

Supplementary Material:  
*Aggregated outputs by linear models:*  
*An application on waste accumulation prediction*

Jerónimo Hernández-González<sup>a</sup>, Iñaki Inza<sup>a</sup>, Igor Granado<sup>b</sup>, Oihane C. Basurko<sup>b</sup>, Jose A. Fernandes<sup>b</sup>, Jose A. Lozano<sup>a,c</sup>

<sup>a</sup>*Department of Computer Science and Artificial Intelligence, University of the Basque Country UPV/EHU, Donostia, Spain.*

<sup>b</sup>*Marine Research Division at AZTI-Tecnalia, Pasaia, Spain.*

<sup>c</sup>*Basque Center for Applied Mathematics, Bilbao, Spain.*

Table 1: Training examples in Municipality #3 for different scenarios. Division per column shows subsets of examples gathered not after each year. Each row shows the number of examples in the bags for a different maximum bag size,  $\check{m}_i$ . The last row shows the number of fully supervised examples, which is not affected by  $\check{m}_i$ .

$\check{m}_i$	$\leq \text{year}$							
	2009	2010	2011	2012	2013	2014	2015	2016
1	66	144	218	245	245	245	280	346
2	132	288	436	490	490	490	560	692
3	189	402	590	668	668	668	769	945
4	232	483	698	795	795	795	923	1131
5	256	526	762	875	875	875	1023	1256
6	273	557	806	932	932	932	1096	1348
7	282	571	827	965	965	965	1141	1408
fully sup.	48	120	210	215	215	215	231	259

Table 2: Training examples in Municipality #2 for different scenarios. Division per column shows subsets of examples gathered not after each year. Each row shows the number of examples in the bags for a different maximum bag size,  $\check{m}_i$ . The last row shows the number of fully supervised examples, which is not affected by  $\check{m}_i$ .

$\check{m}_i$	$\leq \text{year}$							
	2009	2010	2011	2012	2013	2014	2015	2016
1	57	124	185	257	311	352	402	458
2	114	248	370	514	622	704	804	916
3	165	355	527	731	887	1008	1151	1313
4	205	435	642	882	1073	1228	1409	1614
5	238	498	730	998	1216	1398	1612	1848
6	263	549	803	1094	1334	1543	1787	2048
7	283	588	863	1169	1428	1657	1926	2209
fully sup.	9	43	66	84	97	106	121	133

Table 3: Training examples in Municipality #3 for different scenarios. Division per column shows subsets of examples gathered not after each year. Each row shows the number of examples in the bags for a different maximum bag size,  $\check{m}_i$ . The last row shows the number of fully supervised examples, which is not affected by  $\check{m}_i$ .

$\check{m}_i$	$\leq \text{year}$							
	2009	2010	2011	2012	2013	2014	2015	2016
1	70	135	205	268	336	390	401	468
2	140	270	410	536	672	780	802	936
3	195	369	563	746	926	1081	1114	1290
4	230	439	670	899	1113	1301	1345	1552
5	259	498	761	1031	1273	1483	1538	1770
6	277	542	832	1134	1399	1627	1693	1944
7	291	572	876	1200	1477	1720	1796	2063
fully sup.	41	108	145	154	183	205	206	229

Table 4: Training examples in Municipality #4 for different scenarios. Division per column shows subsets of examples gathered not after each year. Each row shows the number of examples in the bags for a different maximum bag size,  $\check{m}_i$ . The last row shows the number of fully supervised examples, which is not affected by  $\check{m}_i$ .

$\check{m}_i$	$\leq \text{year}$							
	2009	2010	2011	2012	2013	2014	2015	2016
1	39	76	108	146	187	218	225	257
2	78	152	216	292	374	436	450	514
3	114	217	312	423	538	628	649	740
4	146	273	399	543	686	799	827	943
5	175	323	480	657	824	955	990	1127
6	201	365	552	762	947	1094	1136	1291
7	224	402	615	854	1055	1214	1263	1433
fully sup.	5	30	37	42	47	53	53	54

Table 5: Training examples in Municipality #5 for different scenarios. Division per column shows subsets of examples gathered not after each year. Each row shows the number of examples in the bags for a different maximum bag size,  $\check{m}_i$ . The last row shows the number of fully supervised examples, which is not affected by  $\check{m}_i$ .

$\check{m}_i$	$\leq \text{year}$							
	2009	2010	2011	2012	2013	2014	2015	2016
1	58	122	178	238	302	370	385	457
2	116	244	356	476	604	740	770	914
3	153	322	475	642	816	994	1039	1227
4	180	385	573	775	986	1196	1254	1472
5	204	440	658	883	1126	1363	1431	1673
6	228	489	734	982	1256	1517	1595	1860
7	251	537	805	1076	1376	1661	1749	2034
fully sup.	72	140	236	293	334	379	379	407

Table 6: Training examples in Municipality #6 for different scenarios. Division per column shows subsets of examples gathered not after each year. Each row shows the number of examples in the bags for a different maximum bag size,  $\check{m}_i$ . The last row shows the number of fully supervised examples, which is not affected by  $\check{m}_i$ .

$\check{m}_i$	$\leq \text{year}$							
	2009	2010	2011	2012	2013	2014	2015	2016
1	77	143	213	278	346	414	426	489
2	154	286	426	556	692	828	852	978
3	206	379	570	752	932	1124	1159	1328
4	241	446	676	894	1107	1339	1385	1585
5	267	497	755	1000	1235	1500	1556	1780
6	286	535	811	1079	1333	1622	1687	1931
7	301	560	850	1135	1402	1710	1784	2045
fully sup.	38	132	184	215	251	275	279	311

Table 7: Training examples in Municipality #7 for different scenarios. Division per column shows subsets of examples gathered not after each year. Each row shows the number of examples in the bags for a different maximum bag size,  $\check{m}_i$ . The last row shows the number of fully supervised examples, which is not affected by  $\check{m}_i$ .

$\check{m}_i$	$\leq \text{year}$							
	2009	2010	2011	2012	2013	2014	2015	2016
1	73	133	181	241	310	374	418	480
2	146	266	362	482	620	748	836	960
3	184	340	463	613	791	951	1069	1232
4	204	381	522	694	893	1076	1218	1407
5	213	404	560	747	959	1160	1320	1529
6	217	417	584	781	1001	1215	1390	1609
7	219	424	600	807	1031	1254	1440	1665
fully sup.	144	297	455	580	698	813	952	1082

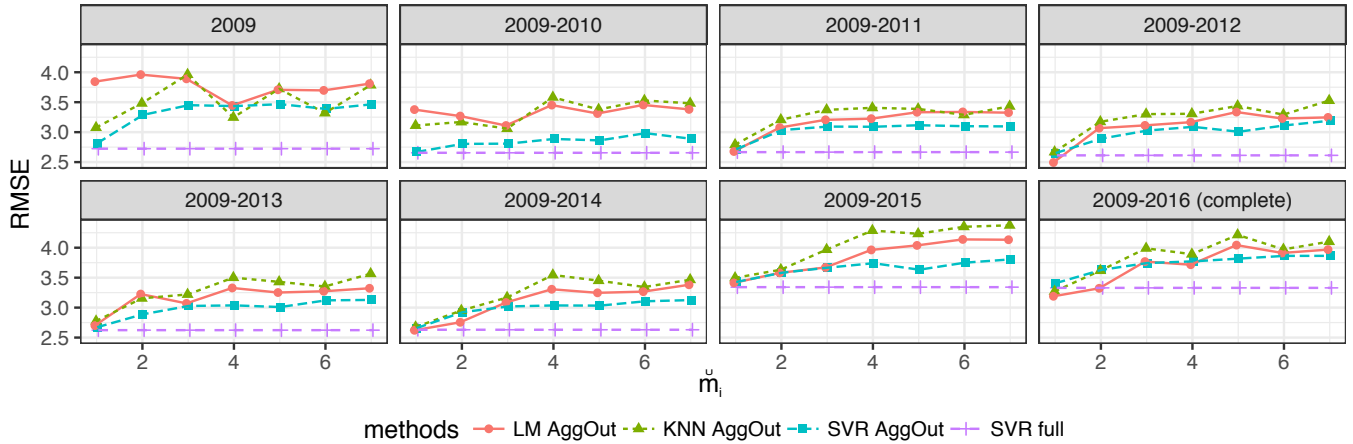


Figure 1: Results in terms of root mean square error of linear, KNN and SVR models learnt from the AO data of the beaches of Municipality #1. SVR models learnt only with the available supervised examples are used as a baseline. Each subfigure shows the results on a different subset (examples collected  $\leq$  year) pruning the large bags by using an increasing maximum value for bag size,  $\check{m}_i$ .

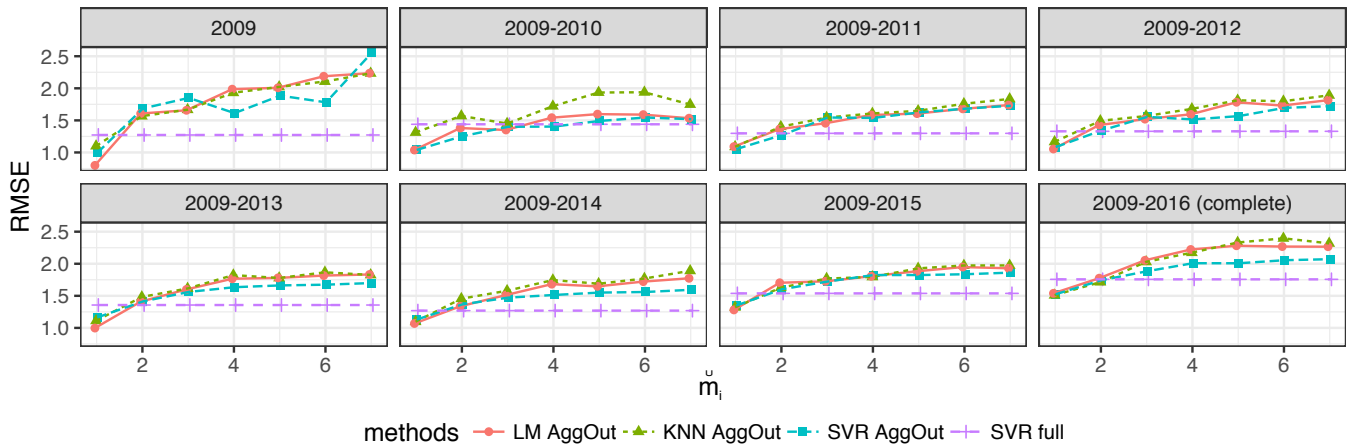


Figure 2: Results in terms of root mean square error of linear, KNN and SVR models learnt from the AO data of the beaches of Municipality #2. SVR models learnt only with the available supervised examples are used as a baseline. Each subfigure shows the results on a different subset (examples collected  $\leq$  year) pruning the large bags by using an increasing maximum value for bag size,  $\check{m}_i$ .

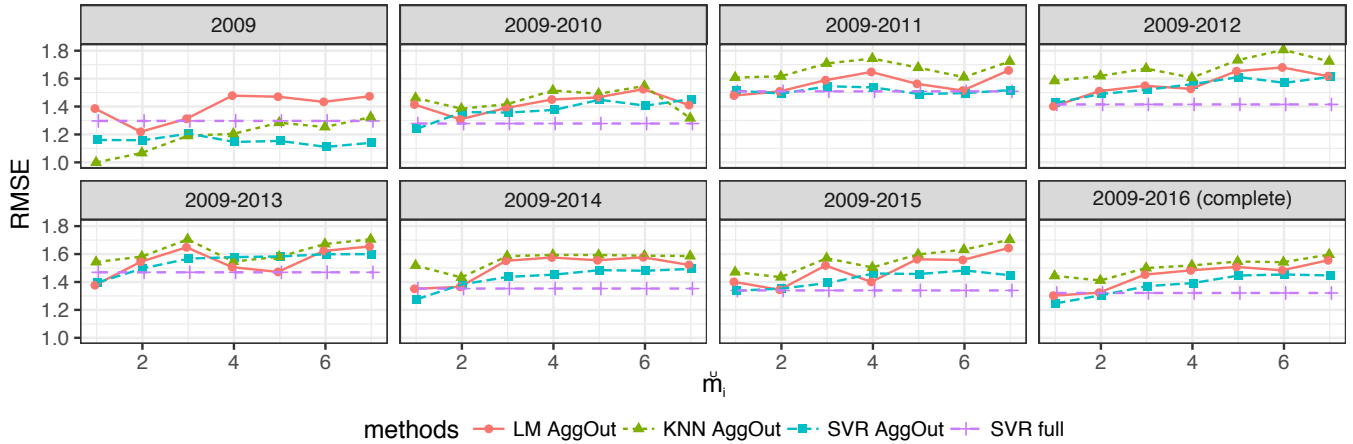


Figure 3: Results in terms of root mean square error of linear, KNN and SVR models learnt from the AO data of the beaches of Municipality #3. SVR models learnt only with the available supervised examples are used as a baseline. Each subfigure shows the results on a different subset (examples collected  $\leq$  year) pruning the large bags by using an increasing maximum value for bag size,  $\check{m}_i$ .

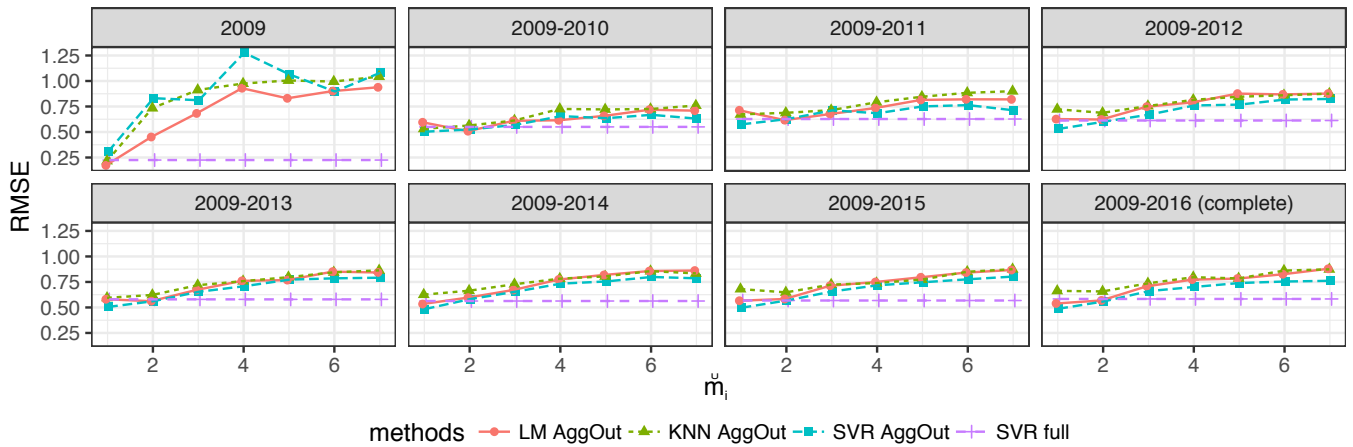


Figure 4: Results in terms of root mean square error of linear, KNN and SVR models learnt from the AO data of the beaches of Municipality #4. SVR models learnt only with the available supervised examples are used as a baseline. Each subfigure shows the results on a different subset (examples collected  $\leq$  year) pruning the large bags by using an increasing maximum value for bag size,  $\check{m}_i$ .

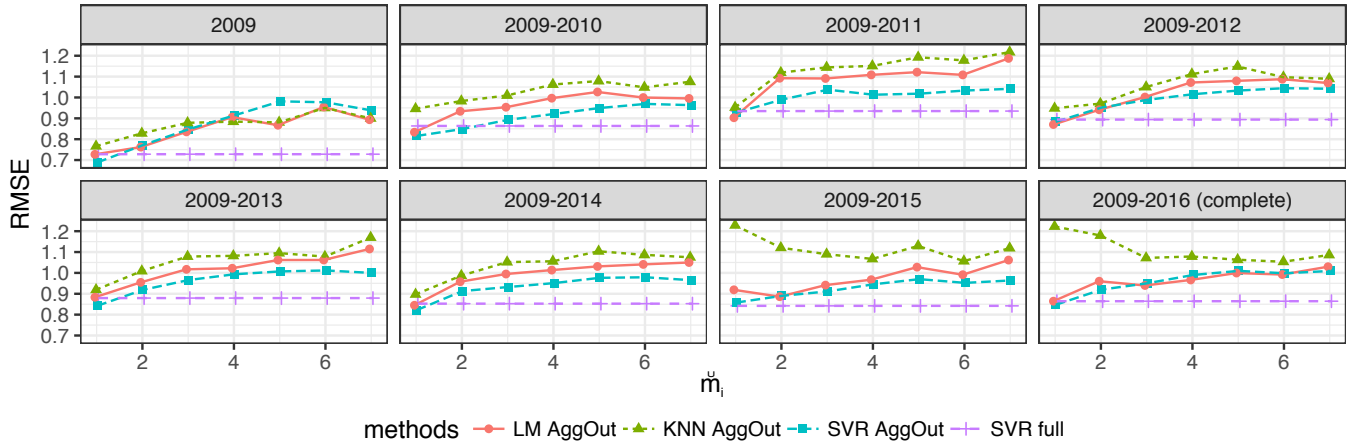


Figure 5: Results in terms of root mean square error of linear, KNN and SVR models learnt from the AO data of the beaches of Municipality #5. SVR models learnt only with the available supervised examples are used as a baseline. Each subfigure shows the results on a different subset (examples collected  $\leq$  year) pruning the large bags by using an increasing maximum value for bag size,  $\check{m}_i$ .

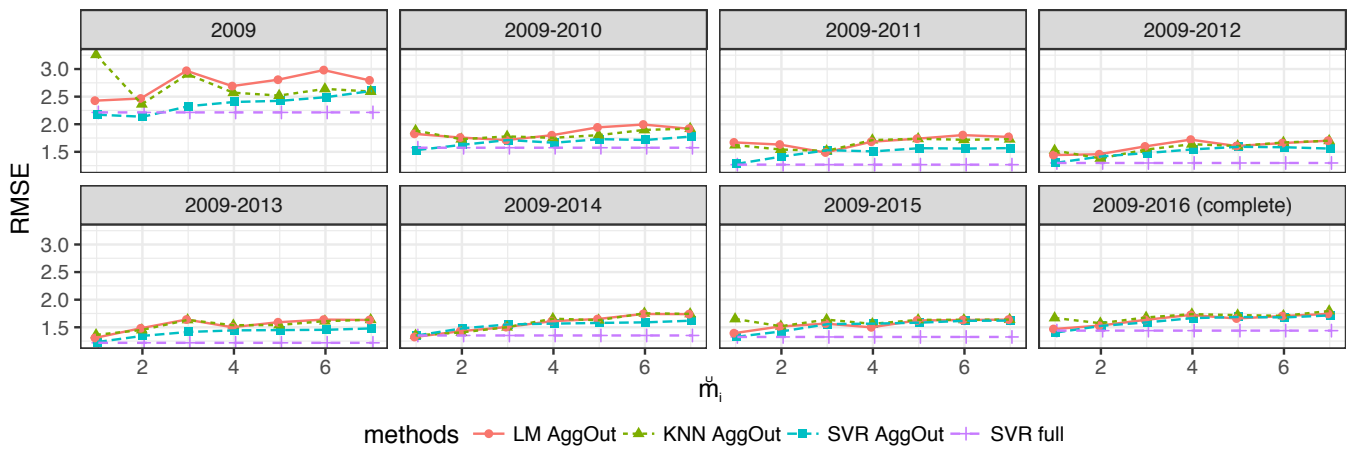


Figure 6: Results in terms of root mean square error of linear, KNN and SVR models learnt from the AO data of the beaches of Municipality #6. SVR models learnt only with the available supervised examples are used as a baseline. Each subfigure shows the results on a different subset (examples collected  $\leq$  year) pruning the large bags by using an increasing maximum value for bag size,  $\check{m}_i$ .

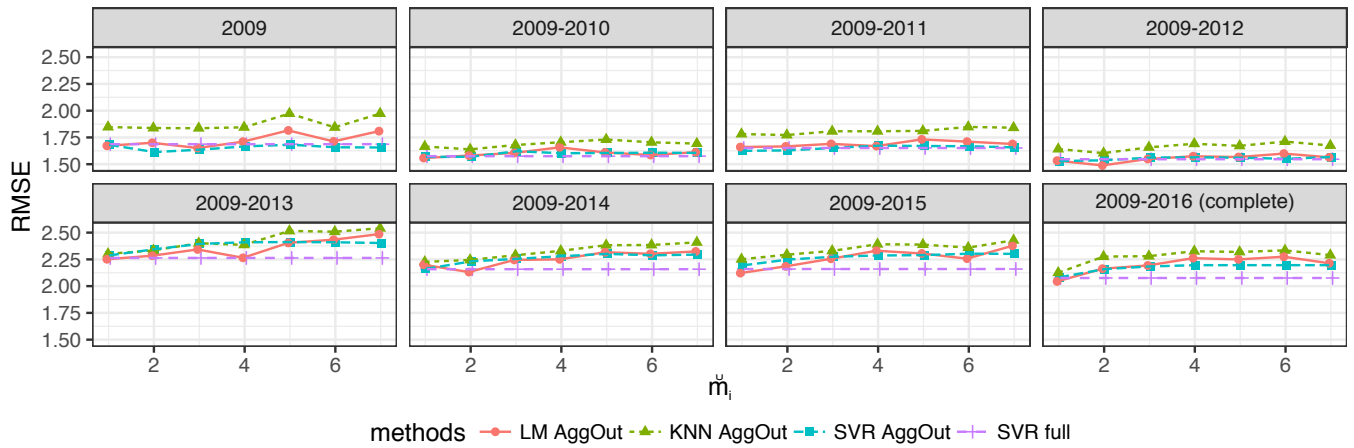


Figure 7: Results in terms of root mean square error of linear, KNN and SVR models learnt from the AO data of the beaches of Municipality #7. SVR models learnt only with the available supervised examples are used as a baseline. Each subfigure shows the results on a different subset (examples collected  $\leq$  year) pruning the large bags by using an increasing maximum value for bag size,  $\bar{m}_i$ .