Evaluating uncertainty and predictive performance of probabilistic models devised for grade estimation in a porphyry copper deposit

SUPPLEMENTARY MATERIAL

Raymond Leung, Alexander Lowe, Arman Melkumyan

Rio Tinto Centre, Faculty of Engineering
The University of Sydney

List of Figures

List of Tables

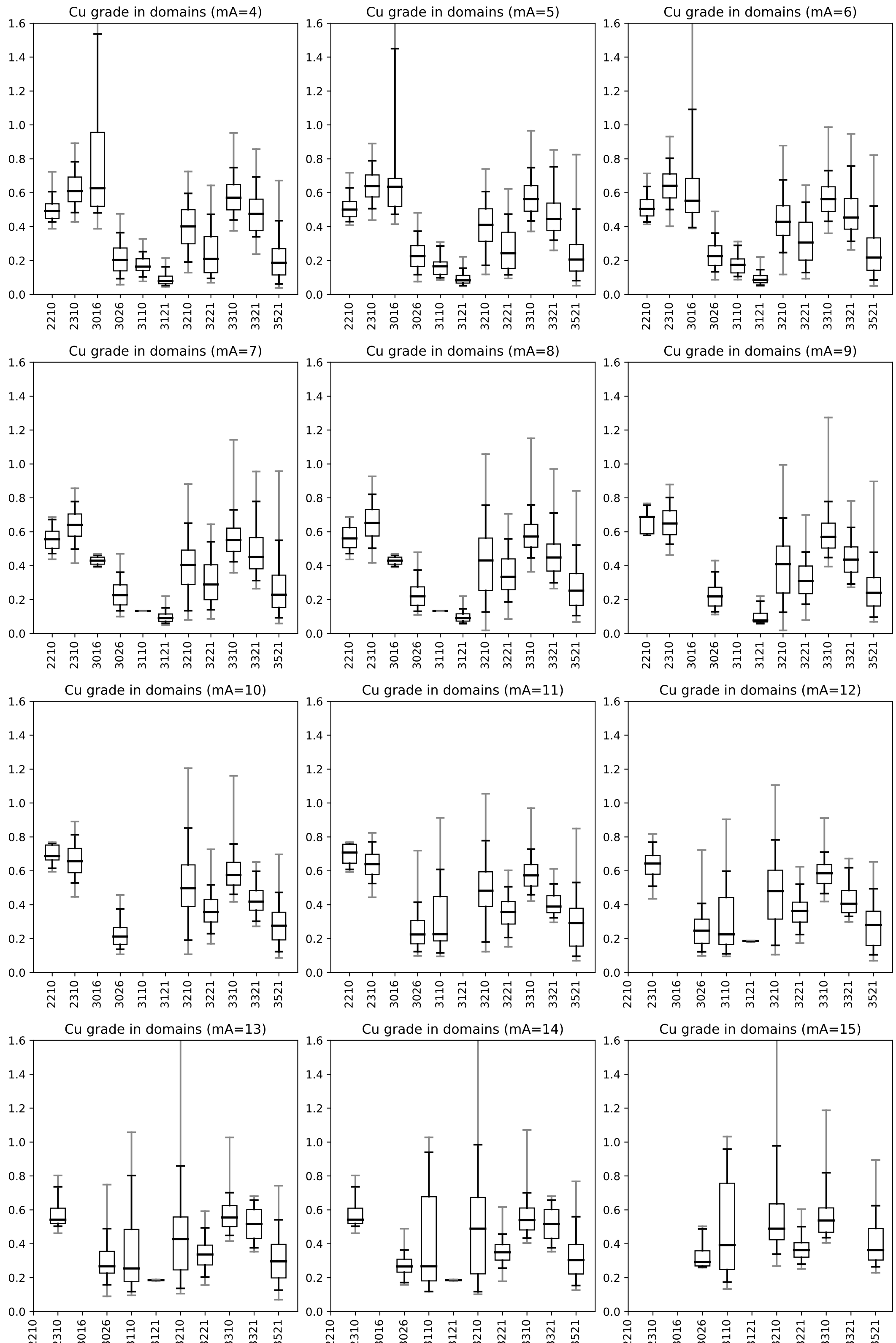


Figure S.1: Copper grade box-plots across different domains (g_D) and inference months (m_A) for locations that require future bench prediction (extrapolation) at the Kennecott porphyry deposit.

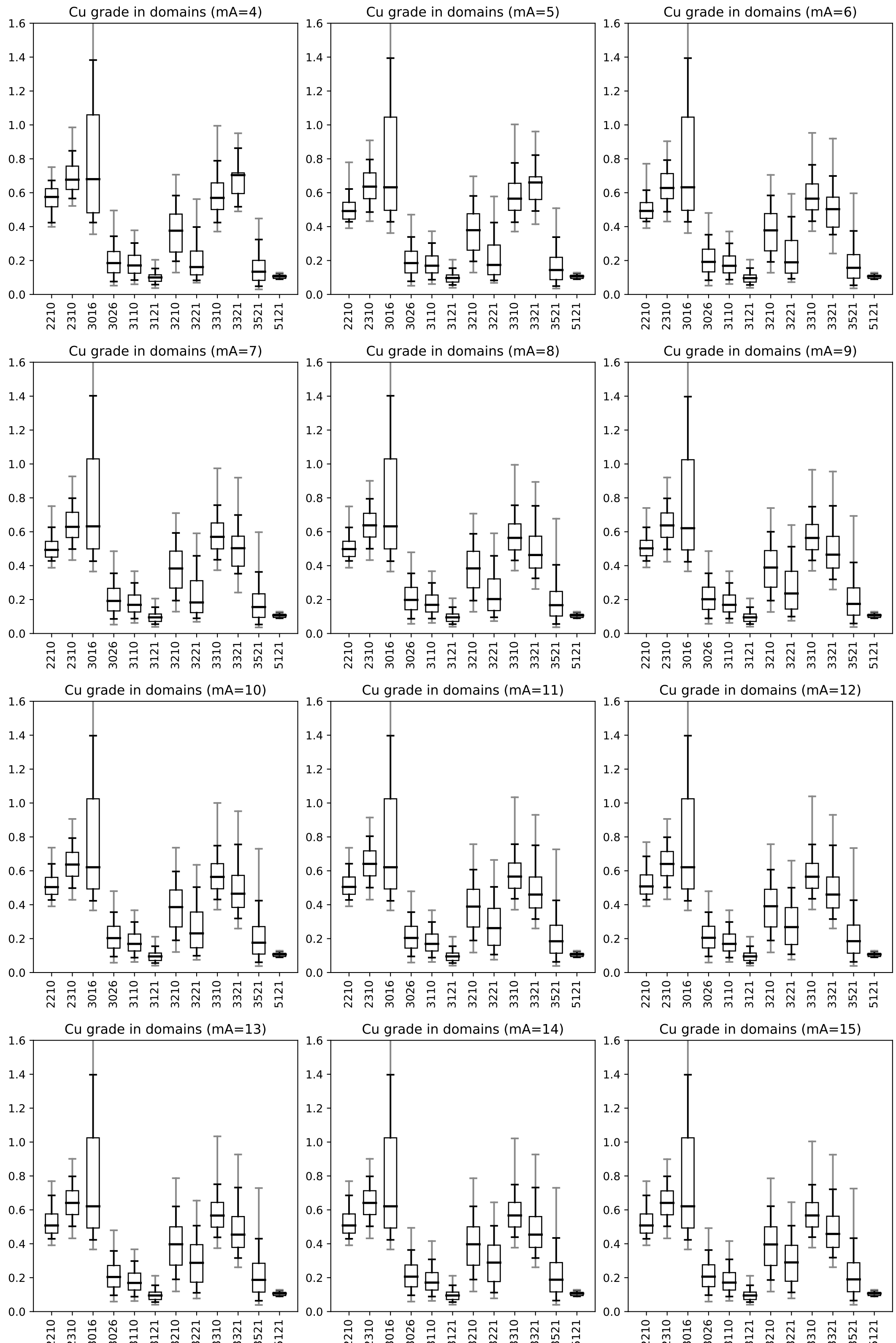


Figure S.2: Copper grade box-plots across different domains (g_D) and inference months (m_A) for locations that require in-situ regression (interpolation) at the Kennecott porphyry deposit.

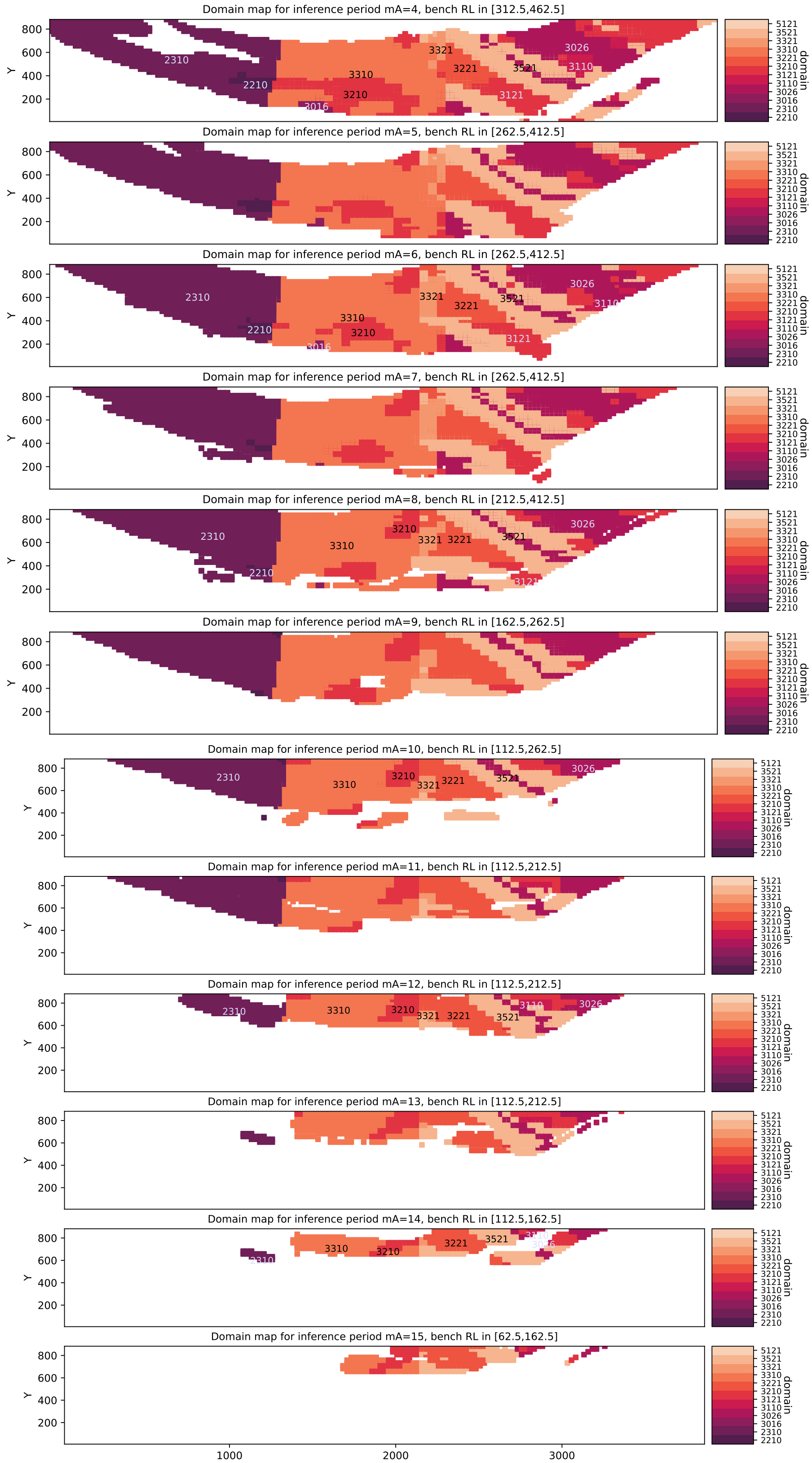


Figure S.3: Domain layout across different inference periods. Blocks pertain to locations that require future bench prediction (extrapolation).

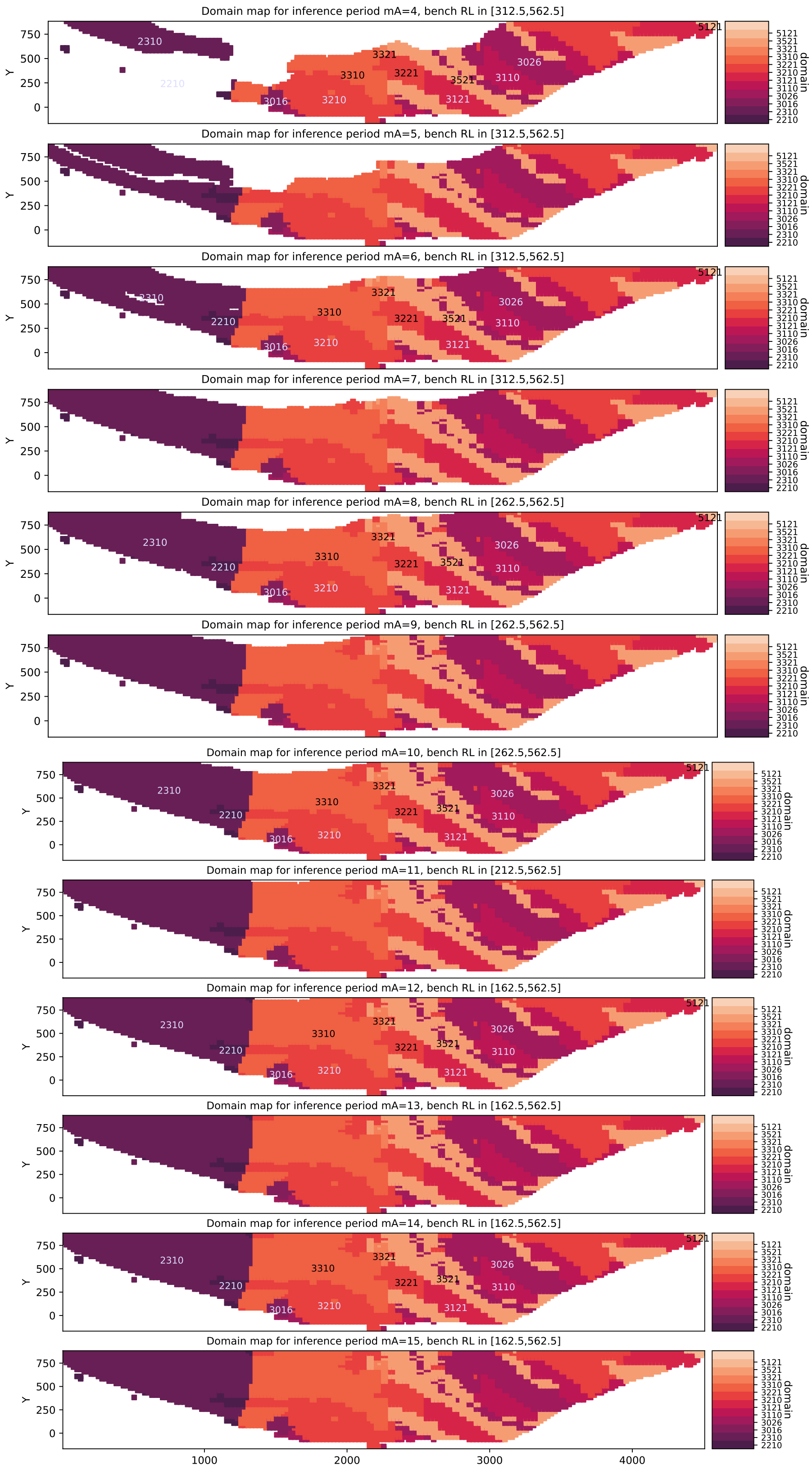


Figure S.4: Top-down view of domain layout across different inference periods. Blocks pertain to locations that require in-situ regression (interpolation).

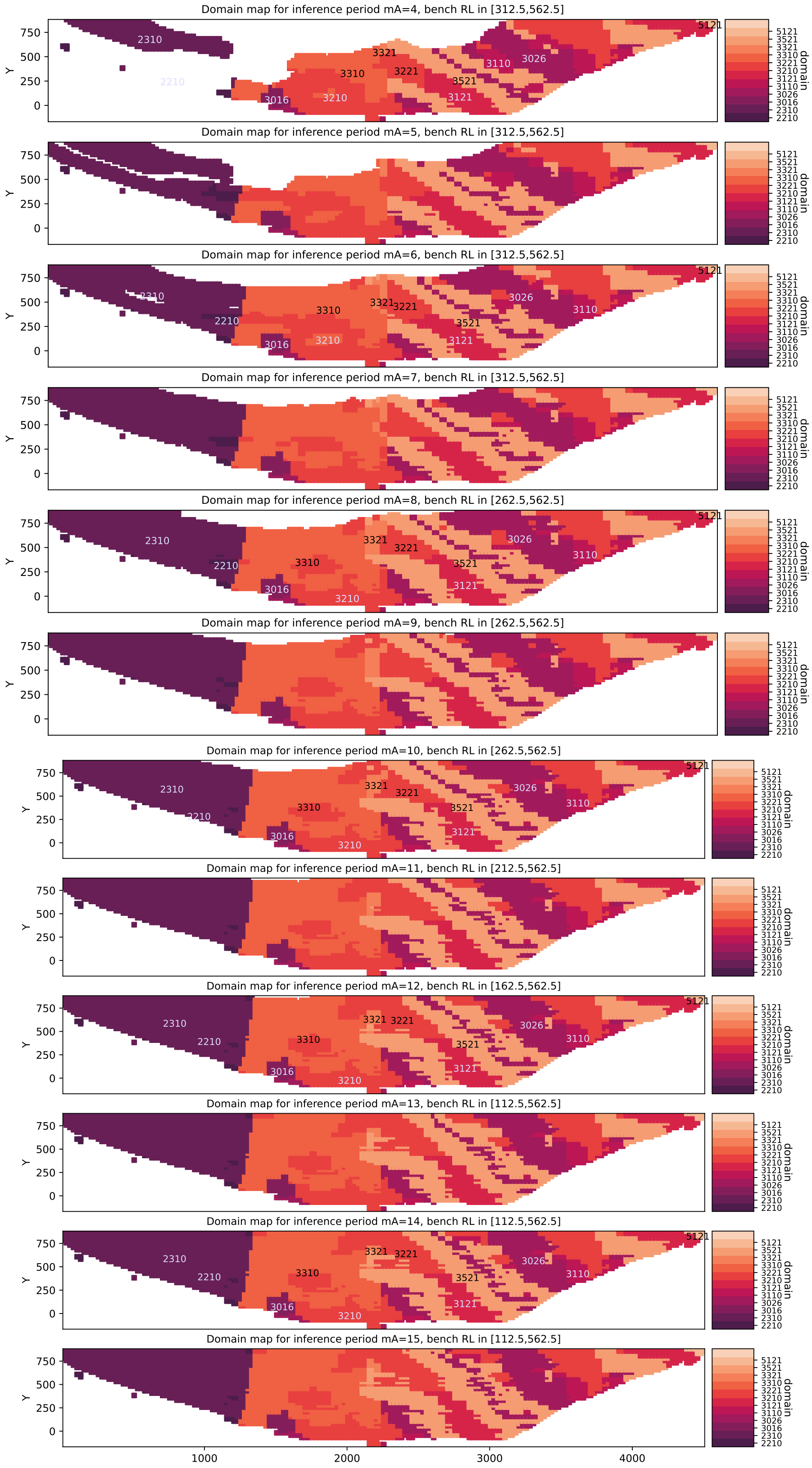
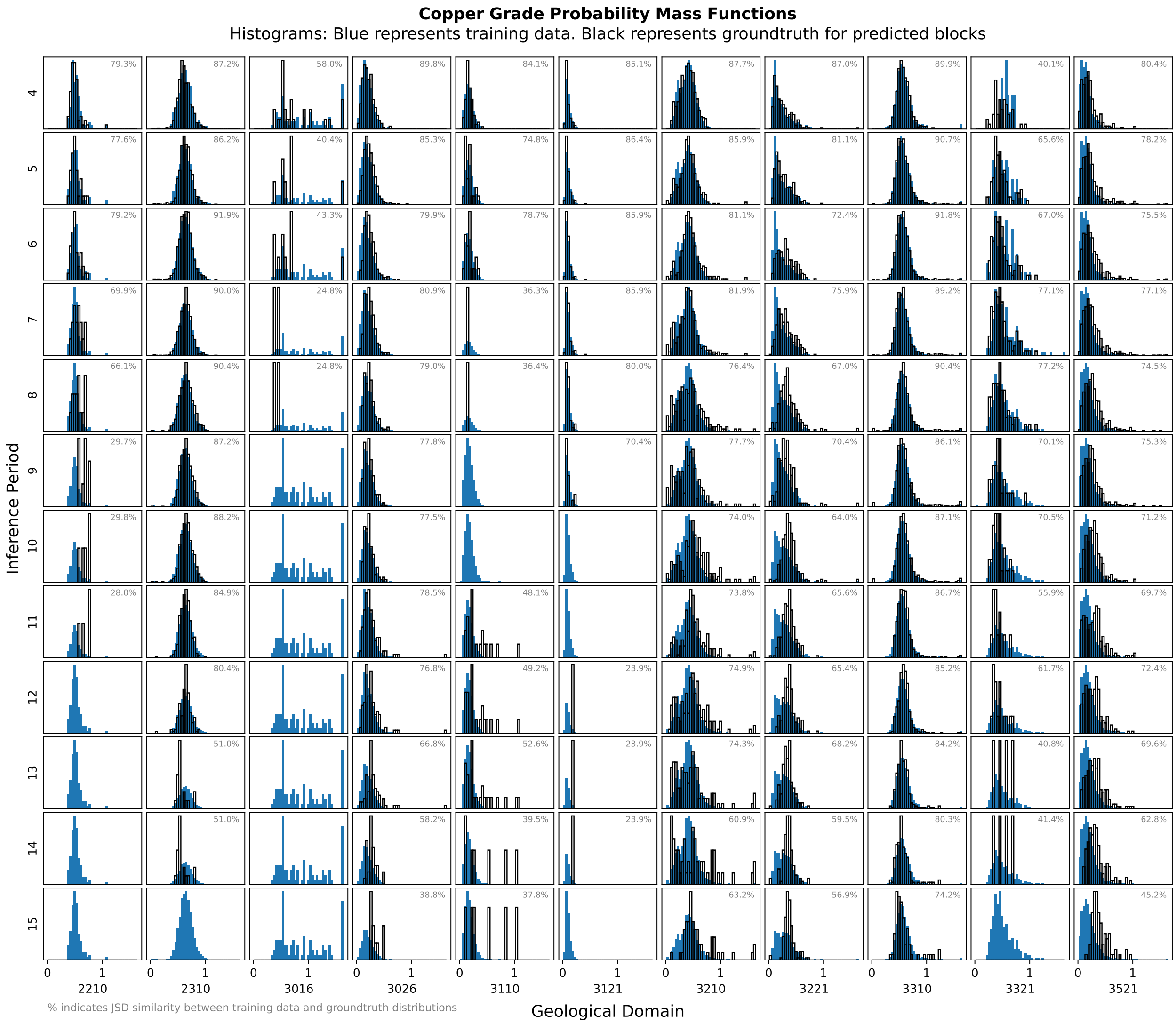
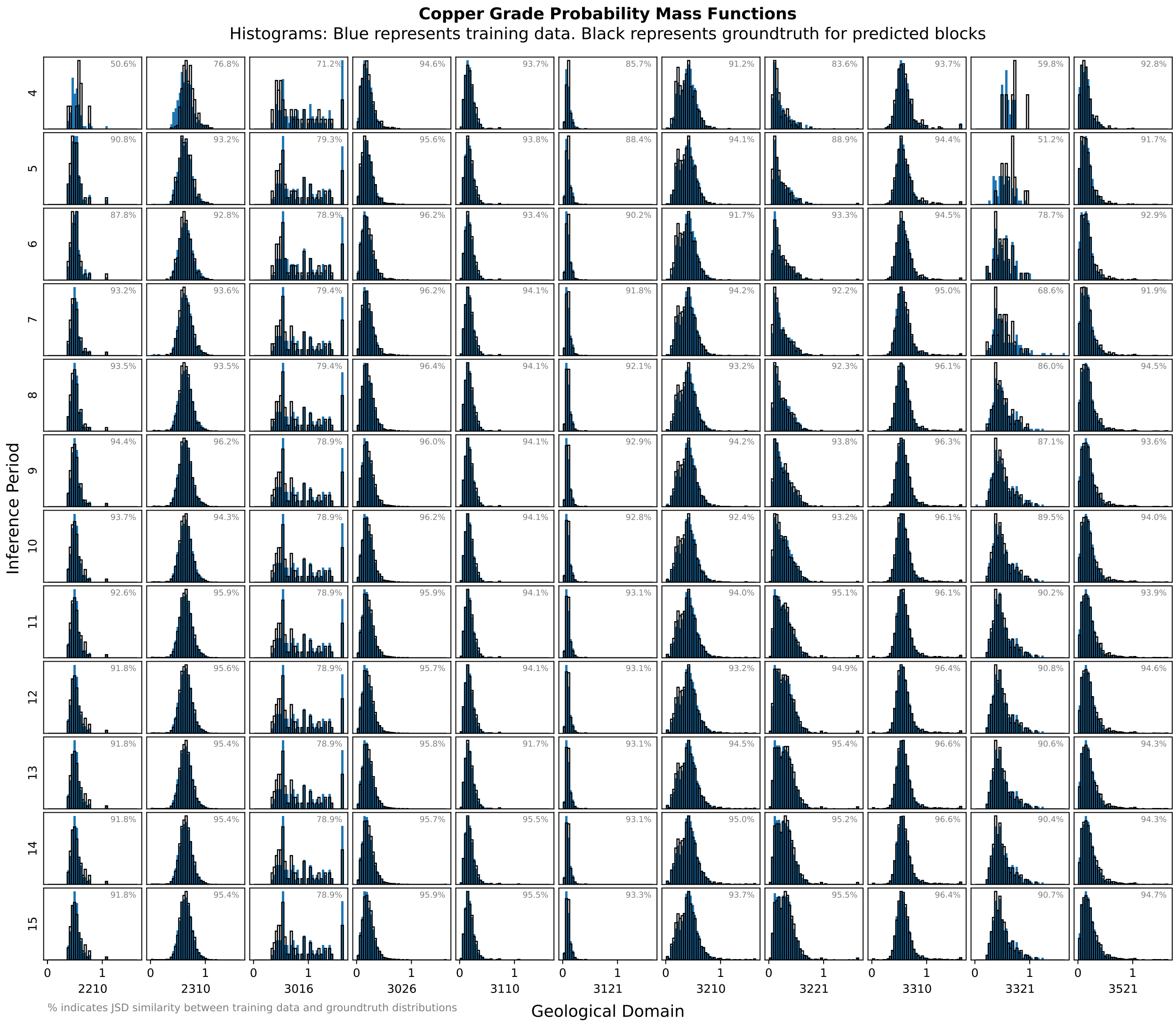


Figure S.5: Bottom-up view of domain layout across different inference periods. Blocks pertain to locations that require in-situ regression (interpolation).



The graphs compare training data (blue-filled) with groundtruth (black-hollow) at the predicted locations.
Numerical value within each subplot indicates JSD similarity between the training data and groundtruth distributions.

Figure S.6: Copper grade distributions across different domains and inference periods for locations that require future bench prediction (extrapolation).



The graphs compare training data (blue-filled) with groundtruth (black-hollow) at the predicted locations.
Numerical value within each subplot indicates JSD similarity between the training data and groundtruth distributions.

Figure S.7: Copper grade distributions across different domains and inference periods for locations that require in-situ regression (interpolation).

Table S.1: Sample size statistics for in-situ regression.

g_D	2210		2310		3016		3026		3110		3121		3210		3221		3310		3321		3521	
	n_T	n_I	n_T	n_I	n_T	n_I	n_T	n_I	n_T	n_I	n_T	n_I	n_T	n_I	n_T	n_I	n_T	n_I	n_T	n_I	n_T	n_I
m_A=4	76	9	641	350	67	43	1078	720	630	473	428	402	1615	1080	300	148	772	420	14	6	1352	918
m_A=5	66	47	837	656	81	63	1307	985	634	540	469	478	1659	1436	418	263	1312	622	61	18	1750	1289
m_A=6	114	50	1249	864	86	63	1367	1215	660	546	518	506	1895	1545	468	368	2000	1032	59	55	1874	1570
m_A=7	109	75	1559	1213	86	71	1653	1275	660	564	637	567	2025	1716	631	383	2351	1505	143	55	2319	1687
m_A=8	130	77	1844	1548	86	71	1794	1537	661	564	666	640	2153	1858	699	513	2894	1775	195	119	2599	2059
m_A=9	121	87	2150	1834	86	73	1984	1644	661	565	685	661	2216	1956	804	607	3069	2176	208	159	2718	2305
m_A=10	123	91	2345	2134	86	73	1990	1814	661	565	696	680	2353	2029	926	669	3444	2296	224	163	2959	2467
m_A=11	123	92	2531	2310	86	73	2014	1858	661	565	692	680	2346	2131	965	810	3641	2630	227	187	2969	2645
m_A=12	123	96	2654	2606	86	73	2050	1886	661	565	692	680	2421	2174	1057	864	3912	2791	242	187	3062	2678
m_A=13	123	96	2653	2746	86	73	2088	1932	690	565	692	680	2436	2238	1101	946	3965	3056	238	199	3143	2762
m_A=14	123	96	2654	2746	86	73	2113	1972	682	584	692	680	2469	2303	1123	1004	4073	3149	242	199	3220	2840
m_A=14	123	96	2654	2765	86	73	2120	1997	682	585	692	682	2506	2328	1152	1024	4150	3221	244	203	3277	2900

Table S.2: Histogram distance (H) summary statistics for in-situ regression.

Model family	Abbrev	h_{JS} mean (SE)	h_{EM} mean (SE)
Simple kriging	SK / SK-SGS	0.2575 (0.0062)	0.1186 (0.0033)
Ordinary kriging	OK / OK-SGS	0.1386 (0.0045)	0.0790 (0.0031)
Gaussian process (global mean)	GP(G) / GP-CRF	0.0917 (0.0040)	0.0409 (0.0016)
Gaussian process (local mean)	GP(L) / GP-SGS	0.0867 (0.0040)	0.0382 (0.0016)

Table S.3: Variogram ratio (R) and spatial fidelity (F) summary statistics for in-situ regression.

Model family	Abbrev	R mean (SE)	F mean (SE)
Simple kriging	SK / SK-SGS	0.4028 (0.0102)	0.5149 (0.0108)
Ordinary kriging	OK / OK-SGS	0.5803 (0.0071)	0.7175 (0.0075)
Gaussian process (global mean)	GP(G) / GP-CRF	0.7247 (0.0035)	0.8482 (0.0021)
Gaussian process (local mean)	GP(L) / GP-SGS	0.7535 (0.0032)	0.8656 (0.0018)

Table S.4: Accuracy (A) and precision (P) summary statistics for in-situ regression.

Model family	Abbrev	A mean (SE)	P mean (SE)
Simple kriging	SK / SK-SGS	0.6142 (0.0148)	0.8840† (0.0046)
Ordinary kriging	OK / OK-SGS	0.9051 (0.0081)	0.8310 (0.0040)
Gaussian process (global mean)	GP(G) / GP-CRF	0.9666 (0.0042)	0.8303 (0.0034)
Gaussian process (local mean)	GP(L) / GP-SGS	0.9723 (0.0038)	0.8109 (0.0037)

* Group averages exclude SGS/CRF $s=2$ and $s=4$, viz., epochs long before convergence.

† Conditional on having an accurate model which is far lower for SK at 0.6142

Table S.5: Likelihood (L) and goodness (G) summary statistics for in-situ regression.

Model family	Abbrev	L median	$[q_L, q_U]$	G mean (SE)
Simple kriging	SK / SK-SGS	0.4731	[0.2364, 0.6897]	0.7876 (0.0052)
Ordinary kriging	OK / OK-SGS	0.5708	[0.3269, 0.7539]	0.8407 (0.0041)
Gaussian process (global mean)	GP(G) / GP-CRF	0.5853	[0.3351, 0.7704]	0.8546 (0.0033)
Gaussian process (local mean)	GP(L) / GP-SGS	0.5954	[0.3517, 0.7802]	0.8487 (0.0032)

Table S.6: Interval tightness (I) summary statistics for in-situ regression.

Model family	Abbrev	I mean (SE)
Simple kriging	SK / SK-SGS	0.6964 (0.0074)
Ordinary kriging	OK / OK-SGS	0.5924 (0.0066)
Gaussian process (global mean)	GP(G) / GP-CRF	0.4402 (0.0028)
Gaussian process (local mean)	GP(L) / GP-SGS	0.4445 (0.0031)

Table S.7: Significance testing of statistical scores for in-situ regression over all domains and inference periods.

	Histogram $H = h_{\text{EM}}$		Spatial Fidelity F		Accuracy A		Precision P	
Family ψ	p	CI	p	CI	p	CI	p	CI
SK/SGS	$< .001$	$[0.1595, 0.1820]$	$< .001$	$[-0.3722, -0.3292]$	$< .001$	$[-0.3867, -0.3295]$	> 0.99	$[0.0705, 0.0870]$
OK/SGS	$< .001$	$[0.0470, 0.0566]$	$< .001$	$[-0.1628, -0.1333]$	$< .001$	$[-0.0801, -0.0542]$	> 0.99	$[0.0167, 0.0272]$
GP(G)/CRF	$< .001$	$[0.0033, 0.0065]$	$< .001$	$[-0.0189, -0.0159]$	0.0012	$[-0.0091, -0.0022]$	> 0.99	$[0.0166, 0.0198]$
Reference	μ	SE	μ	SE	μ	SE	μ	SE
GP(L)/SGS	0.0382	(0.0016)	0.8656	(0.0018)	0.9723	(0.0038)	0.8109	(0.0037)

	Interval I		Goodness G		Likelihood L	
Family ψ	p	CI	p	CI	$p^{\dagger}$	CI†
SK/SGS	$< .001$	$[0.2408, 0.2675]$	$< .001$	$[-0.0692, -0.0528]$	$< .001$	$[-0.1004, -0.0783]$
OK/SGS	$< .001$	$[0.1379, 0.1622]$	$< .001$	$[-0.0123, -0.0035]$	$< .001$	$[-0.0305, -0.0102]$
GP(G)/CRF	> 0.99	$[-0.0073, -0.0013]$	> 0.99	$[0.0043, 0.0074]$	0.0137	$[-0.0202, -0.0011]$
Reference	μ	SE	μ	SE	median	$[q_L, q_U]$
GP(L)/SGS	0.4445	(0.0031)	0.8487	(0.0032)	0.5954	$[0.3517, 0.7802]$

† The more conservative Welch's t-test is used assuming unequal population variance.

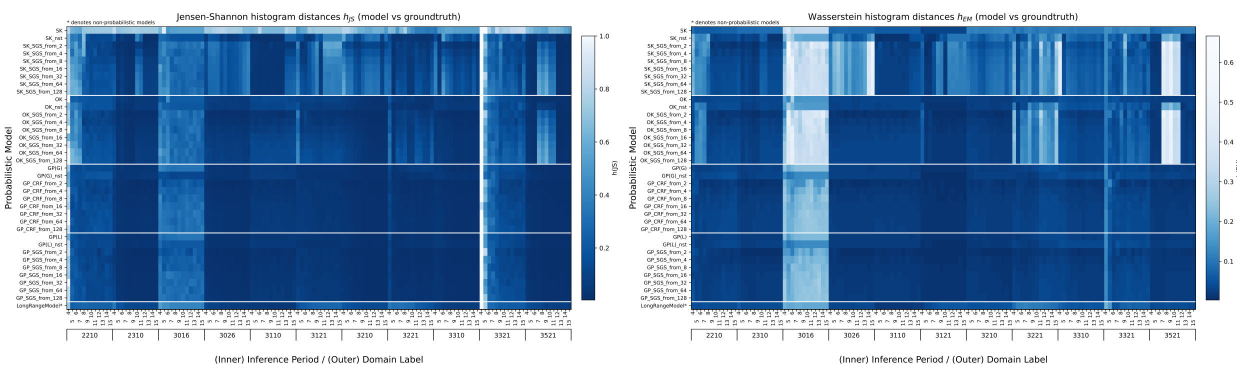


Figure S.8: View of (left) Jensen-Shannon, (right) EM histogram distances for in-situ regression.

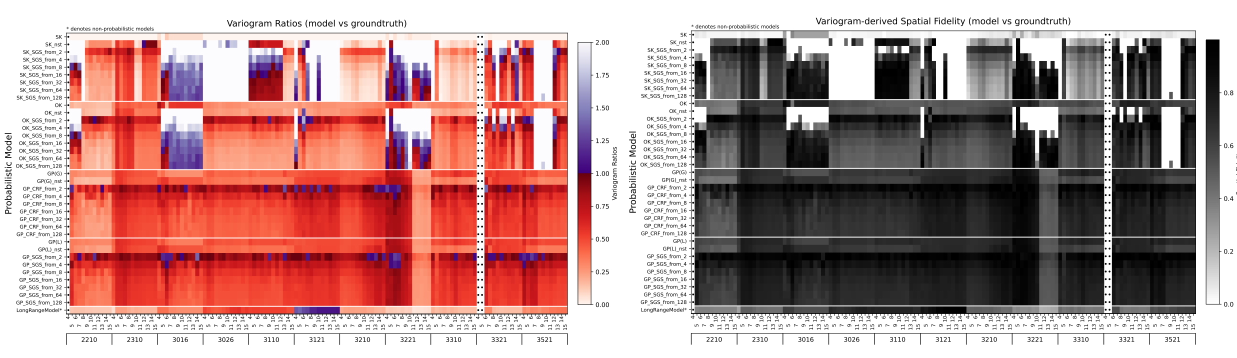


Figure S.9: View of (left) variogram ratios R and (right) spatial fidelity F for in-situ regression.

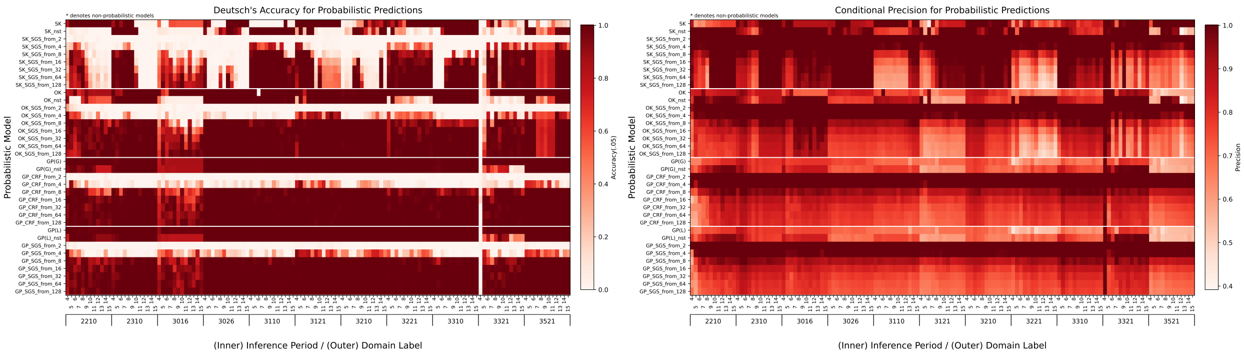


Figure S.10: View of (left) accuracy A and (right) precision P for in-situ regression.

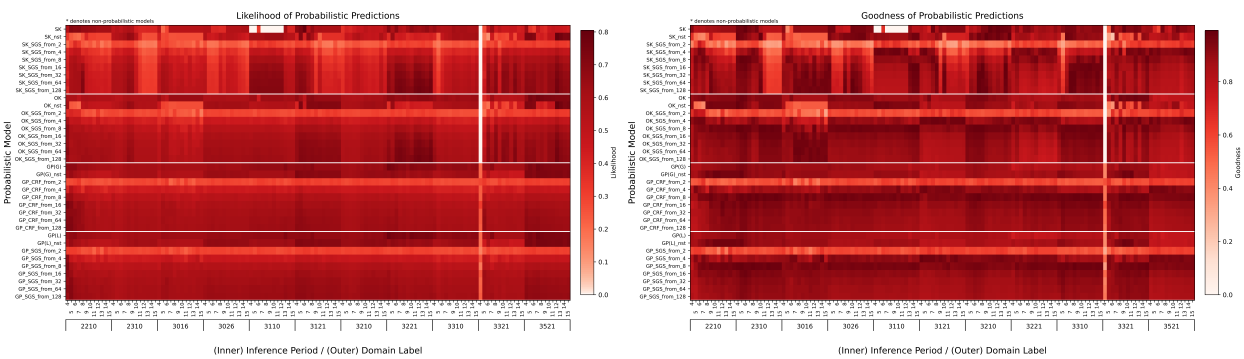


Figure S.11: View of (left) likelihood L and (right) goodness G for in-situ regression.

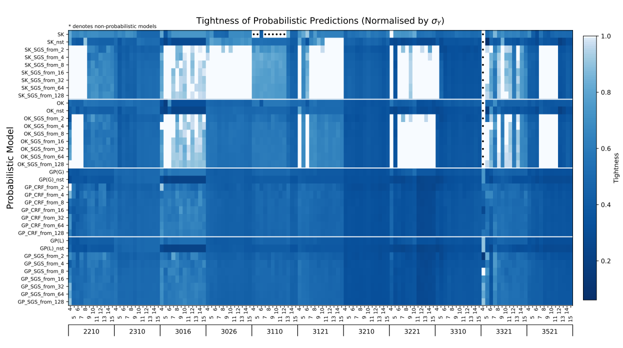


Figure S.12: View of interval tightness I for in-situ regression across domains and inference periods.

Table S.8: Sample-weighted performance statistics for future-bench prediction across domains and inference periods

	Histogram distances		Spatial Fidelity	Abs. Synchronicity		Likelihood	Accuracy	Precision	Goodness	Interval tightness
Model	h_{JS}	h_{EM}	F	$\|S\|_{.25}$	$\|S\|_{.75}$	L	$A_{\xi=.05}$	P	G	I
SK_nst	0.1976	0.0856	0.4060	0.2166	0.7262	0.4738	0.5667	0.9278	0.8391	0.4990
SK_SGS_from_2	0.2159	0.1678	0.6429	0.0013	0.4356	0.2487	0.0044	0.9999	0.4960	1.2714
SK_SGS_from_4	0.2399	0.1821	0.5316	0.1102	0.6219	0.3817	0.1780	0.9948	0.7551	1.3161
SK_SGS_from_8	0.2626	0.1738	0.4642	0.2168	0.6917	0.4583	0.4091	0.9477	0.8378	1.3006
SK_SGS_from_16	0.2961	0.1801	0.4266	0.2640	0.7160	0.4898	0.5505	0.9122	0.8479	1.3319
SK_SGS_from_32	0.3208	0.1822	0.3942	0.2936	0.7253	0.5086	0.5938	0.8863	0.8466	1.3301
SK_SGS_from_64	0.3402	0.1859	0.4031	0.3118	0.7317	0.5187	0.6251	0.8712	0.8442	1.3552
SK_SGS_from_128	0.3478	0.1820	0.4010	0.3255	0.7368	0.5273	0.6443	0.8596	0.8442	1.3277
OK_nst	0.1041	0.0579	0.5686	0.2661	0.7681	0.5129	0.7253	0.9073	0.8866	0.4488
OK_SGS_from_2	0.1122	0.0943	0.7909	0.0010	0.4976	0.2732	0.0029	0.9999	0.5452	0.8518
OK_SGS_from_4	0.1267	0.0964	0.6685	0.1422	0.6855	0.4249	0.1874	0.9945	0.8411	0.8984
OK_SGS_from_8	0.1477	0.1033	0.5980	0.2656	0.7397	0.5029	0.5937	0.9452	0.9235	0.9277
OK_SGS_from_16	0.1678	0.1028	0.5614	0.3327	0.7667	0.5436	0.8871	0.8930	0.9268	0.9393
OK_SGS_from_32	0.1856	0.1049	0.5232	0.3711	0.7806	0.5675	0.9160	0.8501	0.9104	0.9523
OK_SGS_from_64	0.2011	0.1068	0.5282	0.3940	0.7854	0.5805	0.9250	0.8243	0.8977	0.9657
OK_SGS_from_128	0.2094	0.1077	0.5249	0.4078	0.7876	0.5878	0.9262	0.8094	0.8901	0.9723
GP(G)_nst	0.1260	0.0523	0.4856	0.3930	0.8229	0.5892	0.9585	0.8111	0.8954	0.4675
GP(G)_CRF_from_2	0.0557	0.0273	0.8617	0.0009	0.5166	0.2795	0.0047	0.9999	0.5578	0.5170
GP(G)_CRF_from_4	0.0658	0.0320	0.7424	0.1343	0.7035	0.4317	0.1440	0.9983	0.8605	0.5508
GP(G)_CRF_from_8	0.0796	0.0384	0.6740	0.2748	0.7643	0.5138	0.8532	0.9481	0.9498	0.5693
GP(G)_CRF_from_16	0.0904	0.0409	0.6303	0.3495	0.7948	0.5608	0.9606	0.8708	0.9280	0.5658
GP(G)_CRF_from_32	0.0990	0.0440	0.6063	0.3931	0.8111	0.5866	0.9768	0.8211	0.9053	0.5687
GP(G)_CRF_from_64	0.1052	0.0451	0.5851	0.4170	0.8208	0.6020	0.9800	0.7909	0.8908	0.5631
GP(G)_CRF_from_128	0.1094	0.0457	0.5745	0.4297	0.8257	0.6103	0.9803	0.7748	0.8831	0.5615
GP(L)_nst	0.0896	0.0356	0.6368	0.3935	0.8273	0.5933	0.9618	0.8032	0.8918	0.4748
GP(L)_SGS_from_2	0.0576	0.0341	0.8247	0.0012	0.5177	0.2811	0.0036	0.9999	0.5611	0.5814
GP(L)_SGS_from_4	0.0579	0.0331	0.8499	0.1504	0.7001	0.4369	0.1680	0.9980	0.8705	0.6202
GP(L)_SGS_from_8	0.0606	0.0355	0.8354	0.2909	0.7622	0.5208	0.8662	0.9382	0.9490	0.6323
GP(L)_SGS_from_16	0.0687	0.0376	0.8105	0.3651	0.7915	0.5665	0.9533	0.8596	0.9226	0.6372
GP(L)_SGS_from_32	0.0732	0.0392	0.7924	0.4065	0.8066	0.5914	0.9653	0.8116	0.9006	0.6404
GP(L)_SGS_from_64	0.0793	0.0403	0.7791	0.4304	0.8137	0.6067	0.9741	0.7820	0.8866	0.6426
GP(L)_SGS_from_128	0.0815	0.0409	0.7718	0.4454	0.8197	0.6153	0.9781	0.7651	0.8786	0.6417

Table S.9: Sample-weighted performance statistics for in-situ regression across domains and inference periods

	Histogram distances		Spatial Fidelity	Abs. Synchronicity		Likelihood	Accuracy	Precision	Goodness	Interval
Model	h_{JS}	h_{EM}	F	$\lvert S\rvert_{.25}$	$\lvert S\rvert_{.75}$	L	$A_{\xi=.05}$	P	G	I
SK_nst	0.1363	0.0739	0.4880	0.2679	0.7577	0.5111	0.6414	0.8785	0.8399	0.4702
SK_SGS_from_2	0.1170	0.1065	0.7157	0.0008	0.4452	0.2529	0.0040	1.0000	0.5045	0.8039
SK_SGS_from_4	0.1421	0.1081	0.6078	0.1119	0.6302	0.3863	0.1932	0.9927	0.7612	0.8610
SK_SGS_from_8	0.1631	0.1114	0.5600	0.2106	0.6983	0.4584	0.3888	0.9403	0.8269	0.8820
SK_SGS_from_16	0.1822	0.1141	0.5363	0.2665	0.7288	0.4968	0.5521	0.9001	0.8436	0.8955
SK_SGS_from_32	0.1962	0.1152	0.5162	0.2960	0.7442	0.5185	0.6091	0.8720	0.8451	0.8992
SK_SGS_from_64	0.2054	0.1160	0.5052	0.3133	0.7514	0.5300	0.6355	0.8558	0.8438	0.9017
SK_SGS_from_128	0.2111	0.1163	0.4972	0.3222	0.7555	0.5364	0.6448	0.8467	0.8428	0.9034
OK_nst	0.0582	0.0499	0.5464	0.3796	0.8264	0.5888	0.8929	0.7950	0.8704	0.4039
OK_SGS_from_2	0.0411	0.0573	0.8464	0.0013	0.5275	0.2895	0.0046	1.0000	0.5778	0.6559
OK_SGS_from_4	0.0549	0.0606	0.7257	0.1704	0.7099	0.4503	0.2459	0.9918	0.8880	0.7063
OK_SGS_from_8	0.0693	0.0642	0.6519	0.3150	0.7735	0.5375	0.9299	0.9133	0.9450	0.7206
OK_SGS_from_16	0.0832	0.0666	0.6221	0.3911	0.7997	0.5830	0.9725	0.8304	0.9118	0.7330
OK_SGS_from_32	0.0951	0.0683	0.6039	0.4299	0.8133	0.6075	0.9743	0.7817	0.8878	0.7404
OK_SGS_from_64	0.1030	0.0690	0.5909	0.4521	0.8202	0.6210	0.9753	0.7544	0.8741	0.7427
OK_SGS_from_128	0.1077	0.0694	0.5826	0.4640	0.8235	0.6282	0.9753	0.7398	0.8665	0.7455
GP(G)_nst	0.0388	0.0290	0.6255	0.4737	0.8591	0.6429	0.9909	0.7120	0.8544	0.3652
GP(G)_CRF_from_2	0.0169	0.0118	0.8973	0.0006	0.5401	0.2899	0.0058	1.0000	0.5786	0.3827
GP(G)_CRF_from_4	0.0210	0.0169	0.8112	0.1491	0.7185	0.4447	0.1423	0.9994	0.8882	0.4063
GP(G)_CRF_from_8	0.0254	0.0205	0.7537	0.2927	0.7797	0.5294	0.9709	0.9335	0.9591	0.4119
GP(G)_CRF_from_16	0.0293	0.0225	0.7230	0.3691	0.8068	0.5740	0.9962	0.8506	0.9242	0.4120
GP(G)_CRF_from_32	0.0316	0.0237	0.7073	0.4082	0.8215	0.5979	0.9971	0.8034	0.9012	0.4118
GP(G)_CRF_from_64	0.0328	0.0243	0.6995	0.4291	0.8280	0.6107	0.9978	0.7779	0.8887	0.4126
GP(G)_CRF_from_128	0.0338	0.0246	0.6960	0.4389	0.8312	0.6171	0.9983	0.7651	0.8823	0.4130
GP(L)_nst	0.0327	0.0249	0.6486	0.5020	0.8705	0.6587	0.9930	0.6807	0.8389	0.3575
GP(L)_SGS_from_2	0.0152	0.0103	0.9277	0.0007	0.5468	0.2933	0.0061	1.0000	0.5855	0.3897
GP(L)_SGS_from_4	0.0181	0.0143	0.8422	0.1632	0.7298	0.4551	0.2585	0.9985	0.9077	0.4122
GP(L)_SGS_from_8	0.0210	0.0176	0.7838	0.3129	0.7875	0.5413	0.9890	0.9129	0.9520	0.4174
GP(L)_SGS_from_16	0.0245	0.0197	0.7526	0.3920	0.8166	0.5882	0.9971	0.8226	0.9106	0.4173
GP(L)_SGS_from_32	0.0263	0.0208	0.7370	0.4312	0.8291	0.6125	0.9978	0.7743	0.8869	0.4170
GP(L)_SGS_from_64	0.0279	0.0213	0.7287	0.4533	0.8364	0.6261	0.9981	0.7472	0.8734	0.4158
GP(L)_SGS_from_128	0.0289	0.0216	0.7247	0.4645	0.8406	0.6330	0.9989	0.7333	0.8665	0.4156