

Here the returns data is overlapping when rolling window is used so christopherson and DQ test are diagnostic

	VaR violations	violations	n	kupiec_p	christoffel dq_stat	dq_p	acerbi_szekely	acerbi_sz	acerbi_szmae_es	empirical_es_loss
step size 1 and 2 assets					christoffel dq_stat overlapping	dq_p diagnostic				
baseline (previous)	0.044480388			0.18305	0.96045		2.01734755	2.20501	0.4272	0.785174 -3.34341
model with kde when gpd fails	0.049737161	123	2473	0.95206	0.62365	0.22907	0.89178	2.086994488	2.19693	0.1521 0.797055 -3.30847
baseline with same mean equation	0.04286292	106	2473	0.07972	0.78533	2.69976	0.25927	1.991000895	2.20501	0.5318 0.794419 -3.37575
model with no kde when gpd fails	0.048119693	119	2473	0.66218	0.74423	1.62781	0.89787	2.031386587	2.2131	0.3599 0.798314 -3.35155
step size 1 and 4 assets	VaR violations	violations	n	kupiec_p	christoffel dq_stat	dq_p	acerbi_szekely	acerbi_sz	acerbi_szmae_es	empirical_es_loss
						diagnostic				
baseline	0.038114343			0.04999	0.23694		1.861479787	2.34403	0.8574	0.553125 -2.63056
model with kde when gpd fails	0.05135463	127	2473	0.76021	0.18006	6.07598	0.2989	2.019307138	2.20501	0.4006 0.767604 -2.94418
baseline with same mean equation	0.044076021	109	2473	0.15123	0.92677	2.20902	0.33137	2.027828434	2.20501	0.3493 0.741873 -3.08135
model with no kde when gpd fails	0.050141529	124	2473	0.97427	0.1391	5.18969	0.39317	2.009365677	2.20501	0.4542 0.782259 -2.97213
stepsize 10 and 2 assets	VaR violations	violations	n	kupiec_p	christoffel dq_stat	dq_p	acerbi_szekely	acerbi_sz	acerbi_szmae_es	empirical_es_loss
						diagnostic				
baseline	0.024193548			0.00817	0.58464		1.658681283	2.69355	0.8815	1.203286 -3.60159
model with kde when gpd fails	0.032258065	8	248	0.1138	0.46423	1.79542	0.4075	1.734374362	2.69355	0.8029 1.094424 -3.22123
baseline with same mean equation	0.024193548	6	248	0.00817	0.58464	3.51321	0.17263	1.667293687	2.69355	0.876 1.188234 -3.60159
model with no kde when gpd fails	0.024193548	6	248	0.00817	0.58464	3.63991	0.60233	1.546061134	2.69355	0.9666 1.236157 -3.49822
stepsize 10 and 4 assets	VaR violations	violations	n	kupiec_p	christoffel dq_stat	dq_p	acerbi_szekely	acerbi_sz	acerbi_szmae_es	empirical_es_loss
						diagnostic				
baseline	0.040322581			0.4385	0.35822		1.879874312	2.69355	0.7052	0.366633 -2.6984
model with kde when gpd fails	0.044354839	11	248	0.66589	0.31115	0.63604	0.72759	1.89400928	2.69355	0.5922 0.383826 -2.61003
baseline with same mean equation	0.040322581	10	248	0.4385	0.35822	0.83034	0.66023	1.883397874	2.69355	0.6889 0.393854 -2.6984
model with no kde when gpd fails	0.040322581	10	248	0.4385	0.35822	2.71804	0.74336	1.836206549	2.69355	0.6949 0.419182 -2.6984

Case	Model with kde being used when gpd fails(Mean absolute ES error)	Model with gjr garch distribution being used for tails when gpd fails (Mean absolute ES error)	Which model is better
Step size 1 and 2 assets	Failing but by very less value(0.00264) and also this model is passing all tests Baseline : 0.794419439 Model : 0.797055119	Failing by very less value 0.00389 . it is passing all tests . Baseline :0.794419439 Model : 0.798313726	Both models are not better here but by very small value .
Step size 1 and 4 assets	Failing by value 0.02. Baseline : 0.741873496 Model:0.767604324	failing by value 0.04 . Baseline :0.741873496 Model : 0.782259134	Both models are not better here but model with kde is better than other when difference is considered
Step size 10 and 2 assets	Passing by 0.09. Baseline: 1.18823393 Model : 1.094424169	Failing by 0.04 . and it also fails in kupeic test Baseline : 1.18823393 Model :1.236157193	Model with kde is better as it is better than baseline in mae ES and also passing all tests
Step size 10 and 4 assets	Passing by 0.01 Baseline : 0.393853581 Model : 0.383825908	Failing by 0.02 Baseline : 0.393853581 Model : 0.419181509	Model with kde for tails is better because it is better than baseline in mae ES and also passing all tests

By considering all cases we can conclude that model with kde when gpd fails is better and we can also consider that this model is better than baseline because in some cases baseline is not passing some tests

Results link :

https://drive.google.com/drive/folders/1uRowhOOUD8h5obSXPjFyW0-Lscb_60uQ?usp=drive_link

Link for code : [model with no kde](#)

[model with kde](#)