

A MCNF Training

Mini-batch strategy

Algorithm 2 Mini-batch building scheme to train the Normalizing Flow head of MCNF

Inputs: $\mathcal{D} = \{\mathbf{x}_n, y_n\}_{n=1}^N \{y_{i,\text{MCD}}\}_{i=1}^{n_{\text{MCD}}}$

Parameters: $b \equiv$ batch size

Output: \mathfrak{d}

```

1: Initialize  $\mathfrak{d} = \{\}$ , and  $\mathcal{J} = \{1, \dots, N\}$ 
2: while  $|\mathcal{J}| > 0$  do
3:   Initialize  $\mathfrak{d}_i = \{\}$ 
4:   for  $k \leftarrow 1$  to  $b$  do
5:     Generate a random index,  $n \sim \mathcal{U}(1, |\mathcal{J}|)$  and pick  $(\mathbf{x}_n, y_n)$ 
6:     Generate a random index  $i \sim \mathcal{U}(1, n_{\text{MCD}})$  and draw  $y_{n,i,\text{MCD}}$ 
7:     Calculate prediction error  $\delta_n = y_n - y_{n,i,\text{MCD}}$ 
8:     Aggregate n-th observation to current mini-batch,  $\mathfrak{d}_i = \mathfrak{d}_i \cup \{(\mathbf{x}_n, \delta_n)\}$ 
9:     Update indices set,  $\mathcal{J} = \mathcal{J} \setminus \{n\}$ 
10:  Update mini-batches set,  $\mathfrak{d} = \mathfrak{d} \cup \mathfrak{d}_i$ 

```

During MCNF training, in order to effectively propagate epistemic uncertainty to $p_{\theta,\psi}(y|\mathbf{x}, \mathcal{D})$ when the latter is approximated as in Equation (6), we bootstrap the prediction errors. To this end, each mini-batch used to evaluate the gradients of the module based on the NFs is obtained by bootstrapping on a set of n_{MCD} samples, $\{y_{i,\text{MCD}}\}_{i=1}^{n_{\text{MCD}}}$, previously generated from the distribution $p(y_{\text{MCD}}|\mathbf{x})$. Thus, the original training dataset $\mathcal{D} = \{\mathbf{x}_n, y_n\}_{n=1}^N$ is further processed at each iteration of the training steps to recalculate the prediction error δ_n , as indicated in Algorithm 2. Consequently, the implemented mini-batch construction strategy enables every mini-batch to have a different realization of the prediction error per observation at each iteration.

B Synthetic dataset

Romano-Mod dataset. A univariate stochastic process is introduced, by adapting the equation provided in Appendix B in [14]. The distribution that characterizes the uncertainty incorporates heteroskedasticity, which is a particularly relevant validation case to test whether the method adapts correctly to the local distribution of the data. The updated stochastic process is given:

$$y = \text{Poisson}(\sin x + \Delta) + (\beta\epsilon_1 + b) \cdot x + \delta(u \leq \bar{u}) \cdot \gamma \cdot \epsilon_2 \quad (9)$$

where $\delta(\cdot)$ represents the Delta Kronecker function and ϵ_1 and ϵ_2 follow a standard Gaussian distribution. The parameters of the model are β and Δ , which condition the heteroskedasticity of the uncertainty distribution, as well as γ and \bar{u} , which control the magnitude and rate of outliers in the sample generated. We introduce b to induce a linear correlation between the predictor x and the predicted variable y . Table 2 summarizes the values used for these parameters and Figure 4 illustrates the data generated.

Table 2: Parameters for the creation of the synthetic dataset.

Data	Parameters			
	Δ	β	b	\bar{u}
Romano-Original	0.1	0.05	0.0	0.0
Romano-Mod	0.1	0.05	2.0	0.0

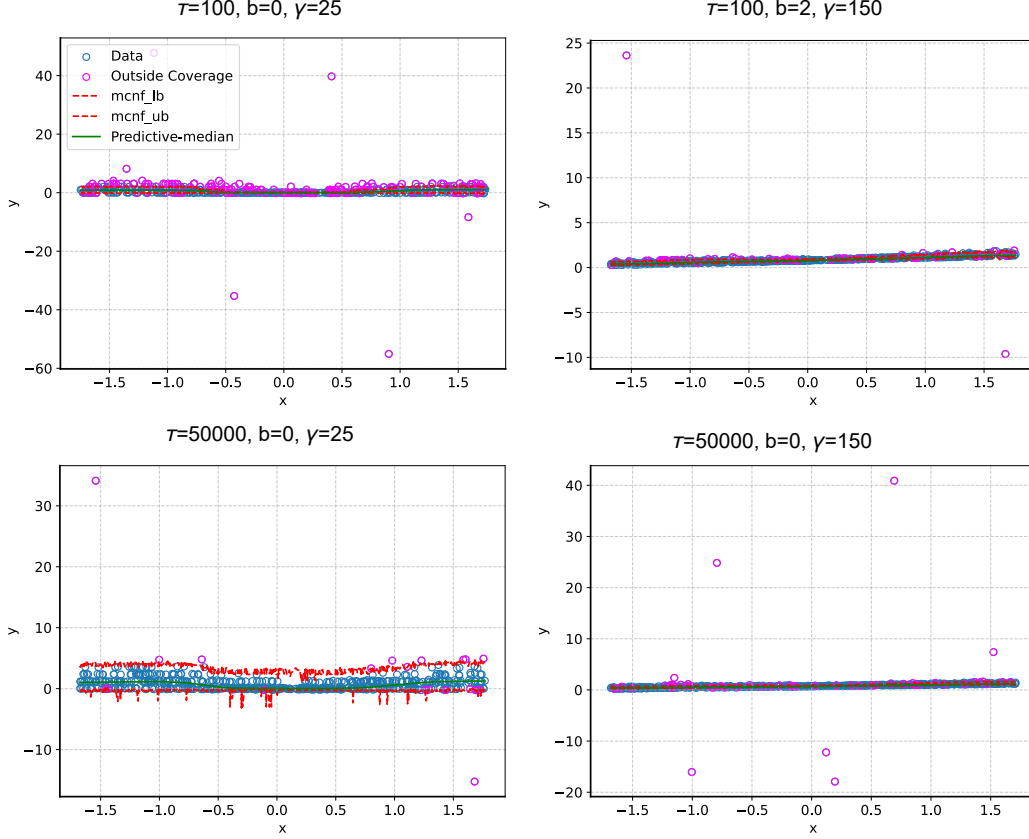


Figure 4: Visualisation of the synthetic data produced based on Equation (9).

C MCNF performance details

C.1 Adaptivity of prediction intervals

In terms of prediction interval adaptivity, all examined UQ methods, including MCNF, are sensitive to the uncertainty inherent in the data, as evidenced by the variability of the quantiles of the interval sizes shown in Table 3. MCD has the narrowest intervals but covers a much smaller proportion of the true values, achieving a very poor coverage overall. Considering the UQ methods that perform well and converge to the expected marginal coverage (90%), MCNF provides the narrowest intervals and maintains similar marginal coverage to the other methods. Quantile MAE (MAE_q), where $MAE_q(X, Y) = \frac{1}{N} \sum_{i=1}^N |y_i - q_{0.05}(x_i)| + |y_i - q_{0.95}(x_i)|$, which provides additional evidence of the improved trade-off between coverage and interval sizes for MCNF, exhibits smaller values for similar marginal coverages.

C.2 Ablation study

C.2.1 Hyperparameter study

We also examined the performance of MCNF for the following hyperparameters: number of epochs (epochs $\in \{20, 50, 100, 150\}$), number of normalizing flow samples ($n_{NF} \in \{100, 200, 500\}$), and number of Monte-Carlo Dropout samples ($n_{MCD} \in \{50, 100, 150\}$). For this evaluation, we employ the same prediction model across all tests, which was selected based on the lowest RMSE value, as described in Section 5.1.

Initially, we evaluate the number of epochs specified for training MCNF, examining epochs $\in \{20, 50, 100, 150\}$, as shown in Figure 5, whilst keeping the number of normalizing flow samples and Monte-Carlos Dropout samples fixed to $n_{NF} = 500$ and $n_{MCD} = 50$, respectively. Figure 6 illustrates

Table 3: Prediction interval sizes $\tilde{\Delta}_v(X, Y)$ for $v \in \{0.05, 0.25, 0.5, 0.75, 0.95\}$, where $\tilde{\Delta}_v(X, Y) = Q_{i=1}^n(q_{0.95}(x_i) - q_{0.05}(x_i))[v]$ and Quantile MAE, where $\text{MAE}_q(X, Y) = \frac{1}{N} \sum_{i=1}^N |y_i - q_{0.05}(x_i)| + |y_i - q_{0.95}(x_i)|$ for all UQ methods and datasets.

Data Metric	CQR	DQR	MCCP	MCD	MCQR	MCNF	NF	
Boston H.	$\tilde{\Delta}_{0.05}$	0.650 \pm 0.481	0.374 \pm 0.031	0.605 \pm 0.409	0.172 \pm 0.018	0.373 \pm 0.030	0.300 \pm 0.027	0.163 \pm 0.019
	$\tilde{\Delta}_{0.25}$	0.730 \pm 0.482	0.458 \pm 0.030	0.690 \pm 0.407	0.212 \pm 0.018	0.456 \pm 0.029	0.352 \pm 0.033	0.218 \pm 0.030
	$\tilde{\Delta}_{0.5}$	0.820 \pm 0.481	0.543 \pm 0.034	0.777 \pm 0.404	0.254 \pm 0.023	0.541 \pm 0.032	0.409 \pm 0.038	0.277 \pm 0.034
	$\tilde{\Delta}_{0.75}$	0.971 \pm 0.494	0.695 \pm 0.058	0.932 \pm 0.400	0.318 \pm 0.029	0.692 \pm 0.057	0.507 \pm 0.059	0.363 \pm 0.047
	$\tilde{\Delta}_{0.95}$	1.315 \pm 0.520	1.048 \pm 0.084	1.269 \pm 0.401	0.497 \pm 0.054	1.044 \pm 0.081	0.834 \pm 0.092	0.617 \pm 0.132
	MAE _q	0.824 \pm 0.482	0.551 \pm 0.038	0.785 \pm 0.402	0.279 \pm 0.023	0.549 \pm 0.037	0.428 \pm 0.037	0.295 \pm 0.029
Concrete	$\tilde{\Delta}_{0.05}$	0.430 \pm 0.081	0.461 \pm 0.019	0.458 \pm 0.105	0.183 \pm 0.015	0.463 \pm 0.019	0.298 \pm 0.027	0.193 \pm 0.024
	$\tilde{\Delta}_{0.25}$	0.542 \pm 0.075	0.574 \pm 0.030	0.569 \pm 0.097	0.238 \pm 0.014	0.573 \pm 0.030	0.394 \pm 0.025	0.283 \pm 0.024
	$\tilde{\Delta}_{0.5}$	0.660 \pm 0.078	0.689 \pm 0.033	0.682 \pm 0.102	0.292 \pm 0.016	0.688 \pm 0.034	0.491 \pm 0.031	0.366 \pm 0.026
	$\tilde{\Delta}_{0.75}$	0.832 \pm 0.099	0.859 \pm 0.052	0.852 \pm 0.115	0.382 \pm 0.028	0.858 \pm 0.052	0.621 \pm 0.053	0.469 \pm 0.024
	$\tilde{\Delta}_{0.95}$	1.045 \pm 0.103	1.076 \pm 0.068	1.072 \pm 0.113	0.528 \pm 0.049	1.072 \pm 0.066	0.818 \pm 0.091	0.676 \pm 0.068
	MAE _q	0.636 \pm 0.079	0.667 \pm 0.040	0.663 \pm 0.102	0.329 \pm 0.026	0.666 \pm 0.039	0.472 \pm 0.032	0.352 \pm 0.020
Abalone	$\tilde{\Delta}_{0.05}$	0.357 \pm 0.044	0.367 \pm 0.020	0.349 \pm 0.045	0.077 \pm 0.007	0.370 \pm 0.020	0.294 \pm 0.017	0.277 \pm 0.030
	$\tilde{\Delta}_{0.25}$	0.446 \pm 0.046	0.457 \pm 0.025	0.439 \pm 0.048	0.109 \pm 0.007	0.459 \pm 0.025	0.380 \pm 0.019	0.371 \pm 0.039
	$\tilde{\Delta}_{0.5}$	0.574 \pm 0.047	0.586 \pm 0.026	0.569 \pm 0.048	0.132 \pm 0.007	0.587 \pm 0.025	0.514 \pm 0.020	0.507 \pm 0.050
	$\tilde{\Delta}_{0.75}$	0.748 \pm 0.043	0.757 \pm 0.027	0.739 \pm 0.048	0.163 \pm 0.007	0.758 \pm 0.027	0.744 \pm 0.040	0.724 \pm 0.045
	$\tilde{\Delta}_{0.95}$	1.117 \pm 0.037	1.128 \pm 0.038	1.109 \pm 0.046	0.249 \pm 0.013	1.127 \pm 0.040	1.098 \pm 0.052	1.092 \pm 0.046
	MAE _q	0.551 \pm 0.043	0.562 \pm 0.021	0.544 \pm 0.045	0.222 \pm 0.008	0.563 \pm 0.021	0.499 \pm 0.017	0.494 \pm 0.039
Protein	$\tilde{\Delta}_{0.05}$	0.709 \pm 0.080	0.744 \pm 0.078	0.720 \pm 0.081	0.113 \pm 0.015	0.749 \pm 0.080	0.341 \pm 0.046	0.270 \pm 0.048
	$\tilde{\Delta}_{0.25}$	1.549 \pm 0.035	1.585 \pm 0.026	1.557 \pm 0.032	0.227 \pm 0.019	1.584 \pm 0.026	1.244 \pm 0.111	1.155 \pm 0.152
	$\tilde{\Delta}_{0.5}$	1.827 \pm 0.022	1.861 \pm 0.019	1.828 \pm 0.019	0.406 \pm 0.011	1.857 \pm 0.019	1.781 \pm 0.033	1.673 \pm 0.062
	$\tilde{\Delta}_{0.75}$	2.010 \pm 0.023	2.044 \pm 0.022	2.009 \pm 0.020	0.524 \pm 0.009	2.038 \pm 0.022	2.061 \pm 0.026	1.933 \pm 0.044
	$\tilde{\Delta}_{0.95}$	2.249 \pm 0.032	2.284 \pm 0.036	2.250 \pm 0.032	0.762 \pm 0.015	2.279 \pm 0.038	2.393 \pm 0.043	2.230 \pm 0.043
	MAE _q	1.349 \pm 0.048	1.383 \pm 0.041	1.354 \pm 0.047	0.576 \pm 0.014	1.382 \pm 0.041	1.172 \pm 0.059	1.048 \pm 0.088
Wave	$\tilde{\Delta}_{0.05}$	0.004 \pm 0.001	0.010 \pm 0.001	0.004 \pm 0.001	0.005 \pm 0.000	0.010 \pm 0.001	0.008 \pm 0.001	0.003 \pm 0.001
	$\tilde{\Delta}_{0.25}$	0.007 \pm 0.001	0.013 \pm 0.002	0.007 \pm 0.001	0.007 \pm 0.000	0.013 \pm 0.002	0.010 \pm 0.001	0.004 \pm 0.001
	$\tilde{\Delta}_{0.5}$	0.010 \pm 0.001	0.016 \pm 0.002	0.010 \pm 0.001	0.009 \pm 0.000	0.016 \pm 0.002	0.013 \pm 0.001	0.006 \pm 0.001
	$\tilde{\Delta}_{0.75}$	0.020 \pm 0.001	0.026 \pm 0.002	0.020 \pm 0.001	0.012 \pm 0.000	0.026 \pm 0.002	0.022 \pm 0.002	0.016 \pm 0.002
	$\tilde{\Delta}_{0.95}$	0.040 \pm 0.002	0.046 \pm 0.003	0.041 \pm 0.002	0.021 \pm 0.001	0.047 \pm 0.003	0.043 \pm 0.004	0.039 \pm 0.005
	MAE _q	0.009 \pm 0.001	0.015 \pm 0.002	0.010 \pm 0.001	0.009 \pm 0.0002	0.016 \pm 0.002	0.013 \pm 0.001	0.006 \pm 0.001
Super	$\tilde{\Delta}_{0.05}$	0.163 \pm 0.015	0.196 \pm 0.012	0.165 \pm 0.014	0.058 \pm 0.007	0.196 \pm 0.012	0.141 \pm 0.014	0.096 \pm 0.008
	$\tilde{\Delta}_{0.25}$	0.286 \pm 0.015	0.318 \pm 0.012	0.289 \pm 0.015	0.111 \pm 0.006	0.319 \pm 0.011	0.270 \pm 0.024	0.220 \pm 0.012
	$\tilde{\Delta}_{0.5}$	0.847 \pm 0.032	0.879 \pm 0.028	0.857 \pm 0.028	0.270 \pm 0.014	0.888 \pm 0.026	0.791 \pm 0.035	0.679 \pm 0.044
	$\tilde{\Delta}_{0.75}$	1.769 \pm 0.058	1.802 \pm 0.061	1.768 \pm 0.059	0.479 \pm 0.020	1.799 \pm 0.061	1.471 \pm 0.089	1.336 \pm 0.078
	$\tilde{\Delta}_{0.95}$	2.074 \pm 0.060	2.105 \pm 0.063	2.066 \pm 0.059	0.735 \pm 0.031	2.097 \pm 0.064	2.087 \pm 0.106	1.988 \pm 0.071
	MAE _q	0.610 \pm 0.025	0.641 \pm 0.025	0.612 \pm 0.029	0.292 \pm 0.007	0.642 \pm 0.024	0.566 \pm 0.023	0.462 \pm 0.022
R-OG	$\tilde{\Delta}_{0.05}$	1.764 \pm 0.181	1.698 \pm 0.145	1.751 \pm 0.207	0.110 \pm 0.040	1.708 \pm 0.146	1.832 \pm 0.345	1.846 \pm 0.269
	$\tilde{\Delta}_{0.25}$	2.374 \pm 0.225	2.309 \pm 0.200	2.361 \pm 0.222	0.138 \pm 0.046	2.316 \pm 0.196	2.852 \pm 0.193	2.785 \pm 0.239
	$\tilde{\Delta}_{0.5}$	3.631 \pm 0.207	3.567 \pm 0.159	3.604 \pm 0.181	0.348 \pm 0.037	3.573 \pm 0.159	3.321 \pm 0.232	3.371 \pm 0.234
	$\tilde{\Delta}_{0.75}$	4.262 \pm 0.245	4.198 \pm 0.225	4.249 \pm 0.226	0.458 \pm 0.035	4.209 \pm 0.223	4.211 \pm 0.091	4.093 \pm 0.124
	$\tilde{\Delta}_{0.95}$	4.516 \pm 0.297	4.446 \pm 0.276	4.505 \pm 0.281	0.557 \pm 0.037	4.463 \pm 0.276	4.373 \pm 0.098	4.296 \pm 0.129
	MAE _q	2.615 \pm 0.350	2.583 \pm 0.350	2.597 \pm 0.242	1.117 \pm 0.166	2.589 \pm 0.350	2.977 \pm 0.395	2.886 \pm 0.254
R-MOD	$\tilde{\Delta}_{0.05}$	0.329 \pm 0.043	0.340 \pm 0.025	0.319 \pm 0.041	0.079 \pm 0.007	0.339 \pm 0.025	0.230 \pm 0.016	0.188 \pm 0.016
	$\tilde{\Delta}_{0.25}$	0.382 \pm 0.042	0.392 \pm 0.023	0.372 \pm 0.041	0.105 \pm 0.007	0.391 \pm 0.022	0.300 \pm 0.025	0.235 \pm 0.036
	$\tilde{\Delta}_{0.5}$	0.491 \pm 0.042	0.495 \pm 0.029	0.483 \pm 0.042	0.141 \pm 0.006	0.494 \pm 0.029	0.438 \pm 0.026	0.382 \pm 0.031
	$\tilde{\Delta}_{0.75}$	0.543 \pm 0.045	0.543 \pm 0.031	0.535 \pm 0.042	0.168 \pm 0.007	0.543 \pm 0.030	0.511 \pm 0.025	0.458 \pm 0.034
	$\tilde{\Delta}_{0.95}$	0.661 \pm 0.044	0.660 \pm 0.036	0.656 \pm 0.044	0.201 \pm 0.009	0.660 \pm 0.038	0.599 \pm 0.036	0.528 \pm 0.048
	MAE _q	0.364 \pm 0.064	0.433 \pm 0.028	0.332 \pm 0.037	0.187 \pm 0.012	0.434 \pm 0.029	0.350 \pm 0.032	0.288 \pm 0.034
Solubility	$\tilde{\Delta}_{0.05}$	1.157 \pm 0.633	0.610 \pm 0.076	1.153 \pm 0.369	0.266 \pm 0.021	0.756 \pm 0.081	0.614 \pm 0.084	0.446 \pm 0.075
	$\tilde{\Delta}_{0.25}$	1.423 \pm 0.611	0.867 \pm 0.083	1.416 \pm 0.365	0.380 \pm 0.033	1.030 \pm 0.086	0.933 \pm 0.130	0.765 \pm 0.122
	$\tilde{\Delta}_{0.5}$	1.700 \pm 0.601	1.155 \pm 0.108	1.685 \pm 0.345	0.504 \pm 0.039	1.289 \pm 0.108	1.184 \pm 0.154	1.021 \pm 0.132
	$\tilde{\Delta}_{0.75}$	2.055 \pm 0.582	1.511 \pm 0.142	2.020 \pm 0.338	0.618 \pm 0.044	1.616 \pm 0.139	1.500 \pm 0.164	1.293 \pm 0.160
	$\tilde{\Delta}_{0.95}$	2.693 \pm 0.600	2.156 \pm 0.214	2.603 \pm 0.373	0.749 \pm 0.043	2.220 \pm 0.206	1.972 \pm 0.152	1.723 \pm 0.255
	MAE _q	1.665 \pm 0.590	1.124 \pm 0.087	1.636 \pm 0.335	0.597 \pm 0.031	1.242 \pm 0.085	1.119 \pm 0.118	0.951 \pm 0.114

the results of evaluating the different normalizing flow sample values $n_{\text{NF}} \in \{100, 200, 500\}$, setting epochs = 100 and $n_{\text{MCD}} = 50$. Lastly, Figure 7 visualizes the ablation results for the different Monte-Carlo Dropout sample values $n_{\text{MCD}} \in \{50, 100, 500\}$, whilst training the MCNF for epochs = 100 and sampling $n_{\text{NF}} = 500$ normalizing flow samples. Our results demonstrate that all tested n_{MCD}

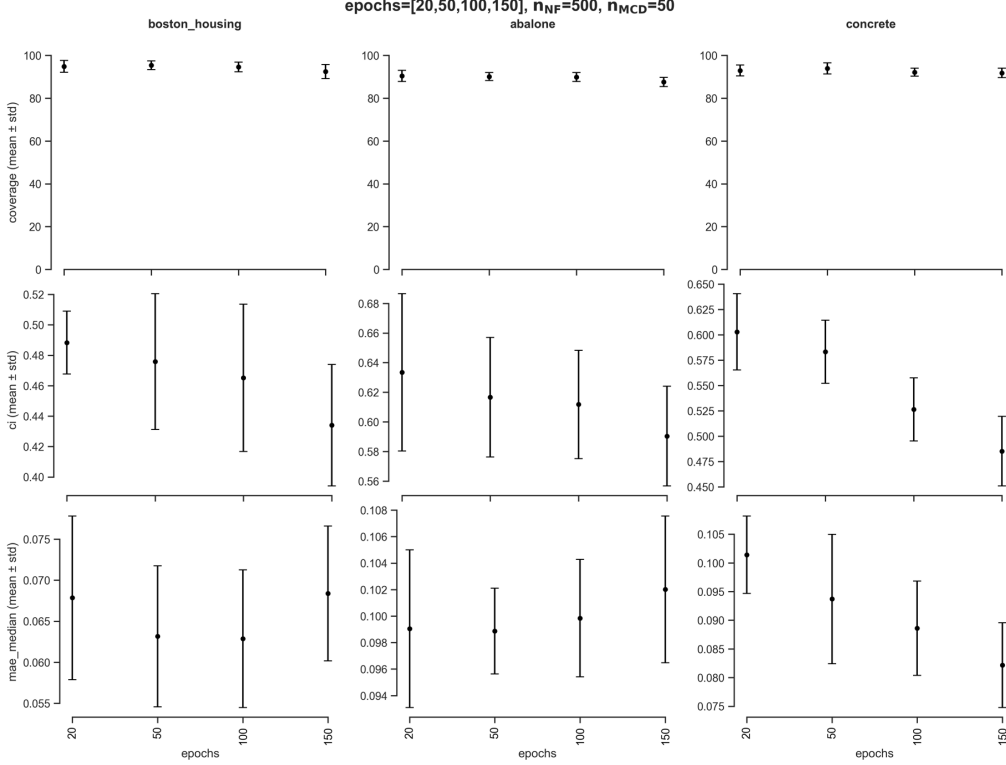


Figure 5: Evaluation of MCNF on Boston Housing, Abalone, and Concrete datasets for different epoch values (20, 50, 100, and 150).

values perform similarly in terms of the coverage metrics, and the difference in the confidence interval and the median MAE is negligible. Similar observations have been derived for the various epoch and n_{NF} values evaluated.

C.2.2 Outliers handling: testing impact of τ hyperparameter

As outlined in recent UQ survey papers [6], UQ methods can be impacted by outliers. Thus, we investigated the performance and response of MCNF to outlier data. Table 4 shows experiments carried out for different τ , b and γ value combinations on MCNF coverage (C), median MAE (\widehat{MAE}), and Quantile MAE (MAE_q). We observe that MCNF was sensitive to outliers when the effect size was close to zero, which likely results from biasing of the maximum likelihood estimator. The effect size relates to the degree of correlation between the feature variables and the response variable. We sought to mitigate this impact, given that low effect sizes (as well as sporadic large outlier values) frequently characterize data sets in many disciplines.

To test the impact of different τ values in Equation (8), with respect to changing effect size, we synthesize data (Section B) by modulating the slope (b) of the governing generative process to generate multiple sets of simulated data (Figure 8), where $b \in \{0, 0.1, 0.2, 0.3, 1, 2\}$. As the scale of the outliers relative to the rest of the data also changed with the slope, we modulated the γ parameter to keep the outliers in the same approximate range. The outliers proportion was 1%. For each level of slope, we predicted from the MCNF method using the following τ values $\tau \in \{10^2, 2 \cdot 10^2, 3 \cdot 10^2, 10^3, 2 \cdot 10^3, 5 \cdot 10^3, 10^4, 10^5, 10^6\}$. The data distributions are shown in Figure 8.

Generally, the marginal coverage increases with the value of τ as the effect of the temperature scaling assigns equal importance to outliers and non-outliers. However, this comes at the cost of overly conservative prediction intervals and less adaptivity to the uncertainty in the data. No single value for τ can be considered universally optimal, as shown in Figure 8. This hyperparameter should be fine-tuned, potentially using a calibration set. When the ratio of the outlier magnitude over the effect size becomes smaller (larger values of the slope in this case) the importance of the former on the regular maximum likelihood procedure decreases, and values of τ can be smaller.

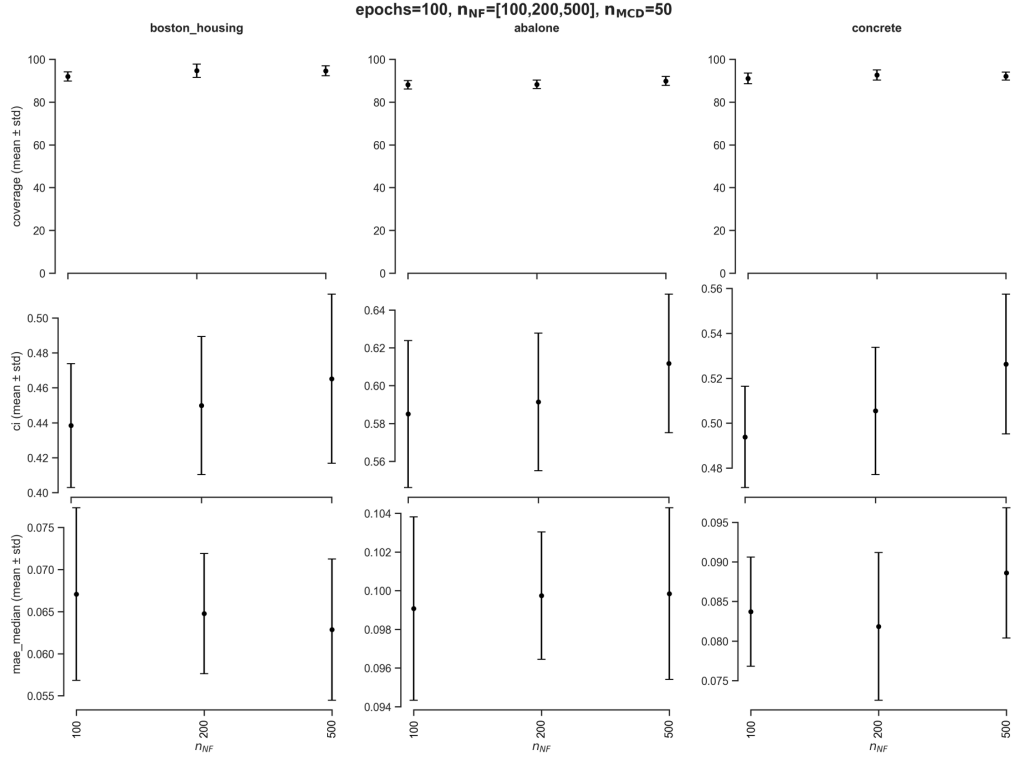


Figure 6: Evaluation of MCNF on Boston Housing, Abalone, and Concrete datasets for different n_{NF} values (100, 200, and 500).

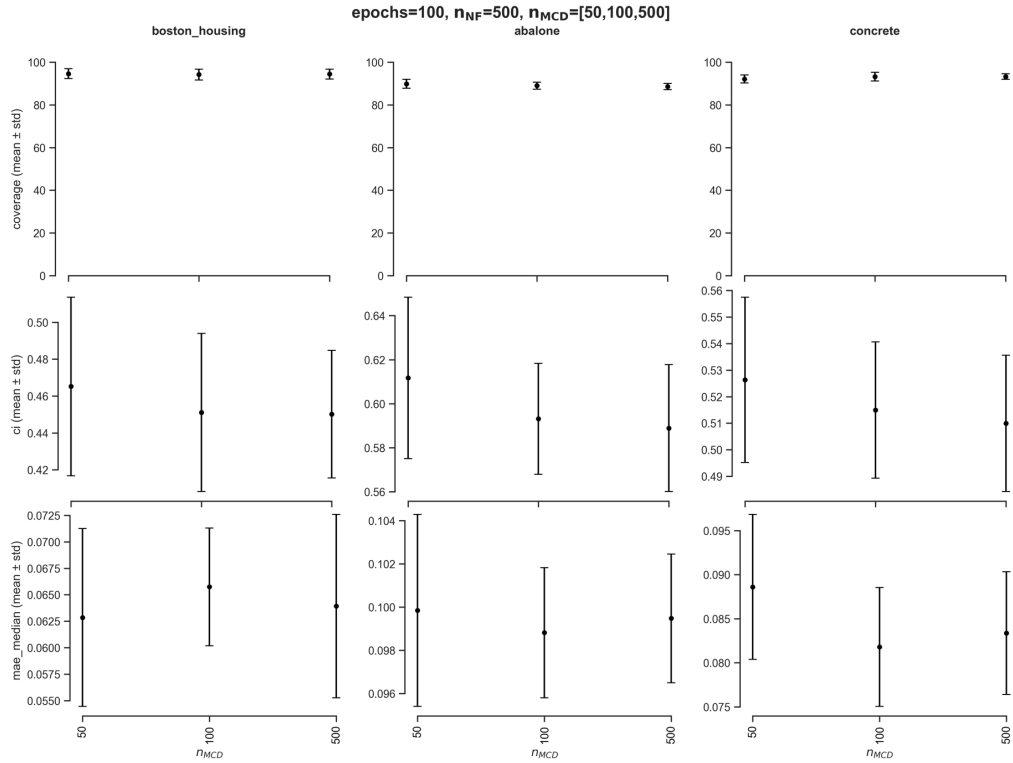


Figure 7: Evaluation of MCNF on Boston Housing, Abalone, and Concrete datasets for different n_{MCD} values (50, 100, and 500).

Table 4: Evaluation of the impact of outliers for different τ , b , and γ value combinations on MCNF coverage (C), median MAE ($\widetilde{\text{MAE}}$), and Quantile MAE (MAE_q).

τ	b	γ	C	$\widetilde{\text{MAE}}$	MAE_q
100	0	25	0.8100	0.2061	1.3039
100	0.1	25	0.7883	0.1845	0.9799
100	0.2	25	0.7900	0.2307	0.9087
100	0.3	25	0.8200	0.1271	0.7115
100	1	100	0.8917	0.1040	0.4217
100	2	150	0.9117	0.2258	0.0474
200	0	25	0.8067	0.2887	1.4055
200	0.1	25	0.8467	0.3229	1.2045
200	0.2	25	0.8567	0.1541	0.9818
200	0.3	25	0.8800	0.1594	0.8330
200	1	100	0.9350	0.1284	0.7240
200	2	150	0.9333	0.0625	0.2758
300	0	25	0.8467	0.2001	1.3449
300	0.1	25	0.8467	0.3229	1.2045
300	0.2	25	0.8783	0.1607	1.3490
300	0.3	25	0.8800	0.1594	0.8330
300	1	100	0.9283	0.1070	0.4160
300	2	150	0.9283	0.0570	0.2793
1000	0	25	0.9067	0.6610	1.9665
1000	0.1	25	0.8867	0.1307	1.1857
1000	0.2	25	0.9017	0.1413	0.9706
1000	0.3	25	0.9133	0.1875	0.8647
1000	1	100	0.9283	0.0978	0.3980
1000	2	150	0.9367	0.0640	0.2923
2000	0	25	0.9000	0.2752	1.6500
2000	0.1	25	0.9400	0.1788	1.4421
2000	0.2	25	0.9233	0.1430	1.1068
2000	0.3	25	0.9183	0.1163	0.8374
2000	1	100	0.9350	0.0807	0.4679
2000	2	150	0.9467	0.0576	0.2600
5000	0	25	0.9083	0.8049	1.6046
5000	0.1	25	0.9467	0.1651	1.2981
5000	0.2	25	0.9300	0.2350	1.1888
5000	0.3	25	0.9183	0.2916	0.7965
5000	1	100	0.9200	0.0957	0.4036
5000	2	150	0.9500	0.0439	0.3584
10000	0	25	0.9200	0.3249	3.4255
10000	0.1	25	0.9467	0.1651	1.2981
10000	0.2	25	0.9133	0.1299	1.0917
10000	0.3	25	0.9300	0.1082	0.8284
10000	1	100	0.9417	0.1113	0.5516
10000	2	150	0.9400	0.0632	0.2897
50000	0	25	0.9800	0.2216	6.4337
50000	0.1	25	0.9450	0.1666	1.7998
50000	0.2	25	0.9667	0.1667	1.8143
50000	0.3	25	0.9433	0.1299	1.2884
50000	1	100	0.9417	0.1113	0.5516
50000	2	150	0.9400	0.0599	0.3046
100000	0	25	0.9867	0.3689	7.2465
100000	0.1	25	0.9867	0.2644	6.5211
100000	0.2	25	0.9800	0.2321	2.9301
100000	0.3	25	0.9650	0.1731	1.8576
100000	1	100	0.9500	0.1085	0.5747
100000	2	150	0.9833	0.0654	0.6146

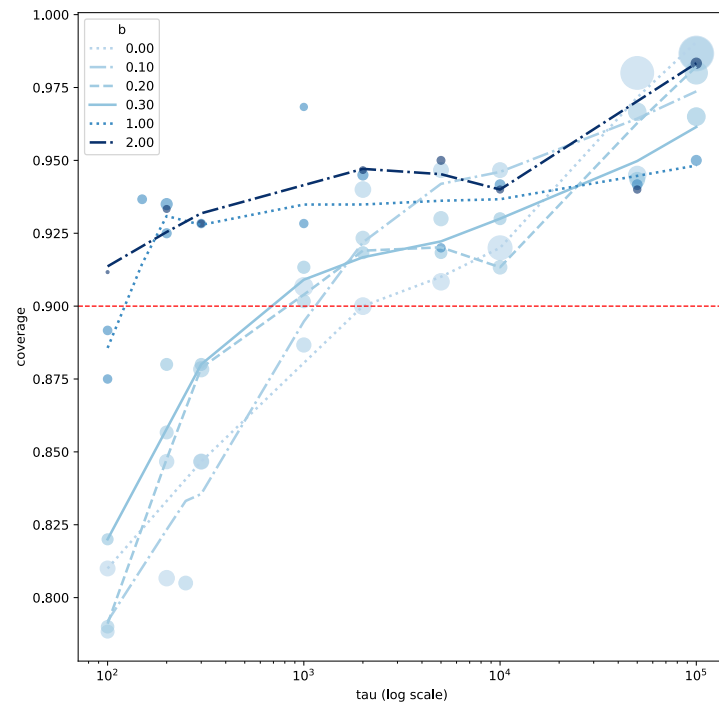


Figure 8: MCNF marginal coverage with respect to τ for different slope b values.