Supplementary Material

Table 1: Accuracy, computation time and size of the SVM learning surrogates

SVM Kernel/Settings	R_{Si}	R_{Idx}	R_{Recon}	Time in min	Compact Model Size (in MB)
SVM Linear	0.7087	0.9616	0.0086	0.571	2.961
SVM Poly2	0.7860	0.9664	0.0192	1.193	2.78
SVM Poly3	0.8555	0.9721	0.0130	5.693	2.786
SVM Poly4	0.8805	0.9794	0.0184	34.336	2.83
SVM RBF kernel	0.8850	0.9803	***0.0466***	2.986	2.751

Table 2: Accuracy, computation time and size of the decision tree learning surrogates

Control Parameter	R_{Si}	R_{Idx}	R_{Recon}	Training time (s)	Compact Model Size (MB)
N_{leaf} = 3	0.9618	0.9955	0.6554	4.69	8.826
N_{leaf} = 5	0.9540	0.9943	0.6132	4.24	8.08
N_{leaf} = 10	0.9346	0.9910	0.2823	3.17	4.104
N_{leaf} = 15	0.9220	0.9888	0.1530	2.99	2.824
N_{leaf} = 20	0.9129	0.9875	0.1747	2.93	2.167
N_{parent} = 3	0.9958	0.9997	***0.8411***	10.84	24.356
N_{parent} = 5	0.9910	0.9988	0.8308	7.38	17.593
N_{parent} = 10	0.9772	0.9966	0.6713	4.99	10.111
N_{parent} = 15	0.9673	0.9952	0.5273	4.08	7.194
N_{parent} = 20	0.9597	0.9933	0.2991	3.67	5.615

Table 3: Accuracy, computation time and size of the tree based ensemble learning surrogates

Ensemble Method	Kernel/Settings	R_{Si}	R_{Idx}	R_{Recon}	Time in min	Compact Model Size (in MB)
Tree with Boosting	N_{learn} = 100	0.9605	0.9948	0.5518	2.080	8.809
	N_{learn} = 250	0.9607	0.9949	0.5708	3.105	21.813
	N_{learn} = 500	0.9605	0.9949	0.5925	9.500	43.479
	N_{learn} = 750	0.9607	0.9949	0.5881	9.140	65.107
	N_{learn} = 1000	0.9607	0.9949	0.5862	17.883	86.74
Tree with Bagging	N_{learn} = 100	0.9367	0.9892	0.1520	0.824	*
	N_{learn} = 250	0.9436	0.9907	0.3154	1.884	*
	N_{learn} = 500	0.9485	0.9916	0.4927	3.765	*
	N_{learn} = 750	0.9512	0.9921	0.5732	5.423	*
	N_{learn} = 1000	0.9531	0.9925	***0.5979***	7.524	*

**Size of bagging models are >1 GB and hence omitted*

Table 4: Accuracy, computation time and size of the GLM, shrinkage and shrinkage based GLM learning surrogates

Polynomial Kernel Order	Method	R_{Si}	R_{Idx}	R_{Recon}	Time in min	Compact Model Size (in KB)
	Ridge ($\alpha = 10^{-6}$)	0.5498	0.8738	0.0011	0.572	14
	Elastic Net ($\alpha = 0.5$)	0.7387	0.9663	0.0713	14.300	12
	Lasso ($\alpha = 1$)	0.7372	0.9661	0.0490	16.094	12
	GLM	0.7913	0.9695	0.0949	0.085	22562
	Ridge GLM ($\alpha = 10^{-6}$)	0.5499	0.8738	0.0011	1.403	13
	Elastic Net GLM ($\alpha = 0.5$)	0.7388	0.9663	0.0712	26.819	12
2	Lasso GLM ($\alpha = 1$)	0.7373	0.9662	0.0495	29.019	12
	Ridge ($\alpha = 10^{-6}$)	0.5598	0.8767	0.0012	0.971	20
	Elastic Net ($\alpha = 0.5$)	0.8171	0.9709	0.0888	29.311	17
	Lasso ($\alpha = 1$)	0.8205	0.9707	0.0829	33.647	17
	GLM	0.8752	0.9751	***0.3431***	0.128	50024
	Ridge GLM ($\alpha = 10^{-6}$)	0.5598	0.8767	0.0012	1.717	19
	Elastic Net GLM ($\alpha = 0.5$)	0.8175	0.9709	0.0917	46.848	16
3	Lasso GLM ($\alpha = 1$)	0.8205	0.9707	0.0978	49.217	16
	Ridge ($\alpha = 10^{-6}$)	0.5646	0.8770	0.0014	1.180	24
	Elastic Net ($\alpha = 0.5$)	0.8392	0.9731	0.0763	40.270	20
	Lasso ($\alpha = 1$)	0.8245	0.9729	0.1159	49.533	20
	GLM	0.8765	0.9820	0.2367	0.267	100846
	Ridge GLM ($\alpha = 10^{-6}$)	0.5646	0.8770	0.0014	2.018	23
	Elastic Net GLM ($\alpha = 0.5$)	0.8442	0.9731	0.1112	59.327	20
4	Lasso GLM ($\alpha = 1$)	0.8293	0.9728	0.1166	63.072	20

Table 5: Accuracy, computation time and size of the neural network learning surrogates

Network Architecture	Layer	No. of Nodes	R_{Si}	R_{Idx}	R_{Recon}	Time in min	Model Size (in MB)
Feedforward Network	1	10	0.8353	0.9718	0.0082	0.2730	0.719
		25	0.8633	0.9739	0.0229	0.3091	0.997
		50	0.8392	0.9749	0.0501	0.5866	1.462
		75	0.8495	0.9748	0.0713	0.9539	1.890
		100	0.8245	0.9746	0.1136	1.3134	2.368
	2	10	0.7068	0.9697	0.0171	2.9511	1.117
		25	0.8685	0.9768	0.0813	0.3547	3.026
		50	0.8775	0.9767	0.0713	0.5658	9.180
		75	0.8380	0.9750	0.0088	0.8588	19.025
		100	0.8444	0.9738	0.0457	1.3307	32.671
	3	10	0.8303	0.9720	0.0512	0.2510	1.484
		25	0.8852	0.9762	0.1517	0.4233	4.997
		50	0.8462	0.9762	0.0589	0.7082	16.813
		75	0.8619	0.9760	0.0484	1.1689	36.063
		100	0.8436	0.9756	0.0482	1.7047	62.858
Cascaded Forward Network	1	10	0.8408	0.9720	0.0133	0.1915	0.748
		25	0.8396	0.9750	0.0437	0.3372	1.025
		50	0.8622	0.9755	0.1007	0.6064	1.491
		75	0.8373	0.9760	0.1163	1.0095	1.930
		100	0.8440	0.9761	0.1638	1.5087	2.396
	2	10	0.8690	0.9759	0.0577	0.2692	1.326
		25	0.8831	0.9769	0.1318	0.4338	3.461
		50	0.8302	0.9773	0.0071	0.8213	9.990
		75	0.8250	0.9751	0.0761	1.2389	20.214
		100	0.7868	0.9746	0.0162	1.5872	34.235
	3	10	0.8647	0.9751	***0.2001***	0.3083	2.194
		25	0.8609	0.9759	0.0729	0.6028	7.745
		50	0.8352	0.9755	0.0515	1.1315	25.963
		75	0.7264	0.9727	0.0042	2.0431	55.309
		100	0.7068	0.9697	0.0171	2.9511	95.959

Table 6: Maximum likelihood estimate for a single microseismic event location at (31, 25, 158) using various receiver placement geometries and noise levels, using bulk likelihoods calculated on voxel by voxel basis

Receiver Placement Geometry	σ	**SNR**	$\widehat{x}_o$	$\widehat{y}_o$	$\widehat{z}_o$	$\|\mathbf{e}\|$
at central location	100	12.8	31	25	158	0
	250	4.3	31	25	158	0
along principal diagonal	100	11.7	31	25	158	0
	250	3.8	31	25	158	0
along anti-diagonal	100	12.5	31	25	158	0
	250	4.7	31	25	158	0
Upper Triangular Region	100	11.2	31	25	158	0
	250	3.3	31	25	158	0
Lower Triangular Region	100	12.4	31	25	158	0
	250	4.4	31	25	158	0
All 23 Uniformly Distributed	100	11.8	31	25	158	0
	250	3.9	31	25	158	0

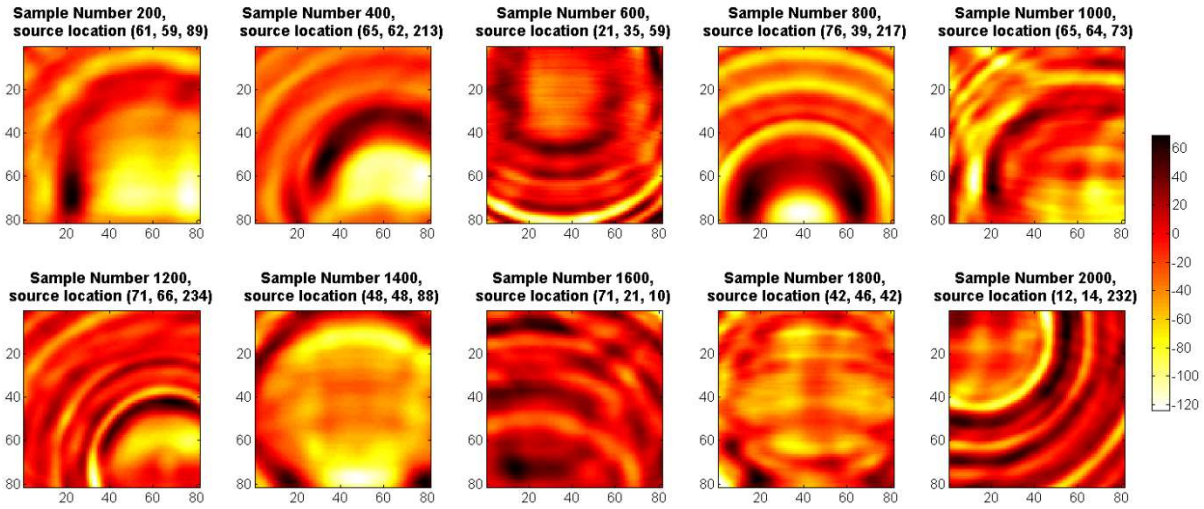


Figure 1: Map view of the acoustic pressure for different locations of explosive microseismic sources for fixed velocity model. Recordings are at T = 1.4 s from the origin time. Sample numbers and source locations are given in the subplot titles.

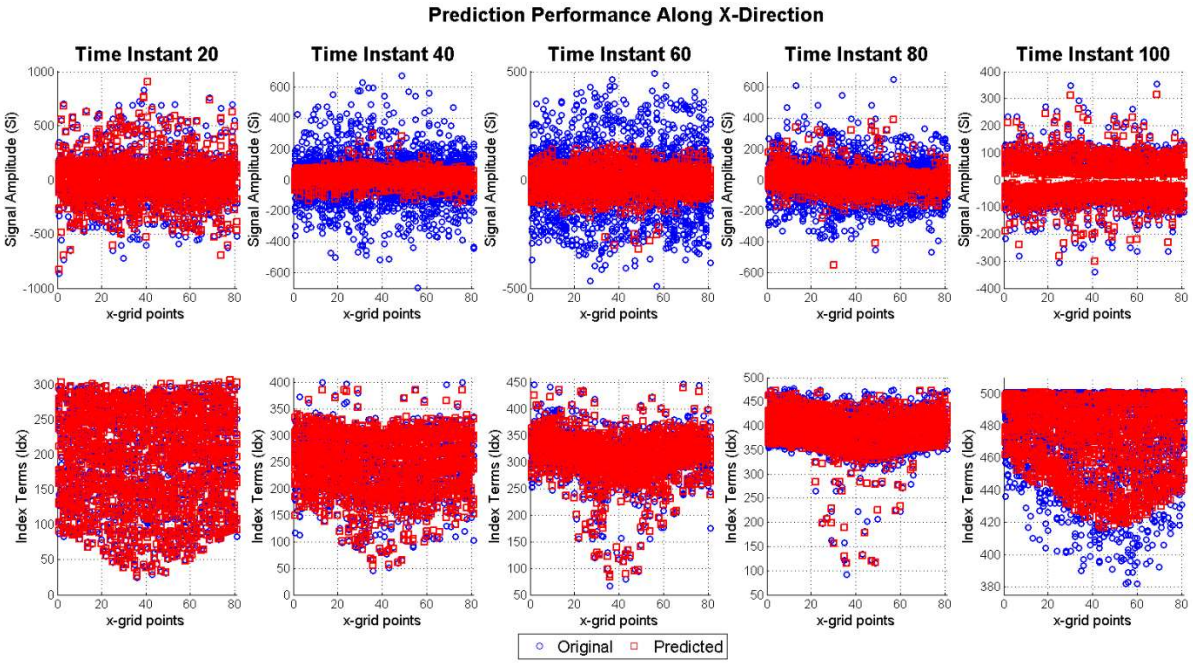


Figure 2: Ground truth vs. predicted variation in Si and Idx along x-direction using Gaussian process surrogates.

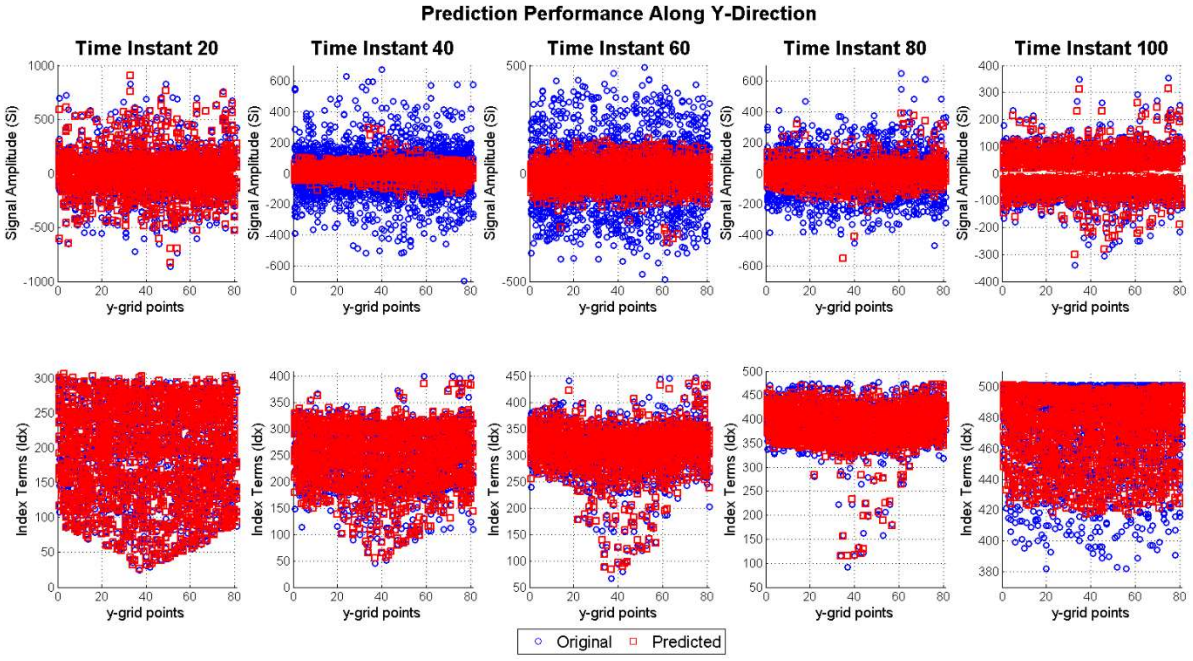
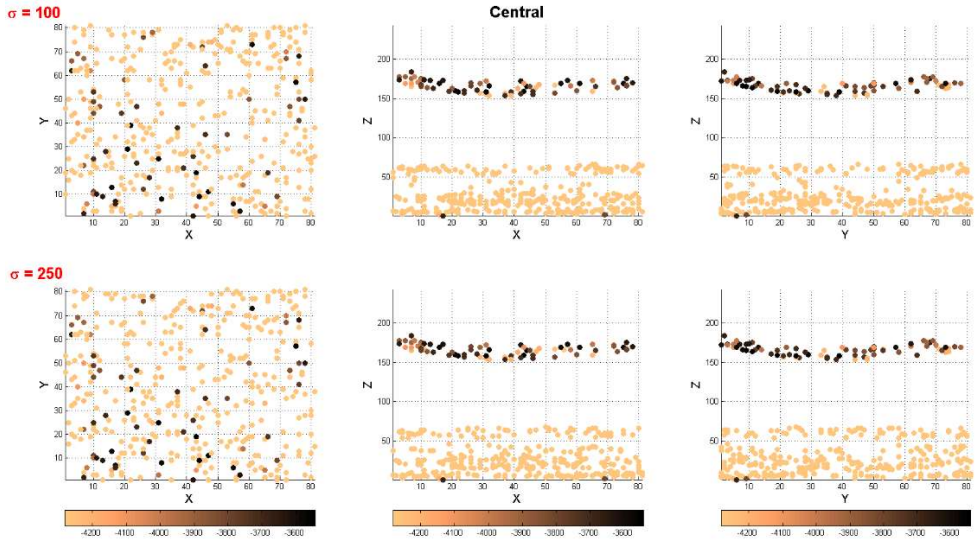
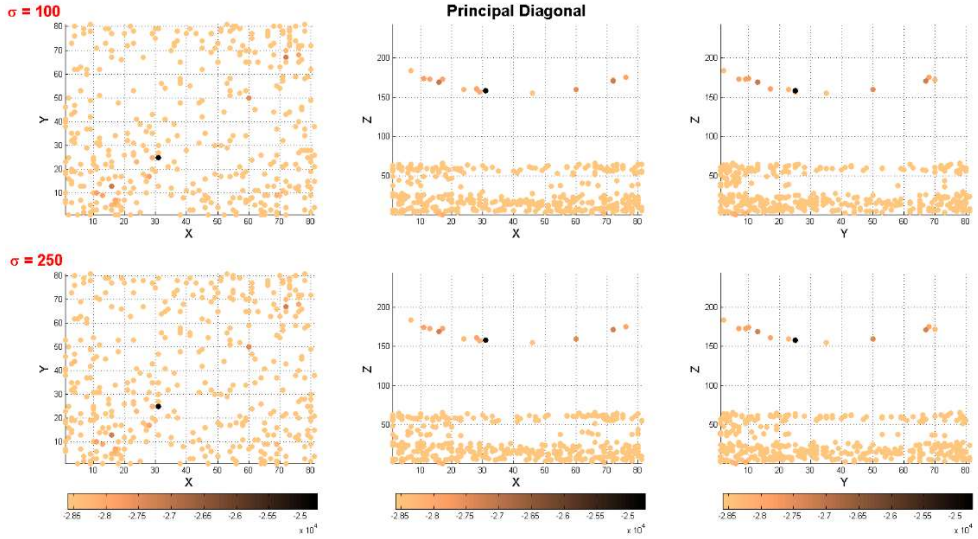


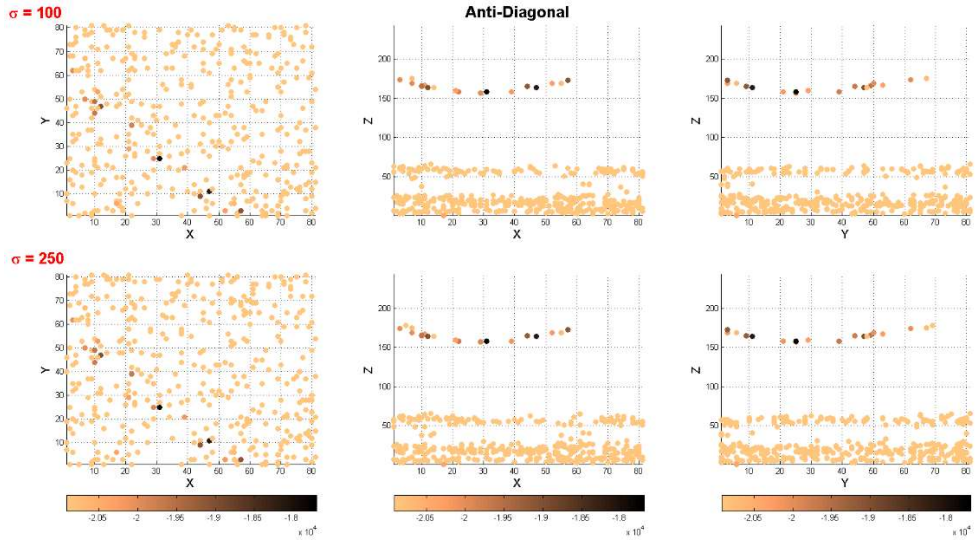
Figure 3: Ground truth vs. predicted variation in Si and Idx along y-direction using Gaussian process surrogates.


σ = 100
Central
σ = 250
X
Y
Z

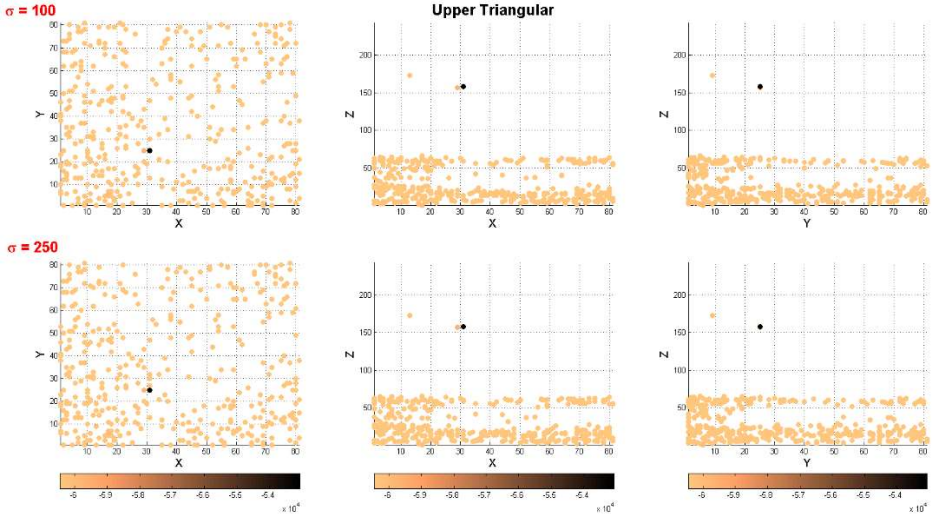
(a)


σ = 100
Principal Diagonal
σ = 250
X
Y
Z

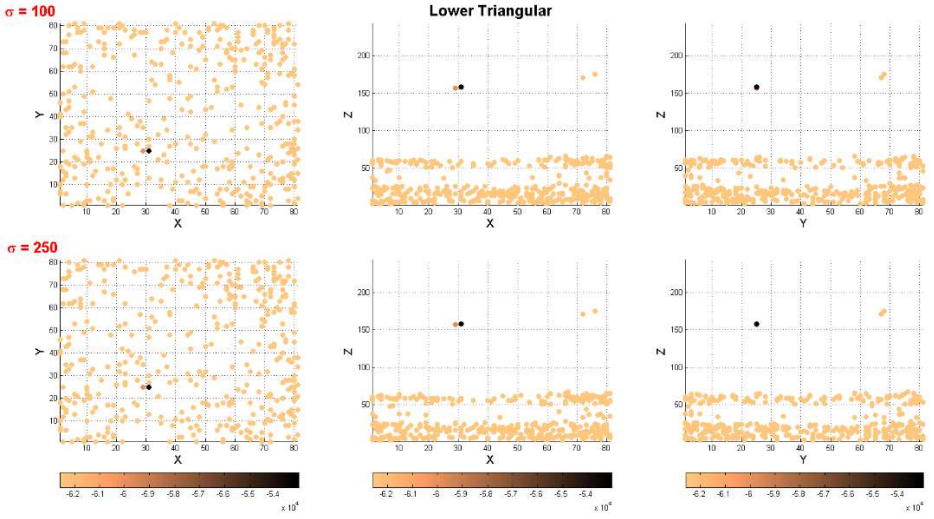
(b)


σ = 100
Anti-Diagonal
σ = 250
X
Y
Z

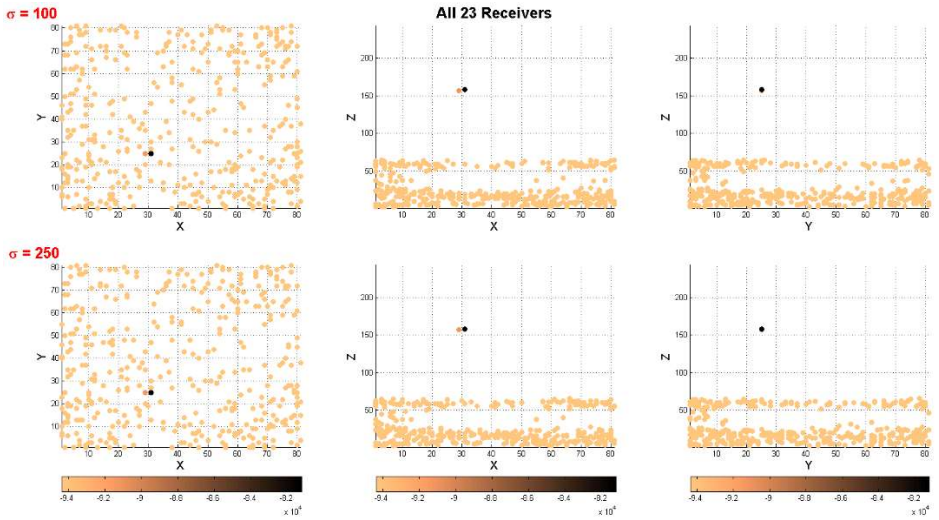
(c)



(d)

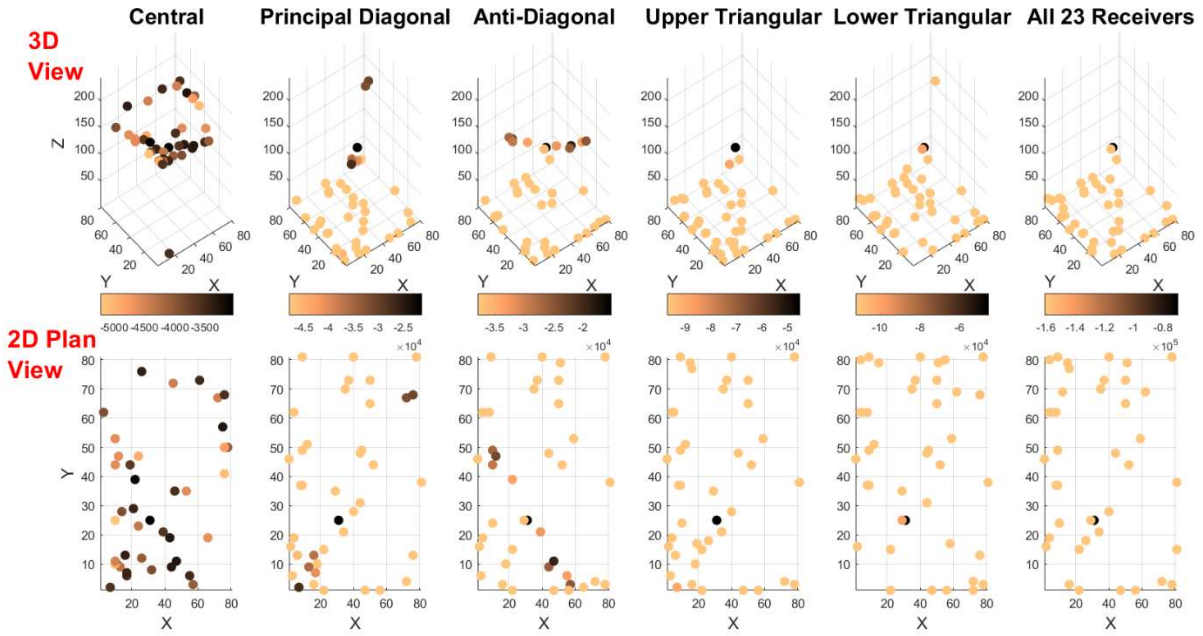


(e)

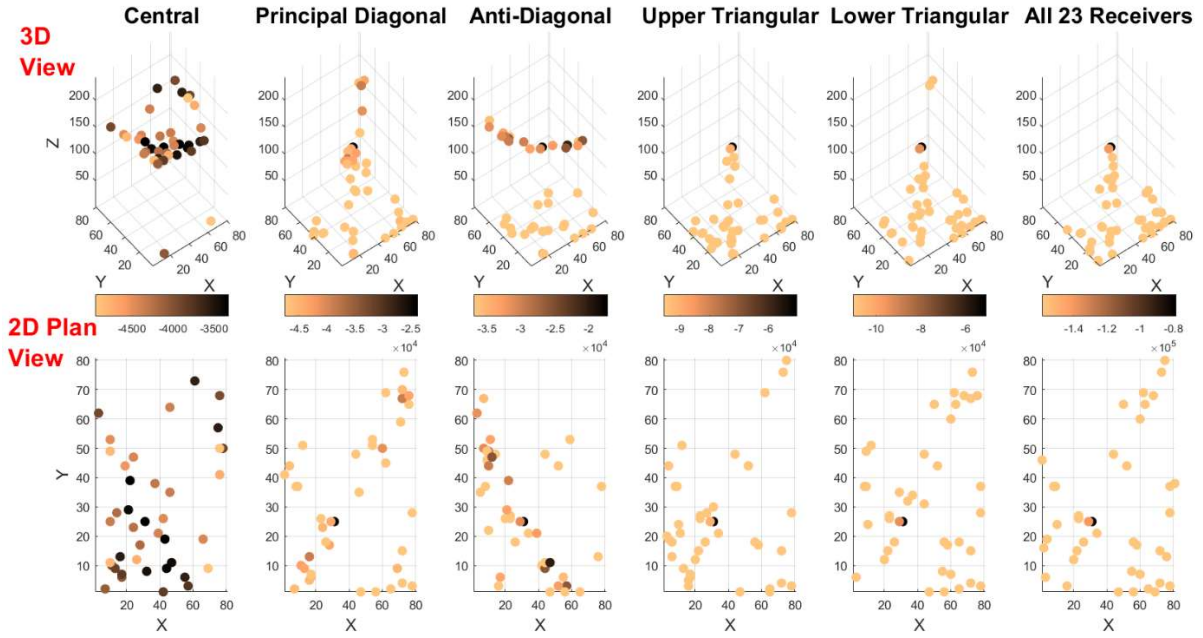


(f)

Figure 4: Scatter-diagram of top 10 percentile of likelihood values for two different noise levels and receiver combinations (a) central, (b) principal diagonal, (c) anti-diagonal, (d) upper-triangular, (e) lower-triangular, (f) all 23 receivers.



(a)



(b)

Figure 5: Most likely event locations using various receiver arrangement with added noise variance σ = 100. Log-likelihood values are shown in the color-bars and the size of the data-points are proportional to the likelihood values: (a) True likelihood, (b) Proxy-based likelihood.