of a cosmological constant (this represents the DE component) and H_0 the Hubble parameter, known as the Hubble constant. The subscript "0" means that they are evaluated at the present time. For a spatially flat model ($\Omega_k = k = 0$), we have $\Omega_m + \Omega_r + \Omega_\Lambda = 1$.

In order to determine a concept of distance between two objects in the Universe, it is convenient to present some common definitions of distance measures in Cosmology [111,112], these include:

1. **Comoving distance**: Due to the homogeneity of the Universe, it is possible to define a coordinate system that considers the expansion of the Universe. The distance between two objects in this system remains constant, so the comoving distance is defined as

$$d_C(z) = d_H \int_0^z dz' \frac{H_0}{H(z')}, \tag{25}$$

where $d_H = \frac{c}{H_0}$ is the Hubble distance.

2. **Transverse comoving distance**: When considering the curvature intrinsic to the geometry of space-time, expressed by the parameter Ω_k, the transversal comoving distance is defined as,

$$d_M = \begin{cases} \frac{d_H}{\sqrt{\Omega_k}} \sinh\left(\frac{\sqrt{\Omega_k} d_C(z)}{d_H}\right) & \text{if } \Omega_k > 0, \\ d_C(z) & \text{if } \Omega_k = 0, \\ \frac{d_H}{\sqrt{-\Omega_k}} \sin\left(\frac{\sqrt{-\Omega_k} d_C(z)}{d_H}\right) & \text{if } \Omega_k < 0, \end{cases} \tag{26}$$

which is equal to the comoving distance in the case of a flat space-time, i.e., for $\Omega_k = 0$.

3. **Luminosity distance**: Comparing the absolute and apparent magnitudes between two objects, that is, the actual brightness emitted by an object compared to the brightness observed from Earth, the luminosity distance is defined, which is written in terms of the transverse comoving distance as:

$$d_L(z) = (1+z) d_M(z). \tag{27}$$

From the above equations, the normalized comoving distance is also obtained as,

$$D(z) = \frac{1}{d_H}\left(\frac{1}{1+z}\right) d_L(z). \tag{28}$$

In the particular case of a flat Universe, a simple expression for the derivative of the normalized comoving distance can be obtained:

$$D'(z) = \frac{H_0}{H(z)}. \tag{29}$$

The cosmological quantities are broadly categorized into two groups—the physical (dynamical) quantities like the DE equation of state parameter w, vs the kinematical (cosmograph-

ical) quantities that are defined as time derivatives of the scale factor a, for example, the Hubble H, deceleration q and jerk j parameters. The deceleration parameter is defined as:

$$q = -\frac{\ddot{a}a}{\dot{a}^2}, \tag{30}$$

which can be written in terms of the derivatives of $D(z)$ with respect to the redshift z, as

$$q(z) = -1 - \frac{D''(z)}{D'(z)}(1+z), \tag{31}$$

or in terms of $H(z)$ and its derivative $H'(z)$, as

$$q(z) = -1 - \frac{H'(z)}{H(z)}(1+z)\,. \tag{32}$$

The deceleration parameter is a measure of the acceleration of the expansion of space, and it is said to be accelerating when q becomes negative [111].

Furthermore, with DE having a time-varying dynamical equation of state $w(z)$ (ignoring the contribution from spatial curvature and radiation), we can write the Hubble parameter $H(z)$ by integrating the Friedmann Equation (23) as,

$$\frac{H^2(z)}{H_0^2} = \Omega_{m,0}(1+z)^3 + (1-\Omega_{m,0})e^{3\int_0^z \frac{1+w(x)}{1+x}dx}. \tag{33}$$

On differentiating the above equation, one can arrive at this expression for the DE equation of state $w(z)$ as

$$w(z) = \frac{2(1+z)H(z)H'(z) - 3H^2(z)}{3H^2(z) - \Omega_{m,0}H_0^2(1+z)^3}\,. \tag{34}$$

As the deceleration parameter q is now estimated and found to be evolving, we focus on the next higher-order derivative, the jerk parameter j, defined as

$$j = \frac{\dddot{a}}{aH^3}\,. \tag{35}$$

It can be rewritten as a function of redshift z, in terms of the Hubble parameter H along with its derivatives $H'(z)$ and $H''(z)$, as

$$j(z) = 1 - 2(1+z)\frac{H'}{H} + (1+z)^2\left(\frac{H''}{H} + \frac{H'^2}{H^2}\right). \tag{36}$$

For the ΛCDM model, j is exactly unity. Therefore, any non-monotonic evolution of j can help in understanding the nature of dark energy in the absence of any convincing physical theory [113,114].

5. Cosmological Functions with GPR

5.1. Hubble Parameter

In cosmology, the aim is to find a mathematical description that explains the characteristics of the Universe and predicts its evolution over time. Thus, determining the dependency of H as a function of z is one of the main topics of study in cosmology.

For the above, the regression method with GP is a very useful tool as it allows to reconstruct the evolutionary model from certain observational data. In this case, the data will be Hubble parameter observations for different redshifts from cosmic chronometers as an alternative to the commonly used data from Type Ia Supernovae. There is a set of 31 data points for $H(z)$ obtained by different authors, which have been gathered and used in many works, such as [115,116]. Using the developed code that contains variances and the Hubble parameter data, the model shown in Figure 7 is obtained.

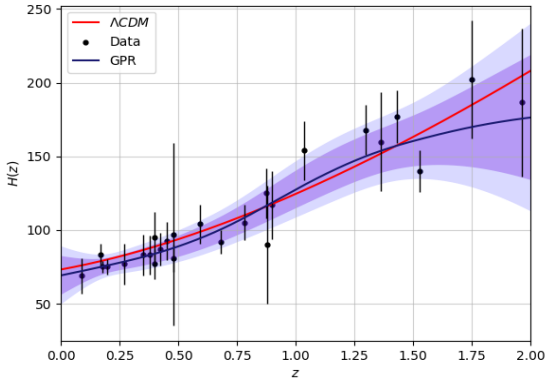

Figure 7. Hubble parameter reconstruction model using a Matern kernel.

The curve for $H(z)$ in the ΛCDM model was created from Equation (24), and the values for the density parameters given by Planck results [2], were obtained under the assumption of a flat Universe as ΛCDM.

Various models were tested with different kernels, as shown in Figure 8. The model that minimizes χ^2 for $H(z)$ was produced by a Matern kernel with the default initial characteristic length of $l = 1$ and an order of $\nu = 1.5$. The optimized hyperparameter after the training is $l = 4.1$. One key result to stress is that from the family of kernels, the least preferred corresponds to the Dot Product, that is, the equation of state for the dark energy is inclined to be anything else except a linear regression, in particular, a constant equation of state.

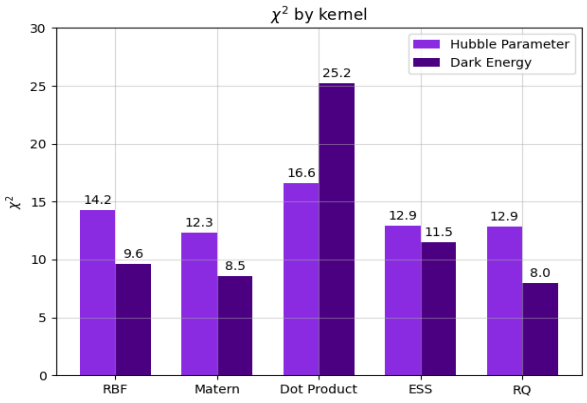

Figure 8. Models with different kernels were produced for both the Hubble Parameter and the Dark Energy functions, and its χ^2 values were computer. As observed, the optimal covariance function was the Matern for the Hubble Parameter and the Rational Quadratic for the Dark Energy.

Evaluating the model for $z = 0$, the value for the Hubble constant $H(0) = H_0 = 68.79 \pm 6.34(1\sigma)$ km Mpc^{-1} s^{-1} is obtained as can be seen in [104]. While the current analysis focuses solely on the cosmic chronometer Hubble data, it is worth noting that incorporating other datasets, such as the Type Ia Supernovae (SNIa) [117,118] or Baryon Acoustic Oscillations (BAO) [119–121], could provide a more comprehensive approach to determining H_0. However, unlike the cosmic chronometers, both these datasets do not offer independent measurements of H_0. For SNIa, astrophysical modeling of the absolute magnitude M_B is required, and for BAO, the sound horizon r_d must be known. This calls for additional modeling to properly handle the GP kernel hyperparameters and cosmological parameters in a coherent manner [122], which is beyond the scope of the present work. However, we plan to expand the analysis by incorporating these additional datasets, which could improve the precision of H_0 estimates [64,77,115,123,124] and offer a more robust comparison in future studies.

5.2. Dark Energy Equation of State

If the DE is considered a dynamic component, then its EoS should be different from -1 (so as to differentiate itself from ΛCDM), or it could present a dependence on redshift as $w(z)$. As proof of the concept and using the previously established methods, we will use a mock dataset, which contains information about the BAO and SNIa datasets (the data points used here come from a model-independent reconstruction of the DE EoS from [125]) from the dark energy equation of state as a function of z to reconstruct it. As such, a non-parametric model of $w(z)$ with GPR using an RQ kernel will be obtained (Figure 9). In Figure 8, a comparison of the values of χ^2 for models obtained using different kernels is shown and the different models can be seen in Figure A4. Note that when reconstructing $H(z)$, the RQ kernel was the best option since it returned the minimum values. As already stated earlier, the ΛCDM model EoS for DE is proposed as a constant of value -1, so if there are cracks in the standard model then our reconstruction should find deviations from this value. In our case, it was found, from the mock dataset that $w = -1$ is well within 1σ bounds of the reconstruction using GPR, which can be interpreted as minor evidence against ΛCDM. For the sake of comparison, we perform the reconstruction of $w(z)$ through Equation (34), but now by using the cosmic chronometers dataset. In general terms, the results obtained in Figure 10 resemble the ones presented in Figure 9, with a maximum value around $z\sim1$ and the crossing of the phantom divide line ($w = -1$) at about $z\sim1.5$. Error bars correspond to $\Omega_{m,0} = 0.3$, and to examine the sensitivity of the reconstruction, we have included two more values of $\Omega_{m,0}$.

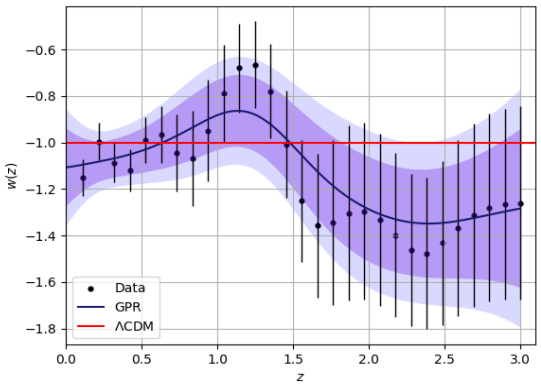

Figure 9. Reconstruction model for the dark energy equation of state, using a mock dataset.

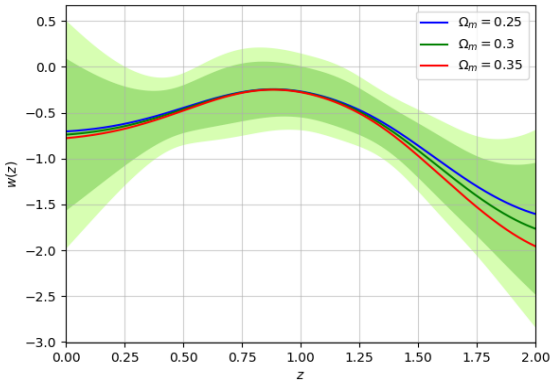

Figure 10. Reconstruction model for the dark energy equation of state, from the reconstructionof $H(z)$ using cosmic chronometers through Equation (34).

5.3. Reconstruction of the Deceleration Parameter

By using $H(z)$ data from cosmic chronometers, the predicted value of $H_0 = 68.798$ km Mpc^{-1} s^{-1} and Equation (29), we can obtain a derived dataset of $D'(z)$.

To obtain the variances/errors of this new dataset, it is straightforward to use the approximation of ratio distribution for uncorrelated variables (the variance of a ratio distribution $\frac{X}{Y}$ of two uncorrelated random variables X and Y can be approximated with a Taylor expansion around μ_X and μ_Y as [126]: $\mathrm{Var}\left(\frac{X}{Y}\right) = \frac{\mu_X^2}{\mu_Y^2}\left[\frac{\mathrm{Var}(X)^2}{\mu_X^2} + \frac{\mathrm{Var}(Y)^2}{\mu_Y^2}\right]$). Thus far, the Matern kernel has presented the most suitable models (at least regarding the χ^2 obtained); henceforth, this kernel will be used for the reconstruction. The resulting GPR prediction for $D'(z)$ from the derived dataset and a comparison with the ΛCDM values computed by combining Equations (24) and (29) with the corresponding density parameters and the value of $H_0 = 67.32$ km Mpc^{-1} s^{-1} from Planck results [2] are shown in Figure 11.

From the same dataset of $D'(z)$, the derivative $D''(z)$ is reconstructed using the `GaPP` package, as explained in Section 3. On the other hand, Equation (29) can be differentiated analytically and evaluated for the ΛCDM density parameters to obtain $D''(z)$. Figure 12 shows the predictions for $D''(z)$ and a comparison with ΛCDM. Finally, from Equation (31) and the GPR predictions of $D'(z)$ and $D''(z)$, a model of the deceleration parameter is produced as in the previous cases. The regression is compared with ΛCDM in Figure 13. We see again some agreement between our reconstruction and the standard model, although an important thing to note is that there is a region where ΛCDM remains outside the 1σ contour, and it is really close to being outside the 2σ one. This could indicate some actual evidence in favor of our reconstruction or at least highlight a tension existing within ΛCDM. Similar discrepancies have been noted in previous studies, where deviations from ΛCDM behavior were observed in various cosmological datasets [58,123,127,128]. These findings suggest potential new physics beyond the standard cosmological model or the need for refined cosmological parameters, which calls for further investigation.

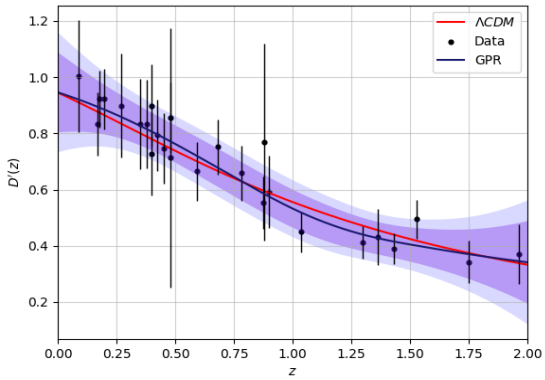

Figure 11. Reconstruction of the first derivative of the normalized comoving distance.

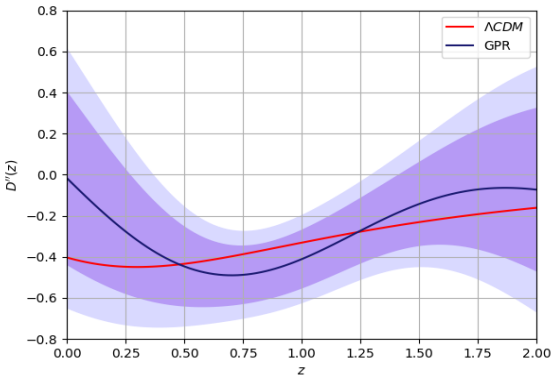

Figure 12. Reconstruction of the second derivative of the normalized comoving distance.

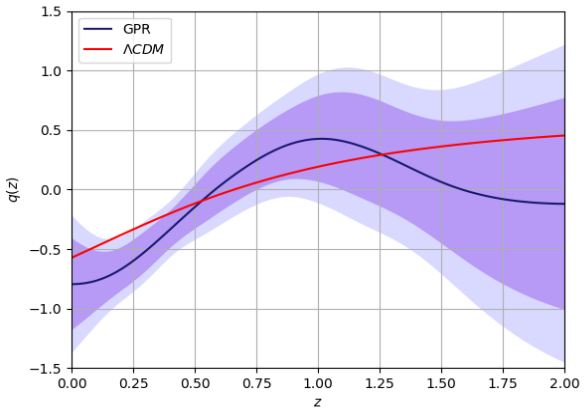

Figure 13. Reconstruction of the deceleration parameter.

5.4. Deceleration Parameter Reconstruction with a Mock Data Set from ΛCDM

If the observables are indeed produced by the ΛCDM model (that is to say that the standard model is the "real" one), a regression using artificial data that were produced by the ΛCDM model should be quite similar to the model obtained from the "real" data. To verify this, we produced a mock dataset of $H(z)$ around the values obtained by evaluating Equation (24) for the density parameters given by ΛCDM cosmology from Planck's results [2]. Then, the whole procedure to obtain $q(z)$ was repeated but this time using the mock dataset so that a comparison with the previous reconstruction could be made. The result and comparison are shown in Figure 14. As expected, the mock data set regression model is into the 2σ confidence level of the reconstruction from the observations. This indicates that, even if the standard model does not reproduce the observables exactly or does it with some caveats, it can emulate the general observed behavior pretty well.

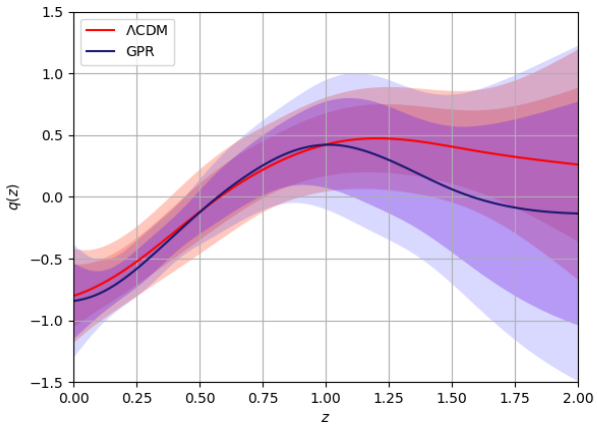

Figure 14. Comparison between the reconstruction of $q(z)$ and ΛCDM model using a mock data set.

5.5. Reconstruction of the Jerk Parameter

From the same $D'(z)$ dataset, derived in Section 5.3, one can further reconstruct the third derivative $D'''(z)$ along with the second derivative $D''(z)$ employing the `GaPP` package, as explained in Section 3. With these reconstructed functions, the evolution of the cosmological jerk parameter can be obtained from Equation (36), as a function of the redshift z. This regression is compared with the ΛCDM case in Figure 15. We find that our reconstruction includes the ΛCDM model (i.e., $j = 1$) within the 1σ confidence level. The mean of the reconstructed function clearly indicates that j has a non-monotonic evolution, which is in agreement with the previous findings [77].

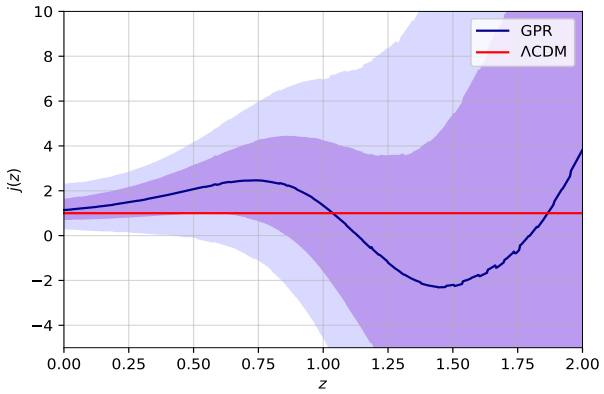

Figure 15. Comparison between the reconstruction of $j(z)$ and ΛCDM model.

5.6. Using GPR as an Interpolation in a Model-Independent Way

Throughout this work, we have demonstrated how a GPR can be utilized in a non-parametric manner to study cosmological quantities. However, there is an alternative approach to leverage the properties of a GPR that we would like to mention. This method also employs a Gaussian process but in a model-independent manner, as it involves inferring parameter values using datasets and Bayesian statistics. By doing so, we can directly compare our model-independent reconstruction against the standard model using Bayesian evidence and maximum log-likelihood.

The GPR in this approach is performed "nodes". These nodes can vary in height (their ordinate position), and these "variable heights" work as the new parameters of the reconstruction [129,130]. For n nodes, we have n variable heights and, as such, n new parameters that need to be inferred. This method has been used before with the equation of state of Dark Energy [131], the interaction kernel of an IDE (interacting Dark Energy) model [80], and the cosmic reionization history [92,93].

6. Discussion and Conclusions

Although Gaussian Process Regression (GPR) does not yield an explicit form of the relationship between variables, it remains a robust method for making predictions given a particular set of observables. It reconstructs functions effectively without needing prior assumptions about their behavior, leveraging libraries like `GaPP` to predict higher derivatives, such as $D''(z)$ and $H'(z)$, which is particularly valuable in cosmological analyses.

GPR has been extensively applied in cosmology, spanning from reconstructing the dark energy equation of state $w(z)$ to cosmographical studies. This flexibility allows GPR to adapt to diverse datasets, making it a powerful tool for probing dark energy and other cosmological phenomena. In gravitational wave cosmology, GPR has been instrumental in reconstructing the luminosity distance from simulated data [132], enabling non-parametric inference of the Hubble parameter and forecasting deviations from the standard ΛCDM model. Additionally, GPR has been employed in large-scale structure studies, such as reconstructing the growth rate of cosmic structures $f\sigma_8(z)$ from redshift space distortions [63,70]. These applications highlight GPR's versatility in handling diverse cosmological datasets and addressing critical questions within the cosmological context.

However, GPR is constrained by the range of observed data, limiting its predictive accuracy outside this interval. Furthermore, uncertainties in derivative function reconstructions increase beyond the data range, which can impact the reliability of extrapolations. The choice of the kernel function in GPR is pivotal, influencing prediction means and covariances significantly. Despite methods like cross-validation and Bayesian model selection to aid kernel selection, the optimal choice remains non-trivial, affecting the quality of reconstructions.

Furthermore, random number generation plays a subtle but crucial role in GPR applications. While tools like `scikit-learn` utilize robust generators like Permuted Congruential Generator 64-bit (PCG-64) [133], the quality of these generators in high-dimensional spaces,

as highlighted in prior studies [134–136], warrants scrutiny. Though we found no immediate issues with the generators used in this work, exploring alternatives like `RANLUX` (http://luscher.web.cern.ch/luscher/ranlux (accessed on 19 November 2024)) [135], known for its high-quality randomness, could further ensure reliability. Such considerations are important in cosmological contexts, where multi-parameter reconstructions are common, and small biases can propagate into significant systematic errors [137,138].

Comparing GPR with other parametric and non-parametric methods, principal component analysis [139] (PCA), logarithmic parametrization [140], rational parametrization [141], Bayesian methods [142], reveals trade-offs between flexibility and interpretability. While PCA simplifies data dimensionality effectively, it may overlook intricate data complexities that GPR can capture. Bayesian methods provide comprehensive probabilistic frameworks but often require detailed prior information and intensive computational resources.

In summary, Gaussian processes offer a powerful and flexible tool for cosmological analyses, enabling model-independent reconstructions and effective uncertainty handling. Despite computational challenges and kernel sensitivity, their widespread application in cosmology demonstrates their potential to provide nuanced insights into the evolutionary history of our Universe.

Author Contributions: All authors contributed equally. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data and code available in the public repository https://github.com/JesusUg2497/GP_in_Cosmology (accessed on 10 May 2024).

Acknowledgments: PM acknowledges funding from Anusandhan National Research Foundation (ANRF), Govt of India under the National PostDoctoral Fellowship (File no. PDF/2023/001986). JJV, LAE and JAV acknowledge the support provided by FORDECYT-PRONACES-CONACYT/304001/2020 and UNAM-DGAPA-PAPIIT IN117723.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. GPR Figures

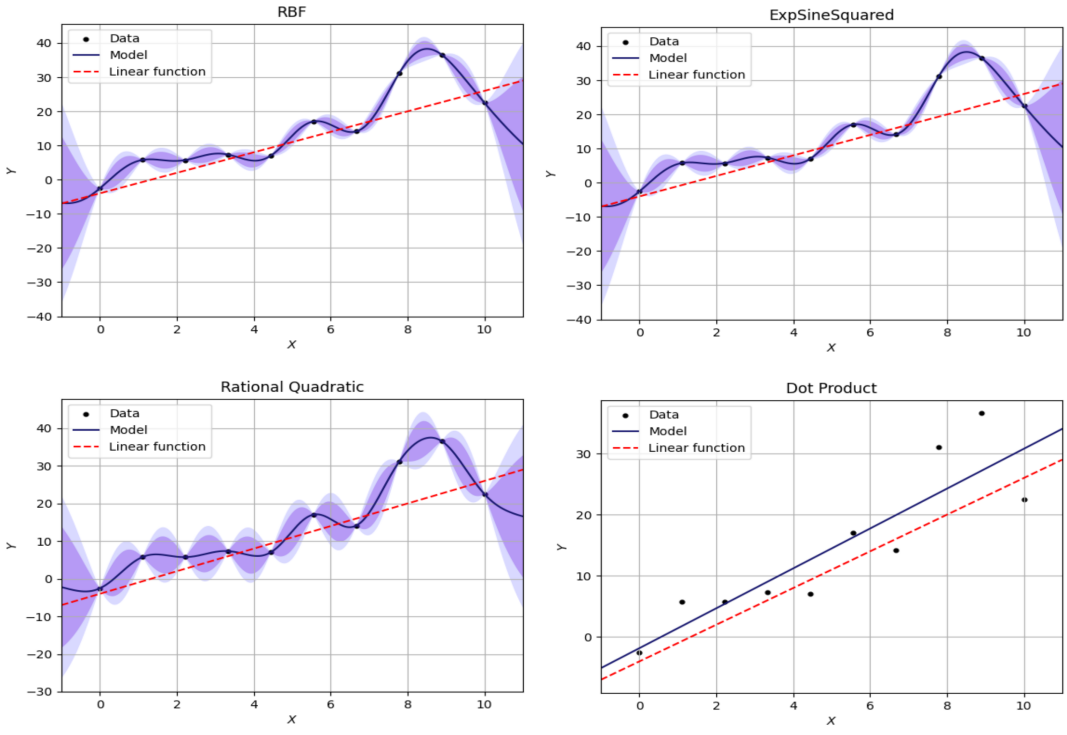

Figure A1. Models obtained from different kernels for the same linear regression with null variances.

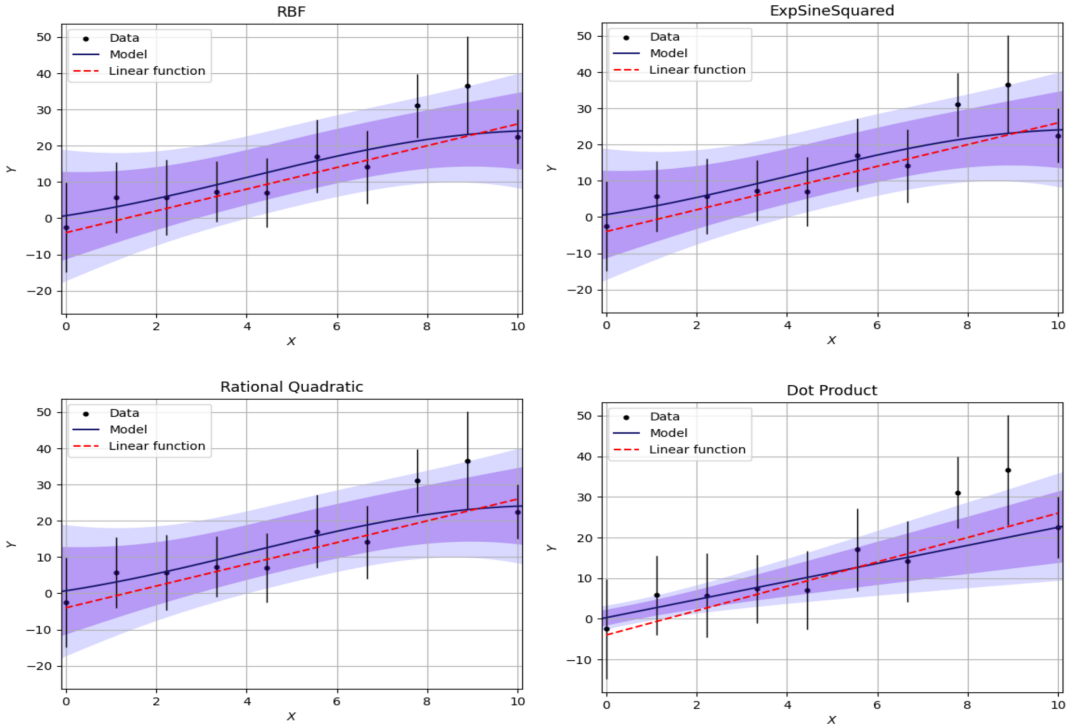

Figure A2. Linear regression models from different kernels with non-zero variances.

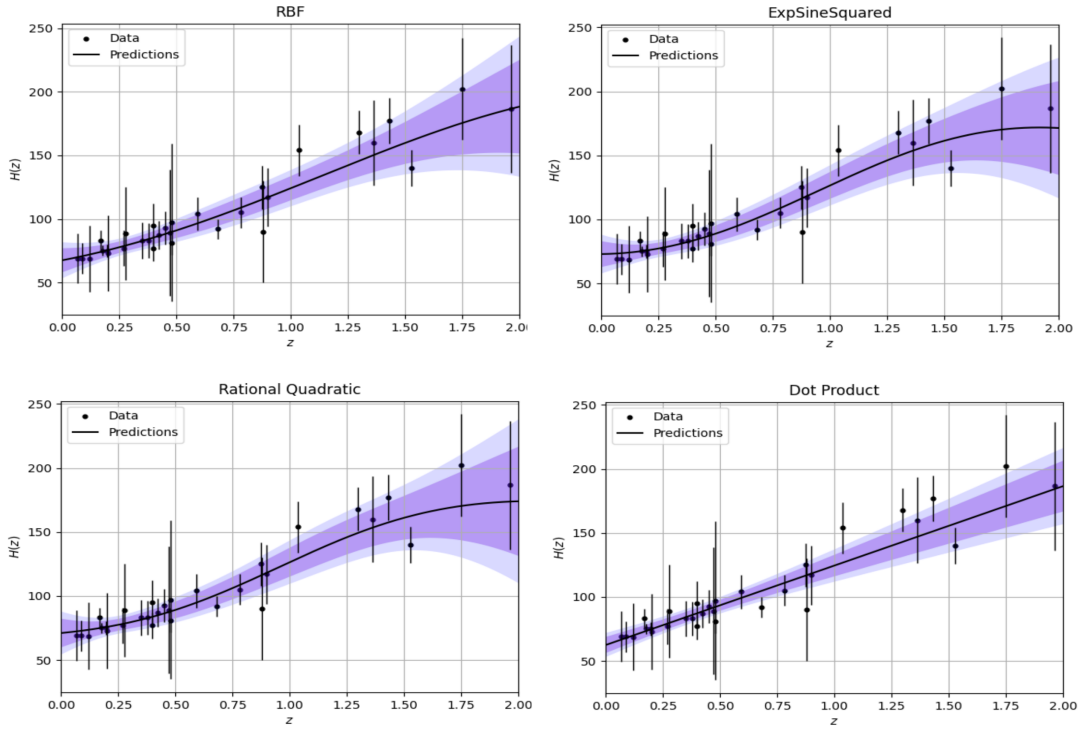

Figure A3. Regression models of Hubble Parameter for different kernels.

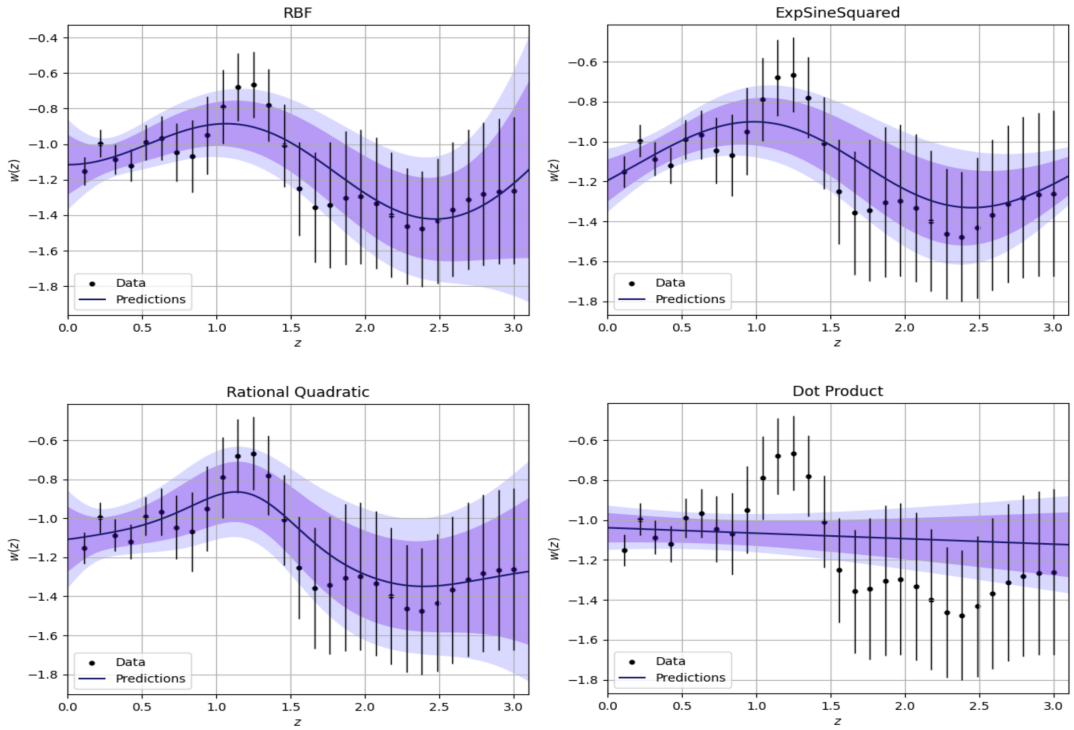

Figure A4. Regression models of Dark Energy for different kernels.

References

1. Boggess, N.W.; Mather, J.C.; Weiss, R.; Bennett, C.L.; Cheng, E.A.; Dwek, E.; Gulkis, S.; Hauser, M.G.; Janssen, M.A.; Kelsall, T.; et al. The COBE mission-Its design and performance two years after launch. *Astrophys. J.* **1992**, *397*, 420–429. [CrossRef]
2. Aghanim, N.; Akrami, Y.; Ashdown, M.; Aumont, J.; Baccigalupi, C.; Ballardini, M.; Banday, A.J.; Barreiro, R.; Bartolo, N.; Basak, S.; et al. Planck 2018 results-VI. Cosmological parameters. *Astron. Astrophys.* **2020**, *641*, A6.
3. York, D.G.; Adelman, J.; Anderson, J.E., Jr; Anderson, S.F.; Annis, J.; Bahcall, N.A.; Bakken, J.; Barkhouser, R.; Bastian, S.; Berman, E.; et al. The sloan digital sky survey: Technical summary. *Astron. J.* **2000**, *120*, 1579. [CrossRef]
4. Astier, P.; Guy, J.; Regnault, N.; Pain, R.; Aubourg, E.; Balam, D.; Basa, S.; Carlberg, R.; Fabbro, S.; Fouchez, D.; et al. The Supernova Legacy Survey: Measurement of, and w from the first year data set. *Astron. Astrophys.* **2006**, *447*, 31–48. [CrossRef]
5. Dey, A.; Schlegel, D.J.; Lang, D.; Blum, R.; Burleigh, K.; Fan, X.; Findlay, J.R.; Finkbeiner, D.; Herrera, D.; Juneau, S.; et al. Overview of the DESI Legacy Imaging Surveys. *Astron. J.* **2019**, *157*, 168. [CrossRef]
6. Gardner, J.P.; Mather, J.C.; Abbott, R.; Abell, J.S.; Abernathy, M.; Abney, F.E.; Abraham, J.G.; Abraham, R.; Abul-Huda, Y.M.; Acton, S.; et al. The James Webb Space Telescope Mission. *Publ. Astron. Soc. Pac.* **2023**, *135*, 068001. [CrossRef]
7. Mellier, Y.; Barroso, J.A.; Achúcarro, A.; Adamek, J.; Adam, R.; Addison, G.E.; Aghanim, N.; Aguena, M.; Ajani, V.; Akrami, Y.; et al. Overview of the Euclid mission. *arXiv* **2024**, arXiv:2405.13491.
8. Chollet, F. *Deep Learning with Python*; Manning: Shelter Island, NY, USA , 2017.
9. Theobald, O. *Machine Learning for Absolute Beginners: A Plain English Introduction*; Ai, Data Science, Python & Statistics for Beginners; Scatterplot Press: London, UK, 2017.
10. Sarker, I.H. Machine learning: Algorithms, real-world applications and research directions. *SN Comput. Sci.* **2021**, *2*, 160. [CrossRef] [PubMed]
11. Ray, S. A quick review of machine learning algorithms. In Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 14–16 February 2019 ; IEEE: Piscataway, NJ, USA, 2019; pp. 35–39.
12. Qiu, J.; Wu, Q.; Ding, G.; Xu, Y.; Feng, S. A survey of machine learning for big data processing. *EURASIP J. Adv. Signal Process.* **2016**, *2016*, 67. [CrossRef]
13. Mahesh, B. Machine learning algorithms—A review. *Int. J. Sci. Res. (IJSR)* **2020**, *9*, 381–386. [CrossRef]
14. Dhall, D.; Kaur, R.; Juneja, M. Machine learning: A review of the algorithms and its applications. In *Proceedings of ICRIC 2019: Recent Innovations in Computing*; Springer: Cham, Switzerland, 2020; pp. 47–63.
15. Sahni, V.; Starobinsky, A. Reconstructing Dark Energy. *Int. J. Mod. Phys. D* **2006**, *15*, 2105–2132. . [CrossRef]
16. Müller, A.; Guido, S. *Introduction to Machine Learning with Python: A Guide for Data Scientists*; O'Reilly Media: Sebastopol, CA, USA, 2016.

17. Sewell, M. Financial applications of Neural Networks. 2023. Available online: http://machine-learning.martinsewell.com/ann/finance.html (accessed on 10 July 2024).
18. Abhishek, K.; Singh, M.; Ghosh, S.; Anand, A. Weather Forecasting Model using Artificial Neural Network. *Procedia Technol.* **2012**, *4*, 311– 318. [CrossRef]
19. Albers, J.; Fidler, C.; Lesgourgues, J.; Schöneberg, N.; Torrado, J. CosmicNet. Part I. Physics-driven implementation of neural networks within Einstein-Boltzmann Solvers. *J. Cosmol. Astropart. Phys.* **2019**, *2019*, 28. [CrossRef]
20. Günther, S.; Lesgourgues, J.; Samaras, G.; Schöneberg, N.; Stadtmann, F.; Fidler, C.; Torrado, J. CosmicNet II: Emulating extended cosmologies with efficient and accurate neural networks. *J. Cosmol. Astropart. Phys.* **2022**, *2022*, 35. [CrossRef]
21. Chantada, A.T.; Landau, S.J.; Protopapas, P.; Scóccola, C.G.; Garraffo, C. Cosmology-informed neural networks to solve the background dynamics of the Universe. *Phys. Rev. D* **2023**, *107*, 063523. [CrossRef]
22. Chantada, A.T.; Landau, S.J.; Protopapas, P.; Scóccola, C.G.; Garraffo, C. Faster Bayesian inference with neural network bundles and new results for f (R) models. *Phys. Rev. D* **2024**, *109*, 123514. [CrossRef]
23. Gómez-Vargas, I.; Esquivel, R.M.; García-Salcedo, R.; Vázquez, J.A. Neural network reconstructions for the Hubble parameter, growth rate and distance modulus. *Eur. Phys. J. C* **2023**, *83*, 304. [CrossRef]
24. Mukherjee, P.; Levi Said, J.; Mifsud, J. Neural network reconstruction of H'(z) and its application in teleparallel gravity. *J. Cosmol. Astropart. Phys.* **2022**, *2022*, 29. [CrossRef]
25. Garcia-Arroyo, G.; Gómez-Vargas, I.; Vázquez, J.A. Reconstructing rotation curves with artificial neural networks. *arXiv* **2024**, arXiv:2404.05833.
26. Dialektopoulos, K.F.; Mukherjee, P.; Levi Said, J.; Mifsud, J. Neural network reconstruction of scalar-tensor cosmology. *Phys. Dark Univ.* **2024**, *43*, 101383. [CrossRef]
27. Mukherjee, P.; Dialektopoulos, K.F.; Said, J.L.; Mifsud, J. A possible late-time transition of M_B inferred via neural networks. *arXiv* **2024**, arXiv:2402.10502.
28. Shah, R.; Saha, S.; Mukherjee, P.; Garain, U.; Pal, S. LADDER: Revisiting the Cosmic Distance Ladder with Deep Learning Approaches and Exploring Its Applications. *Astrophys. J. Suppl.* **2024**, *273*, 27. [CrossRef]
29. Agarwal, S.; Abdalla, F.B.; Feldman, H.A.; Lahav, O.; Thomas, S.A. PkANN–I. Non-linear matter power spectrum interpolation through artificial neural networks. *Mon. Not. R. Astron. Soc.* **2012**, *424*, 1409–1418. [CrossRef]
30. Agarwal, S.; Abdalla, F.B.; Feldman, H.A.; Lahav, O.; Thomas, S.A. Pkann–II. A non-linear matter power spectrum interpolator developed using artificial neural networks. *Mon. Not. R. Astron. Soc.* **2014**, *439*, 2102–2121. [CrossRef]
31. Costanza, B.; Scóccola, C.G.; Zaldarriaga, M. Enhancing CMB map reconstruction and power spectrum estimation with convolutional neural networks. *J. Cosmol. Astropart. Phys.* **2024**, *2024*, 41. [CrossRef]
32. Gómez-Vargas, I.; Vázquez, J.A. Deep learning and genetic algorithms for cosmological Bayesian inference speed-up. *Phys. Rev. D* **2024**, *110*, 083518. [CrossRef]
33. Nygaard, A.; Holm, E.B.; Hannestad, S.; Tram, T. CONNECT: A neural network based framework for emulating cosmological observables and cosmological parameter inference. *J. Cosmol. Astropart. Phys.* **2023**, *2023*, 25. [CrossRef]
34. Sikder, S.; Barkana, R.; Reis, I.; Fialkov, A. Emulation of the cosmic dawn 21-cm power spectrum and classification of excess radio models using an artificial neural network. *Mon. Not. R. Astron. Soc.* **2023**, *527*, 9977–9998. [CrossRef]
35. Jense, H.T.; Harrison, I.; Calabrese, E.; Mancini, A.S.; Bolliet, B.; Dunkley, J.; Hill, J.C. A complete framework for cosmological emulation and inference with CosmoPower. *arXiv* **2024**, arXiv:2405.07903.
36. Gómez-Vargas, I.; Andrade, J.B.; Vázquez, J.A. Neural networks optimized by genetic algorithms in cosmology. *Phys. Rev. D* **2023**, *107*, 043509. [CrossRef]
37. Olvera, J.d.D.R.; Gómez-Vargas, I.; Vázquez, J.A. Observational Cosmology with Artificial Neural Networks. *Universe* **2022**, *8*, 120. [CrossRef]
38. Liu, Y.; Yang, S. Application of Decision Tree-Based Classification Algorithm on Content Marketing. *J. Math.* **2022**, *2022*, 6469054. [CrossRef]
39. Podgorelec, V.; Kokol, P.; Stiglic, B.; Rozman, I. Decision trees: An overview and their use in medicine. *J. Med. Syst.* **2002**, *26*, 445–463. [CrossRef] [PubMed]
40. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
41. Shah, N.; Knee, A.M.; McIver, J.; Stenning, D.C. Waves in a forest: A random forest classifier to distinguish between gravitational waves and detector glitches. *Class. Quantum Grav.* **2023**, *40*, 235008. [CrossRef]
42. Baker, P.T.; Caudill, S.; Hodge, K.A.; Talukder, D.; Capano, C.; Cornish, N.J. Multivariate classification with random forests for gravitational wave searches of black hole binary coalescence. *Phys. Rev. D* **2015**, *91*, 062004. [CrossRef]
43. Mucesh, S.; Hartley, W.G.; Palmese, A.; Lahav, O.; Whiteway, L.; Bluck, A.; Alarcon, A.; Amon, A.; Bechtol, K.; Bernstein, G.; et al. A machine learning approach to galaxy properties: Joint redshift–stellar mass probability distributions with Random Forest. *Mon. Not. R. Astron. Soc.* **2021**, *502*, 2770–2786. [CrossRef]
44. Conceição, M.; Krone-Martins, A.; da Silva, A.; Moliné, Á. Fast emulation of cosmological density fields based on dimensionality reduction and supervised machine learning. *Astron. Astrophys.* **2024**, *681*, A123. [CrossRef]
45. Chacón, J.; Vázquez, J.A.; Almaraz, E. Classification algorithms applied to structure formation simulations. *Astron. Comput.* **2022**, *38*, 100527. [CrossRef]

46. Chacón, J.; Gómez-Vargas, I.; Méndez, R.M.; Vázquez, J.A. Analysis of dark matter halo structure formation in N-body simulations with machine learning. *Phys. Rev. D* **2023**, *107*, 123515. [CrossRef]
47. Ji, S. kNN_Recommender_System. 2021. Available online: https://github.com/jisilvia/kNN_Recommender_System (accessed on 15 July 2024).
48. Banerjee, A.; Abel, T. Nearest neighbour distributions: New statistical measures for cosmological clustering. *Mon. Not. R. Astron. Soc.* **2021**, *500*, 5479–5499. [CrossRef]
49. Banerjee, A.; Kokron, N.; Abel, T. Modelling nearest neighbour distributions of biased tracers using hybrid effective field theory. *Mon. Not. R. Astron. Soc.* **2022**, *511*, 2765–2781. [CrossRef]
50. Yuan, S.; Zamora, A.; Abel, T. 2D k-th nearest neighbour statistics: A highly informative probe of galaxy clustering. *Mon. Not. R. Astron. Soc.* **2023**, *522*, 3935–3947. [CrossRef]
51. Wang, Y.; Banerjee, A.; Abel, T. Detection of spatial clustering in the 1000 richest SDSS DR8 redMaPPer clusters with nearest neighbor distributions. *Mon. Not. R. Astron. Soc.* **2022**, *514*, 3828–3843. [CrossRef]
52. Nair, R.; Jhingan, S.; Jain, D. Testing the consistency between cosmological measurements of distance and age. *Phys. Lett. B* **2015**, *745*, 64–68. [CrossRef]
53. Rana, A.; Jain, D.; Mahajan, S.; Mukherjee, A.; Holanda, R.F.L. Probing the cosmic distance duality relation using time delay lenses. *J. Cosmol. Astropart. Phys.* **2017**, *2017*, 10. [CrossRef]
54. Mukherjee, P.; Mukherjee, A. Assessment of the cosmic distance duality relation using Gaussian process. *Mon. Not. R. Astron. Soc.* **2021**, *504*, 3938–3946. [CrossRef]
55. Mukherjee, P.; Rodrigues, G.; Bengaly, C. Examining the validity of the minimal varying speed of light model through cosmological observations: Relaxing the null curvature constraint. *Phys. Dark Univ.* **2024**, *43*, 101380. [CrossRef]
56. Shafieloo, A.; Kim, A.G.; Linder, E.V. Gaussian Process Cosmography. *Phys. Rev. D* **2012**, *85*, 123530. [CrossRef]
57. Mukherjee, P.; Banerjee, N. Non-parametric reconstruction of the cosmological *jerk* parameter. *Eur. Phys. J. C* **2021**, *81*, 36. [CrossRef]
58. Mukherjee, P.; Banerjee, N. Revisiting a non-parametric reconstruction of the deceleration parameter from combined background and the growth rate data. *Phys. Dark Univ.* **2022**, *36*, 100998. [CrossRef]
59. Jesus, J.F.; Benndorf, D.; Escobal, A.A.; Pereira, S.H. From Hubble to snap parameters: A Gaussian process reconstruction. *Mon. Not. R. Astron. Soc.* **2024**, *528*, 1573–1581. [CrossRef]
60. Dinda, B.R. Analytical Gaussian process cosmography: Unveiling insights into matter-energy density parameter at present. *Eur. Phys. J. C* **2024**, *84*, 402. [CrossRef]
61. Mukherjee, P.; Sen, A.A. Model-independent cosmological inference post DESI DR1 BAO measurements. *arXiv* **2024**, arXiv:2405.19178. [CrossRef]
62. Mukherjee, P.; Shah, R.; Bhaumik, A.; Pal, S. Reconstructing the Hubble Parameter with Future Gravitational-wave Missions Using Machine Learning. *Astrophys. J.* **2024**, *960*, 61. [CrossRef]
63. Calderón, R.; L'Huillier, B.; Polarski, D.; Shafieloo, A.; Starobinsky, A.A. Joint reconstructions of growth and expansion histories from stage-IV surveys with minimal assumptions. II. Modified gravity and massive neutrinos. *Phys. Rev. D* **2023**, *108*, 023504. [CrossRef]
64. L'Huillier, B.; Shafieloo, A.; Polarski, D.; Starobinsky, A.A. Defying the laws of Gravity I: Model-independent reconstruction of the Universe expansion from growth data. *Mon. Not. R. Astron. Soc.* **2020**, *494*, 819–826. [CrossRef]
65. Holsclaw, T.; Alam, U.; Sanso, B.; Lee, H.; Heitmann, K.; Habib, S.; Higdon, D. Nonparametric Reconstruction of the Dark Energy Equation of State from Diverse Data Sets. *Phys. Rev. D* **2011**, *84*, 083501. [CrossRef]
66. Seikel, M.; Clarkson, C.; Smith, M. Reconstruction of dark energy and expansion dynamics using Gaussian processes. *J. Cosmol. Astropart. Phys.* **2012**, *2012*, 36. [CrossRef]
67. Zhang, M.J.; Li, H. Gaussian processes reconstruction of dark energy from observational data. *Eur. Phys. J. C* **2018**, *78*, 460. [CrossRef]
68. Dinda, B.R.; Maartens, R. Model-agnostic assessment of dark energy after DESI DR1 BAO. *arXiv* **2024**, arXiv:2407.17252.
69. Dinda, B.R.; Banerjee, N. A comprehensive data-driven odyssey to explore the equation of state of dark energy. *Eur. Phys. J. C* **2024**, *84*, 688. [CrossRef]
70. Calderón, R.; L'Huillier, B.; Polarski, D.; Shafieloo, A.; Starobinsky, A.A. Joint reconstructions of growth and expansion histories from stage-IV surveys with minimal assumptions: Dark energy beyond Λ. *Phys. Rev. D* **2022**, *106*, 083513. [CrossRef]
71. Yang, Y.; Gong, Y. Measurement on the cosmic curvature using the Gaussian process method. *Mon. Not. R. Astron. Soc.* **2021**, *504*, 3092–3097. [CrossRef]
72. Dhawan, S.; Alsing, J.; Vagnozzi, S. Non-parametric spatial curvature inference using late-Universe cosmological probes. *Mon. Not. R. Astron. Soc.* **2021**, *506*, L1–L5. [CrossRef]
73. Mukherjee, P.; Banerjee, N. Constraining the curvature density parameter in cosmology. *Phys. Rev. D* **2022**, *105*, 063516. [CrossRef]
74. Dinda, B.R. Minimal model-dependent constraints on cosmological nuisance parameters and cosmic curvature from combinations of cosmological data. *Int. J. Mod. Phys. D* **2023**, *32*, 2350079. [CrossRef]
75. Dinda, B.R.; Singirikonda, H.; Majumdar, S. Constraints on cosmic curvature from cosmic chronometer and quasar observations. *arXiv* **2023**, arXiv:2303.15401.

76. Yang, T.; Guo, Z.K.; Cai, R.G. Reconstructing the interaction between dark energy and dark matter using Gaussian Processes. *Phys. Rev. D* **2015**, *91*, 123533. [CrossRef]
77. Mukherjee, P.; Banerjee, N. Nonparametric reconstruction of interaction in the cosmic dark sector. *Phys. Rev. D* **2021**, *103*, 123530. [CrossRef]
78. Cai, R.G.; Tamanini, N.; Yang, T. Reconstructing the dark sector interaction with LISA. *J. Cosmol. Astropart. Phys.* **2017**, *2017*, 31. [CrossRef]
79. Bonilla, A.; Kumar, S.; Nunes, R.C.; Pan, S. Reconstruction of the dark sectors' interaction: A model-independent inference and forecast from GW standard sirens. *Mon. Not. R. Astron. Soc.* **2022**, *512*, 4231–4238. [CrossRef]
80. Escamilla, L.A.; Akarsu, O.; Di Valentino, E.; Vazquez, J.A. Model-independent reconstruction of the interacting dark energy kernel: Binned and Gaussian process. *J. Cosmol. Astropart. Phys.* **2023**, *2023*, 51. [CrossRef]
81. Zhou, L.; Fu, X.; Peng, Z.; Chen, J. Probing the Cosmic Opacity from Future Gravitational Wave Standard Sirens. *Phys. Rev. D* **2019**, *100*, 123539. [CrossRef]
82. Belgacem, E.; Foffa, S.; Maggiore, M.; Yang, T. Gaussian processes reconstruction of modified gravitational wave propagation. *Phys. Rev. D* **2020**, *101*, 063505. [CrossRef]
83. Yang, T. Gravitational-Wave Detector Networks: Standard Sirens on Cosmology and Modified Gravity Theory. *J. Cosmol. Astropart. Phys.* **2021**, *2021*, 44. [CrossRef]
84. Levi Said, J.; Mifsud, J.; Sultana, J.; Adami, K.Z. Reconstructing teleparallel gravity with cosmic structure growth and expansion rate data. *J. Cosmol. Astropart. Phys.* **2021**, *2021*, 15. [CrossRef]
85. Bernardo, R.C.; Levi Said, J. A data-driven reconstruction of Horndeski gravity via the Gaussian processes. *J. Cosmol. Astropart. Phys.* **2021**, *2021*, 14. [CrossRef]
86. Gadbail, G.N.; Mandal, S.; Sahoo, P.K. Gaussian Process Approach for Model-independent Reconstruction of f(Q) Gravity with Direct Hubble Measurements. *Astrophys. J.* **2024**, *972*, 174. [CrossRef]
87. Keeley, R.E.; Shafieloo, A.; Zhao, G.B.; Vazquez, J.A.; Koo, H. Reconstructing the Universe: Testing the Mutual Consistency of the Pantheon and SDSS/eBOSS BAO Data Sets with Gaussian Processes. *Astron. J.* **2021**, *161*, 151. [CrossRef]
88. Ho, M.F.; Bird, S.; Shelton, C.R. Multifidelity emulation for the matter power spectrum using Gaussian processes. *Mon. Not. R. Astron. Soc.* **2021**, *509*, 2551–2565. [CrossRef]
89. Banerjee, N.; Mukherjee, P.; Pavón, D. Checking the second law at cosmic scales. *J. Cosmol. Astropart. Phys.* **2023**, *2023*, 92. [CrossRef]
90. Banerjee, N.; Mukherjee, P.; Pavón, D. Spatial curvature and thermodynamics. *Mon. Not. R. Astron. Soc.* **2023**, *521*, 5473–5482. [CrossRef]
91. Adak, D.; Hazra, D.K.; Mitra, S.; Krishak, A. Model-independent Reconstruction of UV Luminosity Function and Reionization Epoch. *arXiv* **2024**, arXiv:2405.10180. [CrossRef]
92. Krishak, A.; Hazra, D.K. Gaussian Process Reconstruction of Reionization History. *Astrophys. J.* **2021**, *922*, 95. [CrossRef]
93. Mukherjee, P.; Dey, A.; Pal, S. What can we learn about Reionization astrophysical parameters using Gaussian Process Regression? *arXiv* **2024**, arXiv:2407.19481.
94. Buchanan, J.J.; Schneider, M.D.; Armstrong, R.E.; Muyskens, A.L.; Priest, B.W.; Dana, R.J. Gaussian Process Classification for Galaxy Blend Identification in LSST. *Astrophys. J.* **2022**, *924*, 94. [CrossRef]
95. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; Adaptive Computation and Machine Learning; MIT Press: Cambridge, MA, USA, 2006; pp. 1–248.
96. Paciorek, C.; Schervish, M. Nonstationary covariance functions for Gaussian process regression. *Adv. Neural Inf. Process. Syst.* **2003**, *16*, 273–280.
97. Paciorek, C.J.; Schervish, M.J. Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics Off. J. Int. Environmetrics Soc.* **2006**, *17*, 483–506. [CrossRef] [PubMed]
98. Noack, M.M.; Sethian, J.A. Advanced stationary and nonstationary kernel designs for domain-aware gaussian processes. *Commun. Appl. Math. Comput. Sci.* **2022**, *17*, 131–156. [CrossRef]
99. Noack, M.M.; Luo, H.; Risser, M.D. A unifying perspective on non-stationary kernels for deeper Gaussian processes. *APL Mach. Learn.* **2024**, *2*, 010902. [CrossRef]
100. GPy. GPy: A Gaussian Process Framework in Python. 2012. Available online: http://github.com/SheffieldML/GPy (accessed on 20 November 2024).
101. van der Wilk, M.; Dutordoir, V.; John, S.; Artemev, A.; Adam, V.; Hensman, J. A Framework for Interdomain and Multioutput Gaussian Processes. *arXiv* **2020**, arXiv:2003.01115.
102. Gardner, J.R.; Pleiss, G.; Bindel, D.; Weinberger, K.Q.; Wilson, A.G. GPyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 7587–7597.
103. Abril-Pla1, O.; Andreani, V.; Carroll, C.; Dong, L.; Fonnesbeck, C.J.; Kochurov, M.; Kumar, R.; Lao, J.; Luhmann, C.C.; Martin, O.A.; et al. PyMC: A Modern and Comprehensive Probabilistic Programming Framework in Python. *PeerJ Comput. Sci.* **2023**, *9*, e1516. [CrossRef]
104. Ugalde, J. GP_in_Cosmology. 2023. Available online: https://github.com/JesusUg2497/GP_in_Cosmology (accessed on 10 July 2024).
105. Ó Colgáin, E.; Sheikh-Jabbari, M.M. Elucidating cosmological model dependence with H_0. *Eur. Phys. J. C* **2021**, *81*, 892. [CrossRef]

106. Sharov, A.S.; Novikov, I.D. *Edwin Hubble, the Discoverer of the Big Bang Universe*; Cambridge University Press: Cambridge, UK, 1993.
107. Perlmutter, S.; Aldering, G.; Goldhaber, G.; Knop, R.; Nugent, P.; Castro, P.G.; Deustua, S.; Fabbro, S.; Goobar, A.; Groom, D.E.; et al. Measurements of Ω and Λ from 42 high-redshift supernovae. *Astrophys. J.* **1999**, *517*, 565. [CrossRef]
108. Riess, A.G.; Filippenko, A.V.; Challis, P.; Clocchiatti, A.; Diercks, A.; Garnavich, P.M.; Gilliland, R.L.; Hogan, C.J.; Jha, S.; Kirshner, R.P.; et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.* **1998**, *116*, 1009. [CrossRef]
109. Copeland, E.J.; Sami, M.; Tsujikawa, S. Dynamics of dark energy. *Int. J. Mod. Phys. D* **2006**, *15*, 1753–1935. [CrossRef]
110. Mukhanov, V. *Physical Foundations of Cosmology*; Cambridge University Press: Cambridge, UK, 2005.
111. Peebles, P. *Principles of Physical Cosmology*; Princeton Series in Physics; Princeton University Press: Princeton, NJ, USA, 1993.
112. Hogg, D.W. Distance measures in cosmology. *arXiv* **1999**, arXiv:astro-ph/9905116.
113. Alam, U.; Sahni, V.; Saini, T.D.; Starobinsky, A.A. Exploring the expanding universe and dark energy using the Statefinder diagnostic. *Mon. Not. R. Astron. Soc.* **2003**, *344*, 1057. [CrossRef]
114. Sahni, V.; Saini, T.D.; Starobinsky, A.A.; Alam, U. Statefinder: A New geometrical diagnostic of dark energy. *JETP Lett.* **2003**, *77*, 201–206. [CrossRef]
115. Gómez-Valent, A.; Amendola, L. H0 from cosmic chronometers and Type Ia supernovae, with Gaussian Processes and the novel Weighted Polynomial Regression method. *J. Cosmol. Astropart. Phys.* **2018**, *2018*, 51. [CrossRef]
116. Vagnozzi, S.; Loeb, A.; Moresco, M. Eppur è piatto? The Cosmic Chronometers Take on Spatial Curvature and Cosmic Concordance. *Astrophys. J.* **2021**, *908*, 84. [CrossRef]
117. Scolnic, D.; Brout, D.; Carr, A.; Riess, A.G.; Davis, T.M.; Dwomoh, A.; Jones, D.O.; Ali, N.; Charvu, P.; Chen, R.; et al. The Pantheon+ Analysis: The Full Data Set and Light-curve Release. *Astrophys. J.* **2022**, *938*, 113. [CrossRef]
118. Brout, D.; Scolnic, D.; Popovic, B.; Riess, A.G.; Carr, A.; Zuntz, J.; Kessler, R.; Davis, T.M.; Hinton, S.; Jones, D.; et al. The Pantheon+ Analysis: Cosmological Constraints. *Astrophys. J.* **2022**, *938*, 110. [CrossRef]
119. Alam, S.; Aubert, M.; Avila, S.; Balland, C.; Bautista, J.E.; Bershady, M.A.; Bizyaev, D.; Blanton, M.R.; Bolton, A.S.; Bovy, J.; et al. Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory. *Phys. Rev. D* **2021**, *103*, 083533. [CrossRef]
120. Adame, A.G.; Aguilar, J.; Ahlen, S.; Alam, S.; Alexander, D.M.; Alvarez, M.; Alves, O.; Anand, A.; Andrade, U.; Armengaud, E.; et al. DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations. *arXiv* **2024**, arXiv:2404.03002.
121. Adame, A.G.; Aguilar, J.; Ahlen, S.; Alam, S.; Alexander, D.M.; Alvarez, M.; Alves, O.; Anand, A.; Andrade, U.; Armengaud, E.; et al. DESI 2024 III: Baryon Acoustic Oscillations from Galaxies and Quasars. *arXiv* **2024**, arXiv:2404.03000.
122. Hwang, S.g.; L'Huillier, B.; Keeley, R.E.; Jee, M.J.; Shafieloo, A. How to use GP: Effects of the mean function and hyperparameter selection on Gaussian process regression. *J. Cosmol. Astropart. Phys.* **2023**, *2023*, 14. [CrossRef]
123. Haridasu, B.S.; Luković, V.V.; Moresco, M.; Vittorio, N. An improved model-independent assessment of the late-time cosmic expansion. *J. Cosmol. Astropart. Phys.* **2018**, *2018*, 15. [CrossRef]
124. Dinda, B.R.; Banerjee, N. Model independent bounds on type Ia supernova absolute peak magnitude. *Phys. Rev. D* **2023**, *107*, 063513. [CrossRef]
125. Escamilla, L.A.; Vazquez, J.A. Model selection applied to reconstructions of the Dark Energy. *Eur. Phys. J. C* **2023**, *83*, 251. [CrossRef]
126. Kendall, M.; Stuart, A.; Ord, J.; Arnold, S.; O'Hagan, A. *Kendall's Advanced Theory of Statistics, Classical Inference and the Linear Model*; Number v. 2 in A Hodder Arnold Publication; Wiley: Hoboken, NJ, USA, 1994.
127. Lin, H.N.; Li, X.; Tang, L. Non-parametric reconstruction of dark energy and cosmic expansion from the Pantheon compilation of type Ia supernovae. *Chin. Phys. C* **2019**, *43*, 075101. [CrossRef]
128. Gómez-Valent, A. Quantifying the evidence for the current speed-up of the Universe with low and intermediate-redshift data. A more model-independent approach. *J. Cosmol. Astropart. Phys.* **2019**, *2019*, 26. [CrossRef]
129. Alberto Vazquez, J.; Bridges, M.; Hobson, M.P.; Lasenby, A.N. Reconstruction of the Dark Energy equation of state. *J. Cosmol. Astropart. Phys.* **2012**, *2012*, 20. [CrossRef]
130. Hee, S.; Vázquez, J.A.; Handley, W.J.; Hobson, M.P.; Lasenby, A.N. Constraining the dark energy equation of state using Bayes theorem and the Kullback–Leibler divergence. *Mon. Not. R. Astron. Soc.* **2017**, *466*, 369–377. [CrossRef]
131. Gerardi, F.; Martinelli, M.; Silvestri, A. Reconstruction of the Dark Energy equation of state from latest data: the impact of theoretical priors. *J. Cosmol. Astropart. Phys.* **2019**, *2019*, 42. [CrossRef]
132. Shah, R.; Bhaumik, A.; Mukherjee, P.; Pal, S. A thorough investigation of the prospects of eLISA in addressing the Hubble tension: Fisher forecast, MCMC and Machine Learning. *J. Cosmol. Astropart. Phys.* **2023**, *2023*, 38. [CrossRef]
133. O'Neill, M.E. *PCG: A Family of Simple Fast Space-Efficient Statistically Good Algorithms for Random Number Generation*; Technical Report HMC-CS-2014-0905; Harvey Mudd College: Claremont, CA, USA, 2014.
134. Luscher, M. A Portable high quality random number generator for lattice field theory simulations. *Comput. Phys. Commun.* **1994**, *79*, 100–110. [CrossRef]
135. James, F. RANLUX: A FORTRAN implementation of the high quality pseudorandom number generator of Luscher. *Comput. Phys. Commun.* **1994**, *79*, 111–114. Erratum in *Comput. Phys. Commun.* **1996**, *97*, 357. [CrossRef]

136. Shchur, L.N.; Butera, P. The RANLUX generator: Resonances in a random walk test. *Int. J. Mod. Phys. C* **1998**, *9*, 607–624. [CrossRef]
137. Godłowski, W. Remarks on the Methods of Investigations of Alignment Of Galaxies. *Astrophys. J.* **2012**, *747*, 7. [CrossRef]
138. Pajowska, P.; Godłowski, W.; Zhu, Z.H.; Popiela, J.; Panko, E.; Flin, P. Investigation of the orientation of galaxies in clusters: The importance, methods and results of research. *J. Cosmol. Astropart. Phys.* **2019**, *2019*, 5. [CrossRef]
139. Liu, Z.E.; Qin, H.F.; Zhang, J.; Zhang, T.J.; Yu, H.R. Reconstructing equation of state of dark energy with principal component analysis. *Phys. Dark Universe* **2019**, *26*, 100379. [CrossRef]
140. Mamon, A.A.; Das, S. A parametric reconstruction of the deceleration parameter. *Eur. Phys. J. C* **2017**, *77*, 495. [CrossRef]
141. Gadbail, G.N.; Mandal, S.; Sahoo, P.K. Parametrization of Deceleration Parameter in f(Q) Gravity. *Physics* **2022**, *4*, 1403–1412. [CrossRef]
142. Xu, B.; Xia, L.-X. Reconstructing the evolution of deceleration parameter with the non-parametric Bayesian method. *Astrophys. Space Sci.* **2020**, *365*, 44. [CrossRef]