
Constraining cosmological parameters from N-body simulations with Bayesian Neural Networks

Héctor J. Hortúa

Semillero de investigación en Data Science,
Escuela de Ciencias Básicas Tecnologías e Ingenierías ECBTI,
Universidad Nacional Abierta y a Distancia UNAD
CEAD José Acevedo y Gómez - Bogotá- Colombia
hector.hortua@unad.edu.co

Abstract

In this paper we use The Quijote simulations in order to extract the cosmological parameters through Bayesian Neural Networks. This kind of models has a remarkable ability of estimating the associated uncertainty, which is one of the ultimate goals in the precision cosmology era. We demonstrate the advantages of BNNs for extracting more complex output distributions and non-Gaussianities information from the simulations.

1 Introduction

Cosmological N-body simulations provide a unique means to test gravity on large-scales. It opens the possibility to fully explore the growth of structure in both the linear and non-linear regime [1]. Currently, the concordance cosmological model, Λ -Cold Dark Matter (Λ CDM), gives an accurate description of most the observations from early universe as well as late stages using a set of six parameters. The most recent observations from Cosmic Microwave Background (CMB) and cosmic distance ladder estimations have shown a tension between some of the parameters: σ_8 being the amplitude of the power spectrum normalization and h being the normalized Hubble constant. At the same time exists a well known degeneracy with the total non-relativistic matter density parameter Ω_m [2].

Recently, Neural networks have been used on a variety of tasks because of their potential to solving under determined inverse problems. However, neural networks are prone to overfitting due to the excessive number of parameters to be adjusted and do not provide a measure of uncertainty for its predictions. This problem can be addressed by following a probabilistic approach that permit quantifying the uncertainty of the model predictions [3, 4]. In order to do that, The Quijote simulations [5] are used to extract cosmological information via two approaches: a traditional neural network and a Bayesian one, which has the advantage of estimating the associated uncertainty. The main contributions of this paper are the following:

1. Propose probabilistic neural networks approaches for cosmological analysis and provide alternative techniques for extracting complex information from simulations.
2. Compare the performance and advantages of BNNs with respect to the traditional architectures in physics.

2 Dataset and Network

In this work, we leverage 2000 latin-hypercubes simulation taken from The Quijote project [5]. These simulations vary the value of each parameter as following: $\Omega_m \in [0.1, 0.5]$, $h \in [0.6, 0.9]$

and $\sigma_8 \in [0.6, 1]$. To account for the presence of cosmic variance, the simulations were running with different random seeds and high-resolution [5]. We make use of 32^3 - and 64^3 -volxes as inputs for the Neural Net (interpolated from the standard resolution simulations 512^3 points) that contain information of the three-dimensional density field. The output of the network consists in nine neurons: three for the mean of the three parameters, and six for the elements of the covariant matrix. Those outputs fed the next layer which is a Multivariate Gaussian distribution. TensorFlow and TensorFlow-probability¹ were used for building the traditional and stochastic Neural Networks. We have used the Mean Square Error (MSE) as loss function for the deterministic network, while for BNN the minimization is applied to the negative-log-likelihood of the distribution on the top of the network. In both scenarios, the coefficient of determination (R^2) was used to assess the performance:

$$R^2 = 1 - \frac{\sum_i (\bar{\mu}(x_i) - \bar{y})^2}{\sum_i (y_i - \bar{y})^2}, \quad (1)$$

where $\bar{\mu}(x_i)$ are the predicted values of the trained Network, \bar{y} is the average of the true parameters and the summations are performed over the entire test set. R^2 ranges from 0 to 1, where 1 represents perfect inference.

3 Method

The architecture employed in this work is shown in Fig.3 where the weights were updated via backpropagation. Then, after obtaining a decent performance in the model, we froze the entire network and put on the top a dense layer which leverages the Flipout estimator to make a Monte Carlo approximation of the distribution in the kernel and bias [6]. BNNs are useful for obtaining the aleatoric uncertainties that represent the intrinsic randomness in the input dataset and they can be reduced enhancing the quality of the data [4]. On the other hand, epistemic uncertainty quantifies the ignorance about the correct model that generated the data, it includes the uncertainty in the model, parameters, and convergence. This uncertainty is caused by the limited training data with respect to the entire feature space. Collecting more data in regions where there is a low density of training examples will reduce this uncertainty, while the aleatoric will remain unchanged. Methods for estimating epistemic uncertainties are different from the aleatoric ones, and this is where BNNs can offer a mathematically grounded base for computing this uncertainty and be able to estimate the performance of the model [4].

4 Results

Fig.1 reports the prediction versus the true values for samples. As we observed, the model can capture most of the information encoded in the dataset. The performance can be also seen in Table1, where we observed that σ_8 is predicted with high accuracy as it was also reported in [7]. On the other hand, the Bayesian Neural Networks offer better probabilistic inference estimates, in terms of their consideration of uncertainty. The results of this model are displayed in Fig.2. Here, a sample from the test set was taken in order to visualize the parameter constraints. Correlations among parameters are displayed in the triangular plot, which permit to analyze the features of the parameters and describe cosmological probes which allow to break degeneracies. Additionally, during inference, the model provides faster samples with respect to classical methods like MCMC which can bring significant advantages in huge datasets. Normalizing flows and other architectures like Resnet-3D are other ingredients used during this preliminary research which will be described in more detail in a future report.

5 Discussion

In cosmology, the large-scale, three-dimensional matter distribution, modeled with N-body simulations, plays an important role in understanding the evolution of the universe. In this work we have shown the use of deep convolutional and Bayesian neural network to estimate cosmological parameters from those N-body simulations. Finally, other architectures like ResNet-3D was also used

¹<https://www.tensorflow.org/>

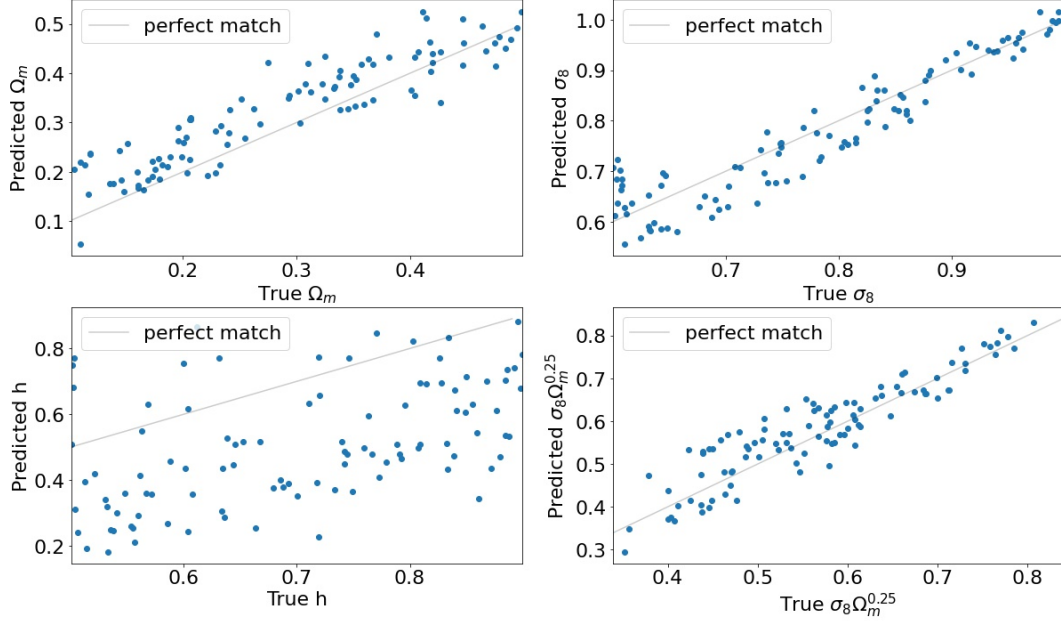


Figure 1: Plot of predicted values of three parameters over the target values of the convolutional neural network. The solid blue line represents a perfect predictor.

Table 1: Best performance described by coefficient of determination R^2 reached for convolutional and Bayesian neural network; along with the constrains for the parameters at 95% C.L. predicted by the BNN. True values correspond to the real values for the sample used for the BNN test visualization.

Parameter	95% limits	True value	R^2 BNN	R^2 CNN
Ω_m	$0.129^{+0.067}_{-0.068}$	0.115	0.833	0.871
σ_8	$0.935^{+0.084}_{-0.084}$	0.955	0.881	0.915
h	$0.582^{+0.098}_{-0.097}$	0.564	0.401	0.456
$\sigma_8 \Omega_m^{0.25}$	$0.555^{+0.080}_{-0.094}$	0.557	—	—
$\Omega_m h^2$	$0.044^{+0.029}_{-0.027}$	0.368	—	—

to carry out this task along with masked autoregressive flows. The results driven by those additional ingredients will be shown in a future publication.

References

- [1] Borgani, S. and Kravtsov, A. (2011). Cosmological simulations of galaxy clusters. *Advanced Science Letters*, 4(2), 204-227.
- [2] Tinker, J. L., Sheldon, E. S., Wechsler, R. H. et al. (2011). Cosmological constraints from galaxy clustering and the mass-to-number ratio of galaxy clusters. *The Astrophysical Journal*, 745(1), 16.
- [3] Daniel T. Chang. Bayesian Neural Networks: Essentials. *arXiv:2106.13594* (2021).
- [4] Hortúa, Héctor J. and Volpi, Riccardo and Marinelli, Dimitri and Malagò, Luigi. Parameter estimation for the cosmic microwave background with Bayesian neural networks. *Phys. Rev. D* **102**, 103509 (2020).
- [5] Villaescusa-Navarro, F., Hahn, C., Massara, E., Banerjee, A., Delgado, A. M., Ramanah, D. K. et. al. (2020). The quijote simulations. *The Astrophysical Journal Supplement Series*, 250(1), 2.
- [6] Yeming Wen and Paul Vicol and Jimmy Ba and Dustin Tran and Roger Grosse. Flipout: Efficient Pseudo-Independent Weight Perturbations on Mini-Batches. *ICLR, 2018*
- [7] Lazanu, A. (2021). Extracting cosmological parameters from N-body simulations using machine learning techniques. *arXiv preprint arXiv:2106.11061*.

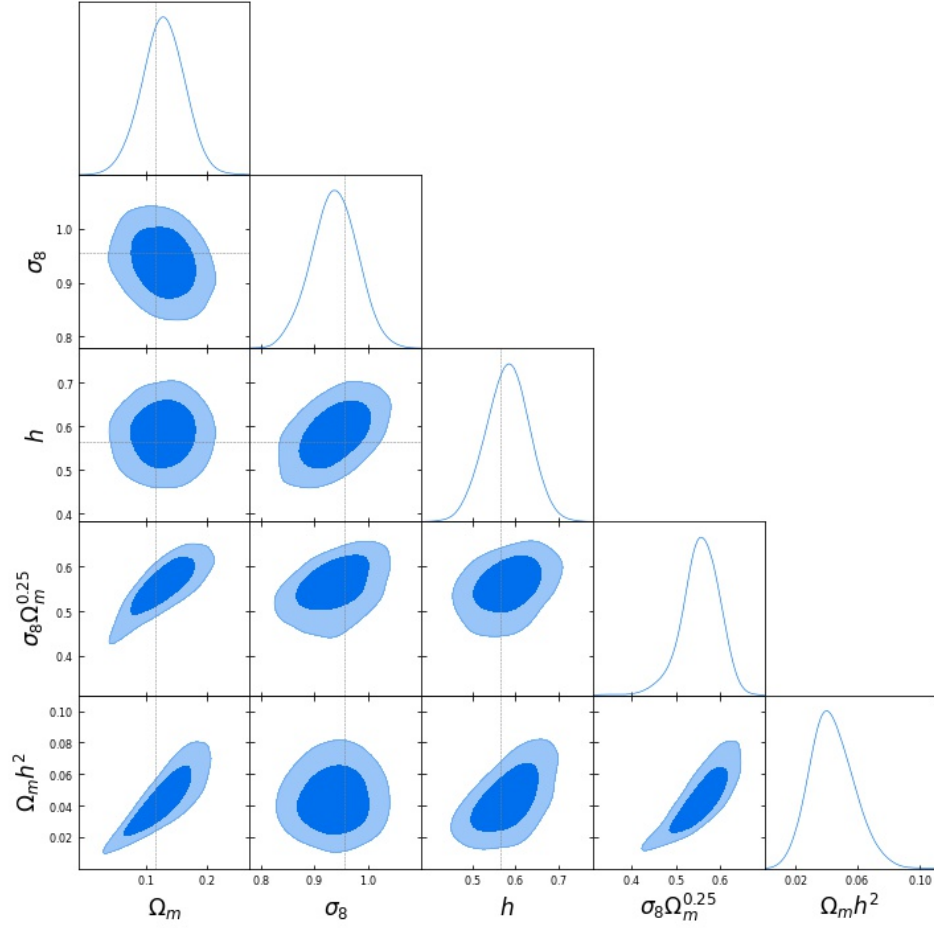


Figure 2: Constraints on the value of the cosmological parameters from the simulations. The light-blue and dark-blue regions represent the 1σ and 2σ constraints, respectively. The panels with the solid lines represent the true values.

A Appendix

Fig.3 shows the architecture used in this paper and also, works as backbone for the BNN.

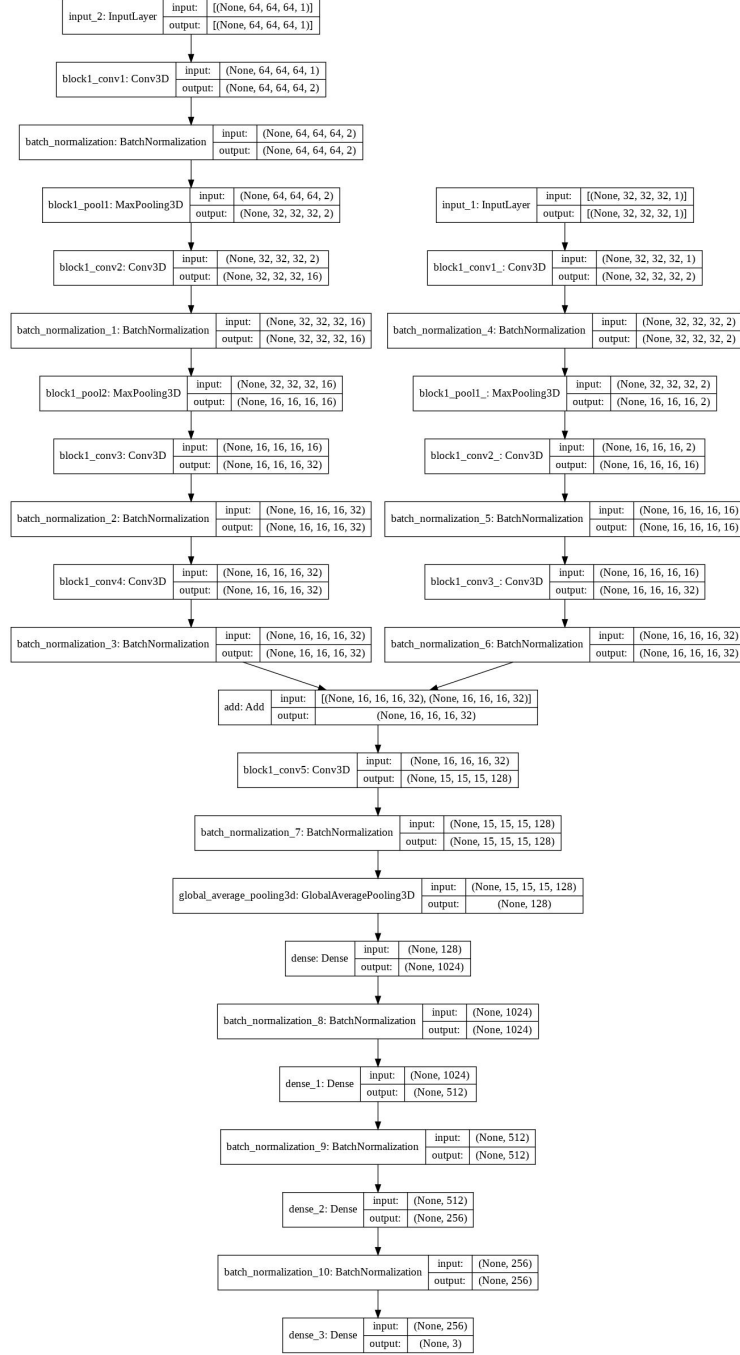


Figure 3: CNN architecture used in the paper. This is an multi-input model with 32^3 - and 64^3 -volxes.