

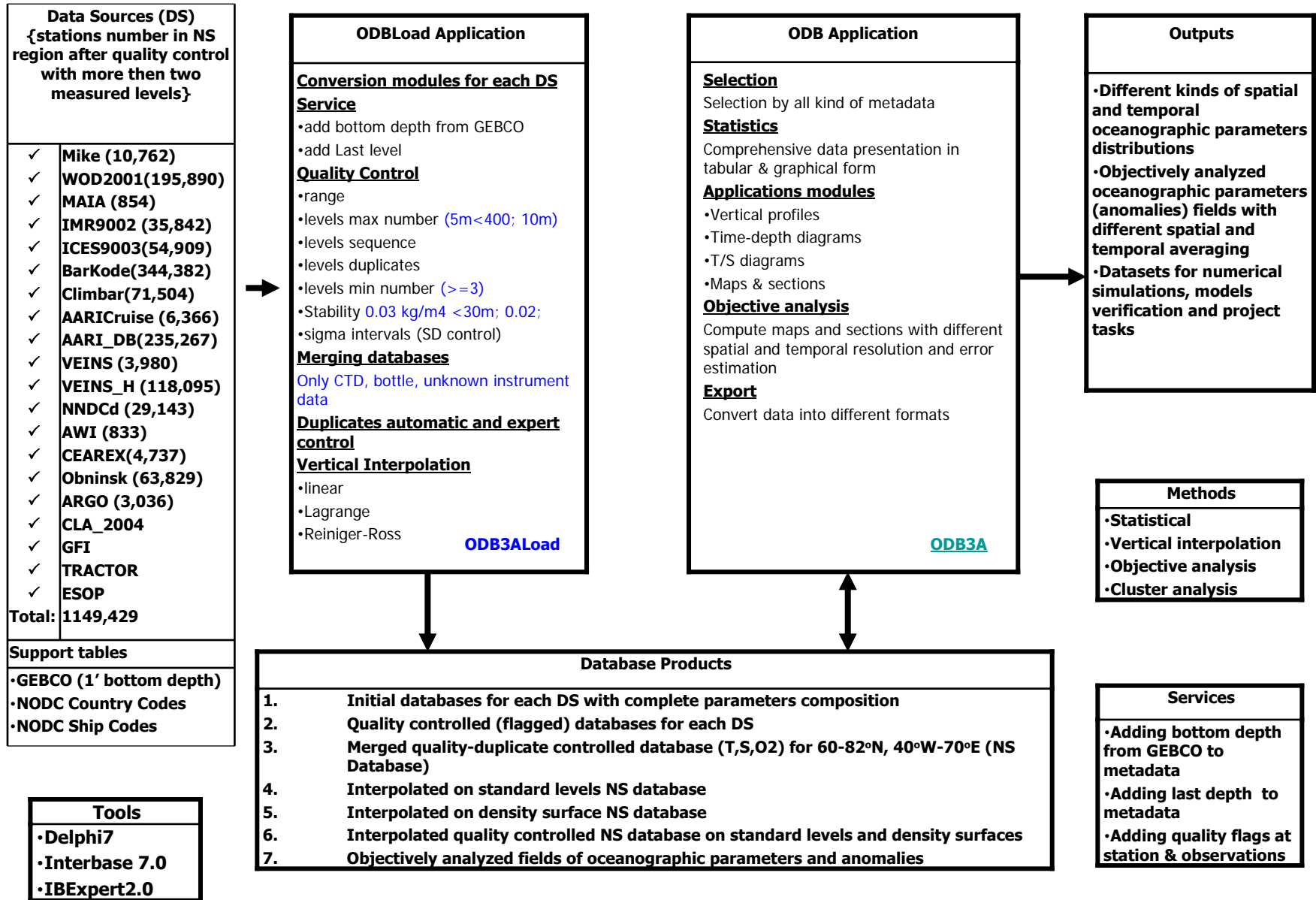
Alexander Korablev NERSC/AARI
Alexander Smirnov AARI/NIERSC
Andrey Pnyushkov AARI
Ola M. Johannesen NERSC/GFI

Compilation of the oceanographic database for the Nordic Seas

- ✓ Initial data collection and database structure
- ✓ Data processing and quality control
- ✓ Duplicates analysis
- ✓ Vertical interpolation
- ✓ Software for data processing and analysis
- ✓ AARI contribution into UNESCO/IOC/GADAR project

NODC/NOAA Silver Spring, 24 April, 2008

Implementation schema for software and NS database developing



Initial sources databases (original region, all parameters)

	Database	Latitude	Longitude	Period	Stations	Size, Kb
1	AARIOD_i.IB	48.30->82.47	-69.44->67.50	28.01.1901->18.08.2000	295998	665540
2	Argo_i.ib	60.00->79.01	-39.56->26.20	26.02.1998->06.01.2005	2494	15408
3	AWI_CD_I.IB	69.49->86.28	-18.18->161.37	10.06.1987->02.10.2003	1119	141672
4	BARKODE_I.IB	63.58->84.00	5.00->109.60	23.05.1898->16.09.1998	639144	1136504
5	CEAREX_I.IB	59.32->89.33	-176.35->144.24	06.08.1900->17.05.1989	6705	101744
6	CLIMATL_i.IB	64.00->90.00	0.00->149.58	19.04.1810->17.08.2001	379071	1279172
7	CLIMATLZ_I.IB	36.39->90.00	-11.02->161.41	19.04.1810->22.07.2001	339747	1289604
8	Climbar_i.ib	65.52->81.37	16.28->68.30	23.05.1898->01.07.1993	74280	162440
9	ESOP_I.IB	62.21->79.15	-27.29->23.01	11.02.1993->27.09.1998	3077	314744
10	GSP_i.IB	62.22->80.00	-26.20->18.01	16.04.1987->21.03.1993	1589	12524
11	Ices_ctd_i.ib	60.00->81.53	-30.00->64.30	17.01.1990->10.03.2003	42398	-1610432
12	ICES9003LR_I.IB	60.00->81.58	-30.01->70.00	01.01.1990->10.03.2003	56897	260776
13	IMR9002_I.IB	60.00->81.15	-39.30->64.30	17.01.1990->12.12.2002	36051	1216120
14	MAIA_I.IB	59.50->74.17	-6.37->20.01	29.04.2000->30.10.2001	859	36728
15	MIKE4803.IB	62.02->69.15	-4.00->8.04	01.10.1948->11.07.2003	10904	18647
16	NANSEN_i.IB	49.60->74.30	-31.21->14.05	01.06.1986->06.06.1989	1602	9344
17	NNDCD_I.IB	50.00->86.12	-50.00->79.58	05.04.1900->11.06.1977	113314	157116
18	NPI_I.IB	76.45->81.16	-12.28->31.60	29.08.2005->14.09.2005	117	10120
19	Npcd_i.ib	74.13->81.20	6.06->26.26	19.09.1959->28.11.1992	3193	7476
20	OBNINSK_I.IB	61.00->79.59	-27.43->25.58	20.05.1896->25.02.1986	60138	208068
21	Overflow_i.IB	56.15->70.00	-40.52->12.25	30.05.1960->20.09.1973	2313	8884
22	PnSec_I.IB	60.35->80.00	-14.00->27.33	31.12.1980->29.12.1990	1093	4656
23	TRACTOR_I.IB	59.36->82.21	-35.14->20.03	23.02.2001->13.10.2003	1206	142212
24	VEINS_H_I.IB	58.50->84.42	-45.00->25.01	15.04.1900->16.02.2000	156458	624944
25	VEINS_I.IB	51.55->82.29	-49.23->20.18	11.04.1997->16.08.2000	2397	202568
26	WOD2001_I.IB	60.00->90.00	-39.59->70.00	18.09.1866->11.12.1999	261930	565296
		Total:			2494094	6981875

Metadata (19 fields in two separate tables)

Date / Time: 8 äi ääëý 2004 ä.
 Database: C:\ODB\Database\MIKE4803.IB
 User: SYSDBA
 Table: STATION

Table: STATION

Fields			
Name	Type	Domain	Not Null
ABSNUM	INTEGER	RDB\$1	NOT NULL
STFLAG	SMALLINT	RDB\$2	
Default Value:	0		
STLAT	NUMERIC(8,5)	RDB\$3	NOT NULL
STLON	NUMERIC(9,5)	RDB\$4	NOT NULL
STDATE	DATE	RDB\$5	NOT NULL
STTIME	TIME	RDB\$6	NOT NULL
STSOURCE	VARCHAR(10)	RDB\$7	NOT NULL
STVERSION	SMALLINT	RDB\$8	NOT NULL
STCOUNTRYNAME	VARCHAR(40)	RDB\$9	
Default Value:	'UNKNOWN'		
STVESSELNAME	VARCHAR(40)	RDB\$10	
Default Value:	'UNKNOWN'		
STDEPTHSOURCE	INTEGER	RDB\$11	
Default Value:	-9		
STLASTLEVEL	SMALLINT	RDB\$12	
Default Value:	-9		
STDEPTHGRID	SMALLINT	RDB\$13	
Default Value:	-9		
STDEPTHGRIDMIN	SMALLINT	RDB\$14	
Default Value:	-9		
STDEPTHGRIDMAX	SMALLINT	RDB\$15	
Default Value:	-9		

Date / Time: 8 äi ääëý 2004 ä.
 Database: C:\ODB\Database\MIKE4803.IB
 User: SYSDBA
 Table: STATION_INFO

Table: STATION_INFO

Fields			
Name	Type	Domain	Not Null
ABSNUM	INTEGER	RDB\$31	NOT NULL
COUNTRYCODE	VARCHAR(2)	RDB\$32	
Default Value:	"		
VESSELCODE	VARCHAR(4)	RDB\$33	
Default Value:	"		
STNUMINCUISE	VARCHAR(10)	RDB\$34	
Default Value:	"		
PROJECTCODE	INTEGER	RDB\$35	
Default Value:	-9		
INSTITUTECODE	INTEGER	RDB\$36	
Default Value:	-9		
INSTRUMENT	INTEGER	RDB\$37	
Default Value:	-9		
SOURCEUNIQUEID	VARCHAR(10)	RDB\$38	
Default Value:	"		
SOURCEDATAORIGIN	VARCHAR(15)	RDB\$39	
Default Value:	"		
VESSELCRUISEID	VARCHAR(10)	RDB\$40	
Default Value:	"		

Date / Time: 8 äi ääëý 2004 ä.
 Database: C:\ODB\Database\MIKE4803.IB
 User: SYSDBA
 Table: P_TEMPERATURE

Table: P_TEMPERATURE

Fields			
Name	Type	Domain	Not Null
ABSNUM	INTEGER	RDB\$59	NOT NULL
LEVELNUM	SMALLINT	RDB\$60	NOT NULL
LEVEL_	NUMERIC(5,1)	RDB\$61	NOT NULL
VALUE_	NUMERIC(5,3)	RDB\$62	NOT NULL
FLAG_	SMALLINT	RDB\$63	NOT NULL

Vessel codes problem!

New vessel codes table was compiled using vessel codes from all initial DS

Quality control flags

On station

1	32768	Erroneous station
2	16384	Questionable station
3	8192	Hydrochemistry didn't pass quality control
4	4096	Metadata compiled from different DS
5	2048	Profiles compiled from different DS
6	1024	
7	512	NODC ship code not unique or not found according table ShipCode_List (SupportTables.ib)
8	256	Precision of temperature probe < 0.01°C
9	128	Precision of salinity probe < 0.02psu
10	64	Station date was changed (hour=24 or minutes=60)
11	32	Station date was changed (days in month exceed the limit)
12	16	Station time is missing (hours is absent)
13	8	
14	4	Time interval between stations in cruise exceed the limit
15	2	Station on land
16	1	Last level at station greater that minimal depth inside 5 km radius from GEBCO or station on land
	0	Flags not signed

On measured value

1	32768	Erroneous value
2	16384	Questionable value
3	8192	
4	4096	Value outside 5 σ interval
5	2048	Value outside 4 σ interval
6	1024	Value outside 3 σ interval
7	512	Value from gradient zone
8	256	Value compiled from duplicates
9	128	Vertical stability (t,s) or vertical gradient error
10	64	Range error (value outside coarse interval)
11	32	Level depth computed from pressure
12	16	Hydrochemistry units recomputed (see DS documentation)
13	8	Interpolated value (according DS)
14	4	Erroneous value (according DS)
15	2	Questionable value (according DS)
16	1	Accuracy lost (according DS)
	0	Flags not assigned

Merged, quality controlled initial databases prepared for duplicate control procedure. The Nordic Seas area: 60°-82°N, 40°W-70°E.
(Only stations with more than two levels with temperature or salinity)

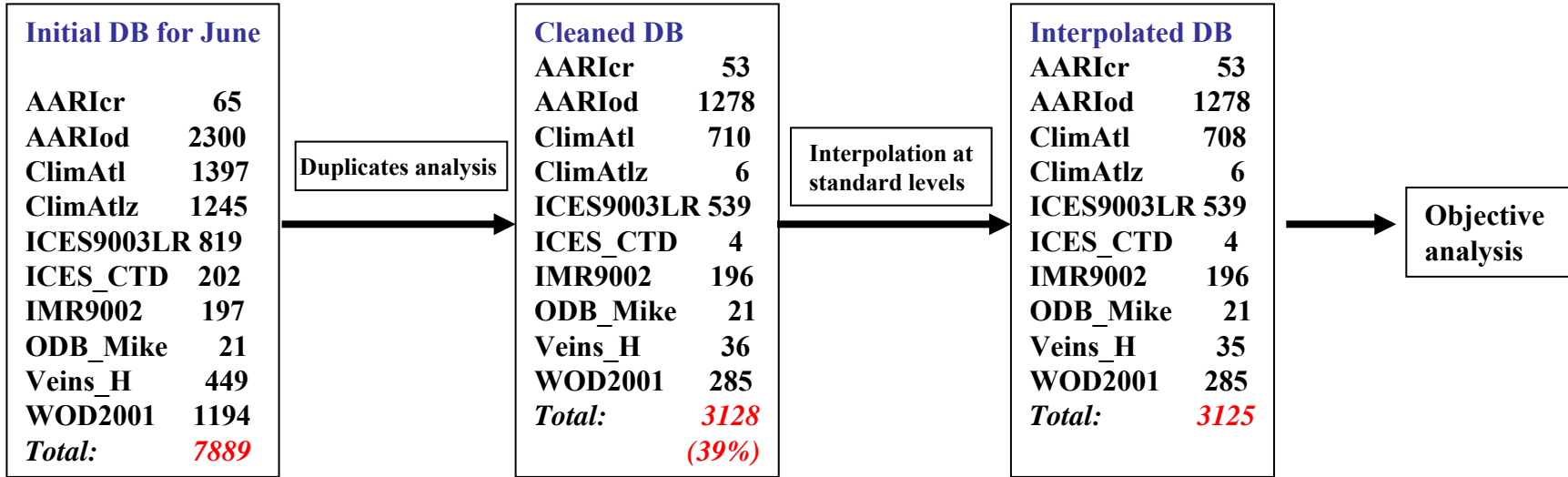
	Database	Latitude	Longitude	Period	Stations	Size, Kb
1	AARICR.IB	59.59->80.23	-15.00->23.10	28.03.1984->21.03.1993	6366	27616
2	AARIOD.IB	60.00->81.57	-40.00->67.50	28.01.1901->18.08.2000	250217	565176
3	ARGO.IB	60.00->79.01	-39.56->26.20	26.02.1998->06.01.2005	2489	15688
4	AWI_CD.IB	69.49->82.00	-18.18->68.20	10.06.1987->02.10.2003	834	20012
5	CEAREX.IB	60.38->81.60	-18.23->69.47	14.08.1900->17.05.1989	4737	48876
6	ESOP.IB	62.21->79.15	-27.29->23.01	11.02.1993->27.09.1998	2988	51616
7	GSP.IB	62.22->80.00	-26.20->18.01	16.04.1987->21.03.1993	1587	11708
8	IMR9002.IB	60.00->81.15	-39.30->64.30	17.01.1990->12.12.2002	35843	330604
9	MAIA.IB	60.02->74.17	-6.37->20.01	29.04.2000->30.10.2001	854	7264
10	NANSEN.IB	60.00->74.30	-27.40->14.05	01.06.1986->06.06.1989	1418	6460
11	NNDCD.IB	60.00->81.59	-40.00->70.00	05.04.1900->09.06.1977	29143	68588
12	NPICD.IB	74.13->81.20	6.06->26.26	19.09.1959->28.11.1992	3110	6912
13	OBNINSK.IB	61.00->79.59	-27.43->27.33	20.05.1896->25.02.1986	57570	219664
14	ODB_CA.IB	60.57->82.00	0.00->70.00	10.01.1900->22.07.2001	209312	581740
15	ODB_ICES.IB	60.00->81.58	-30.01->70.00	01.01.1990->10.03.2003	56036	463724
16	ODB_Mike.IB	62.02->69.15	-4.00->8.04	01.10.1948->11.07.2003	10762	22008
17	OVERFLOW.IB	60.00->70.00	-38.40->-1.37	30.05.1960->20.09.1973	2136	8268
18	PnSec.IB	60.35->80.00	-14.00->27.33	31.12.1980->29.12.1990	1088	4668
19	TRACTOR.IB	60.16->81.58	-35.14->20.00	23.02.2001->13.10.2003	1040	22600
20	VEINS.IB	60.10->81.51	-39.50->20.18	11.04.1997->14.08.2000	2324	28912
21	VEINS_H.IB	60.00->81.57	-40.00->25.01	15.04.1900->16.02.2000	124194	493028
22	WOD2001.IB	60.00->81.59	-39.59->70.00	23.08.1875->04.12.1999	191026	489568
	Total:				995074	3494700

Duplicates analysis module (checking all stations inside one minute grid box)

Purpose: find best metadata and profiles composition for each oceanographic station by means of automatic and expert controls

Alg	StAccess Flag	Write to ODB	Applied duplicates control algorithm
	0	No	Initial Flag.
	1	Yes	No duplicates.
1	2	Yes	Absolute Duplicate. All profiles and metadata the same.
	200	No	
2	3	Yes	Full Duplicate. T,S,O2 profiles the same, metadata coincide partly. Merging profiles and MD (StNumInCruise, InstituteCode, ProjectCode, Instrument)
	300	No	
3	4	Yes	TSO2 Duplicate. T, S, O2 profiles the same, metadata does not coincide Merging profiles and MD (StCountryName, StVesselName, StDepthSource, CountryCode, VesselCode, StNumInCruise, InstituteCode, ProjectCode, Instrument).
	400	No	
4	5	Yes	Interpolated or reduced resolution profiles. Merging profiles and MD (StCountryName, StVesselName, StDepthSource, CountryCode, VesselCode, StNumInCruise, InstituteCode, ProjectCode, Instrument).
	500	No	
5	6	Yes	Multi-day staions. Time is different. Merging profiles and MD (StCountryName, StVesselName, StDepthSource, CountryCode, VesselCode, StNumInCruise, InstituteCode, ProjectCode, Instrument).
	600	No	
6	7	Yes	Multi-day staions. Time is unknown. Merging profiles and MD (StCountryName, StVesselName, StDepthSource, CountryCode, VesselCode, StNumInCruise, InstituteCode, ProjectCode, Instrument).
	700	No	
	8		Reserved
	800		Reserved
7	9	Yes	Expert control. Merging selected profiles and MD
	900	No	

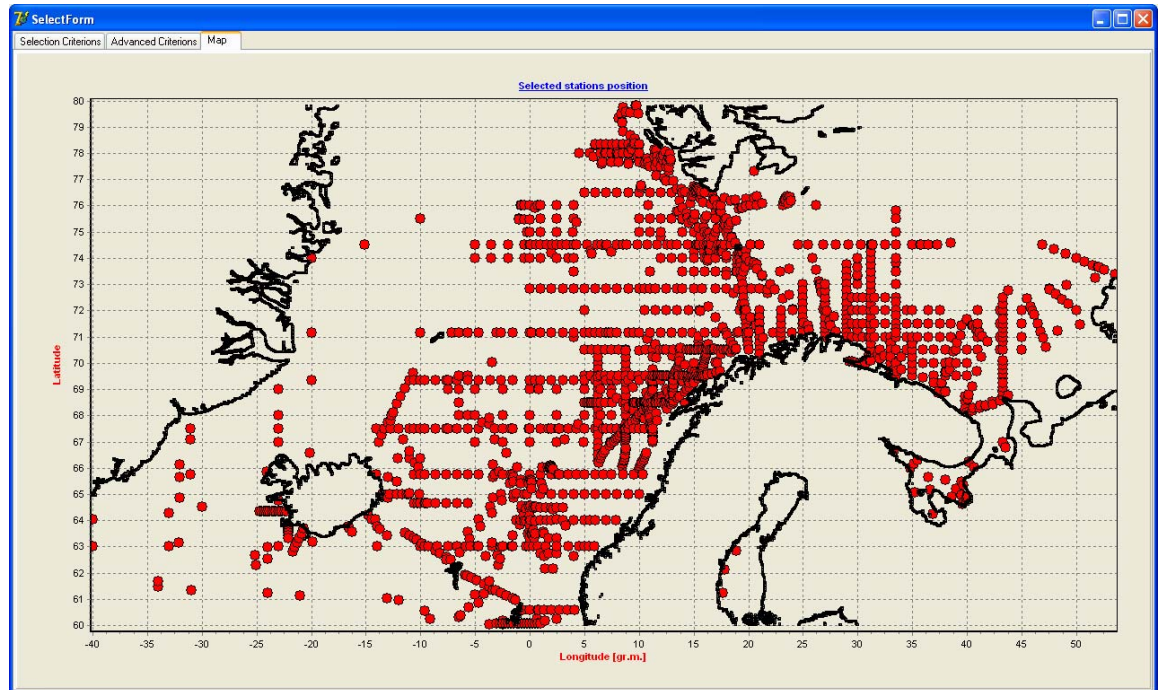
Merging databases (an example for June 1990)



Duplicates Analysis

1	No duplicates	580
2	Absolute duplicates	643
3	Full duplicates	53
4	TSO2 duplicates	162
5	Interpolated/ reduced resolution	46
6	Multi day stations time <	1491
7	Multi day stations time unknown	45
9	Expert control	115
200		1846
300		157
400		1846
500		624
900		251

7523 stations (>95%) have found automatically



Duplicate control, visual control, SD control, vertical interpolation procedures

Monthly ODB Initial->Interpolated	After duplicates control		After Visual cont		Set SD flags (in situ)			Interpolated Profiles		Set SD flags (interpolated)			Stations summary in DB				Comments
													ODB_ (initial)		_i (interpolated)		
	Par.	Profiles #	Prf.#	Meg.#	3 sd	4 sd	5 sd	Prf.#	Meg.#	3 sd	4 sd	5 sd	Duplicate	Station	Initial	Visual C.	
January (SK) ODB_Jan.ib->Jan_i.ib	T	17032	17032	327506	2185	869	425	17030	202857	1529	610	261	51761	17040 32.9%	16885		
	S	14758	14758	294496	3419	1284	478	14704	177562	1533	560	238					
	O2	2348	2348	23130	62	19	0	2345	32319	92	28	0					
February (SK,AS) ODB_Feb.ib->Feb_i.ib	T	23502	23502	506609	2784	879	365	23501	275049	1434	560	193	68541	23511 34.3%	23365		
	S	20875	20875	469990	4250	1546	694	20871	246505	1672	527	297					
	O2	3409	3409	33929	80	16	3	3403	45232	121	12	4					
March (AK) ODB_Mar.ib->Mar_i.ib	T	28631	28460	692312	3967	1315	632	28459	298259	1893	670	305	81548	28464 34.9%	28252		
	S	25264	25076	635453	5698	1921	842	25053	336283	2343	885	366					
	O2	4539	4480	43294	275	119	34	4474	58606	360	129	38					
April (AK) ODB_Apr.ib->Apr_i.ib	T	36800	36800	827491	4900	1357	554	36798	424084	2204	662	264	120861	36788 30.4%	36586		
	S	33199	33199	777169	7942	2696	1141	33117	380058	3012	1050	462					
	O2	5872	5872	52081	312	94	42	5844	67217	372	71	32					
May (AS) ODB_May.ib->May_i.ib	T	42428	42260	982331	5676	1942	883	42245	516871	3030	1160	559	114203	42259 37%	41995		
	S	37576	37380	921349	9183	3390	1644	37258	457583	4154	1390	712					
	O2	8121	7761	87919	627	182	66	7649	96564	610	166	55					
June (AS) ODB_Jun.ib->Jun_i.ib	T	52030	52030	1064117	8299	3052	536	52021	689567	5872	2076	871	135151	52009 38.5%	51734		
	S	44957	44957	956692	10952	4052	942	44800	593807	5778	1824	842					
	O2	12403	12403	138445	1255	432	94	12283	164433	1604	412	147					
July (AK) ODB_Jul.ib->Jul_i.ib	T	41555	41495	1025937	6662	1915	757	41484	507805	3107	838	325	98749	41513 42.1%	41087		
	S	35083	34996	927033	10909	4066	1555	34860	434725	3720	1369	710					
	O2	6053	5981	70682	360	98	31	5972	81490	403	85	24					
August (SK) ODB_Aug.ib->Aug_i.ib	T	53069	52962	1317487	8062	2368	981	52933	629669	3877	1074	459	121134	53005 43.8%	52398		
	S	43716	43577	1177368	12495	4095	1845	43405	525141	4910	1669	797					
	O2	6990	6849	92037	754	245	91	6796	89719	475	69	12					
September (SK) ODB_Sep.ib->Sep_i.ib	T	41882	41761	1004340	5581	1524	559	41743	478938	2368	707	305	100639	41790 41.5%	41261		
	S	34956	34778	905563	8812	2881	1358	34700	404376	3083	976	509					
	O2	5791	5623	70786	526	196	84	5566	66231	246	63	26					
October (AS) ODB_Oct.ib->Oct_i.ib	T	29102	28981	529946	2781	941	447	28976	327433	1488	432	196	65607	28992 44.2%	28750		
	S	25515	25309	481090	4861	1838	941	25126	285439	2196	715	304					
	O2	7979	7867	75590	410	78	14	7810	88324	77	43	5					
November (AP) ODB_Nov.ib->Nov_i.ib	T	21784	21638	522978	2992	979	402	23634	264009	1483	382	147	51021	21642 42.5%	21460		
	S	19182	18966	492127	4311	1483	667	18939	235476	1674	410	169					
	O2	4299	4036	42898	102	22	7	3991	51024	146	20	5					
December (AS) ODB_Dec.ib->Dec_i.ib	T	17147	17044	252460	1448	428	183	17044	203388	1114	389	177	36736	17047 46.4%	16908	ODB3ALoad	
	S	14615	14501	220952	1602	512	234	14476	174201	1311	482	204					
	O2	2474	2361	23006	77	15	4	2359	36616	102	24	7					
														404060	400581		

995074 st. / 600 (st/hour) = 1658 hours / 8 = 207 days / 30 = 6.9 month

Duplicates analysis (1)

Stations in DUPLICATE table: 3980

Duplicate control

Write into ODB

StAccess field content

Yes No

<input checked="" type="checkbox"/> not analysed	[0]	2379	0
<input checked="" type="checkbox"/> no duplicates	[1]	150	0
<input checked="" type="checkbox"/> Absolute duplicates	[2]	60	306
<input checked="" type="checkbox"/> Full duplicates	[3]	59	100
<input checked="" type="checkbox"/> TSQ2Coincidence	[4]	77	319
<input checked="" type="checkbox"/> Interpolated profiles	[5]	77	145
<input checked="" type="checkbox"/> Multi-day stations t <>	[6]	44	0
<input checked="" type="checkbox"/> Multi-day stations t < >	[7]	177	0
<input type="checkbox"/> ...reserved			
<input type="checkbox"/> Expert con			

Update

Preview

ShowCDSProfiles

Col Index: 3 Title: s_n Profile Name: SALINITY Number of profiles: 2

1	IMR9002 -> 26069	2	ICES9003 -> 10432
0.0	35.063 0	0.0	35.063 32
10.0	35.063 0	10.0	35.063 32
15.0	35.062 0	15.0	35.062 32
20.0	35.051 0	20.0	35.051 32
25.0	34.822 4	25.0	34.822 32
30.0	35.113 0	30.0	35.113 32
35.0	35.117 0	35.0	35.117 32
40.0	35.123 0	40.0	35.123 32
45.0	35.128 0	45.0	35.128 32
50.0	35.122 0	50.0	35.122 32
55.0	35.115 0	55.0	35.115 32
60.0	35.115 0	60.0	35.115 32
65.0	35.109 0	65.0	35.109 32
70.0	35.111 0	70.0	35.111 32
75.0	35.110 0	75.0	35.110 32
80.0	35.113 0	80.0	35.113 32
85.0	35.111 0	85.0	35.111 32
90.0	35.109 0	90.0	35.109 32
95.0	35.109 0	95.0	35.109 32
100.0	35.105 0	100.0	35.105 32
105.0	35.107 0	105.0	35.107 32

All profiles RecN

1
2

All profiles RecN

1
2

Memo1

*** \$STEP: 1 Cur

2 duplicate

.....MD coinc

lat lon

YES YES

[after Abs

[after Full

[after T,S

[after Mult

DuplicateCDSExpertControl

TabSheet1 TabSheet2

CDSMD

RecN	AbsD	AbsS	StAccess	StFlng	StLat	StLon	StDate	StTime	StSource	StVersion
1	884	26069	0	0	74	24.5	18.08.1990	7:02:00	IMR9002	1 NOI
2	1165	10432	2	0	74	24.5	18.08.1990	7:02:00	ICES9003	1 NOI

RecN	AbsD	AbsS	CountryCode	VesselCode	StNumInCruise	ProjectCode	Institute Code	Instrument	SourceUniqueID
1	884	26069	58	58MS	786	902	973	4	44914
2	1165	10432	58	58MS	##	-9	-9	-9	10432

Step: 1

Filter on: 2 st. set remove

PopulateODB

Current Record: 1

CDSPRF

number of levels

RecN	StAccess	t_n	s_n	o2_n	si_n	p_n	n_n	al_n	ch_n	ph_n
1	0	84	84	0	0	0	0	0	0	0
2	0	36	36	0	0	0	0	0	0	0

mean level

RecN	StAccess	t_jmd	s_jmd	o2_jmd	si_jmd	p_jmd	n_jmd	al_jmd	ch_jmd	ph_jmd
1	0	212.97	212.97							
2	0	147.5	147.5							

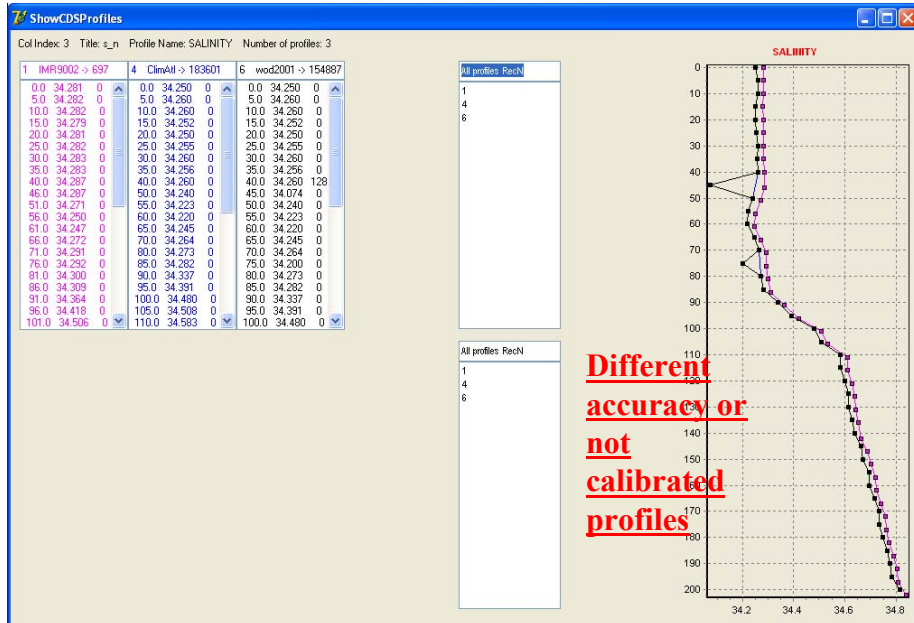
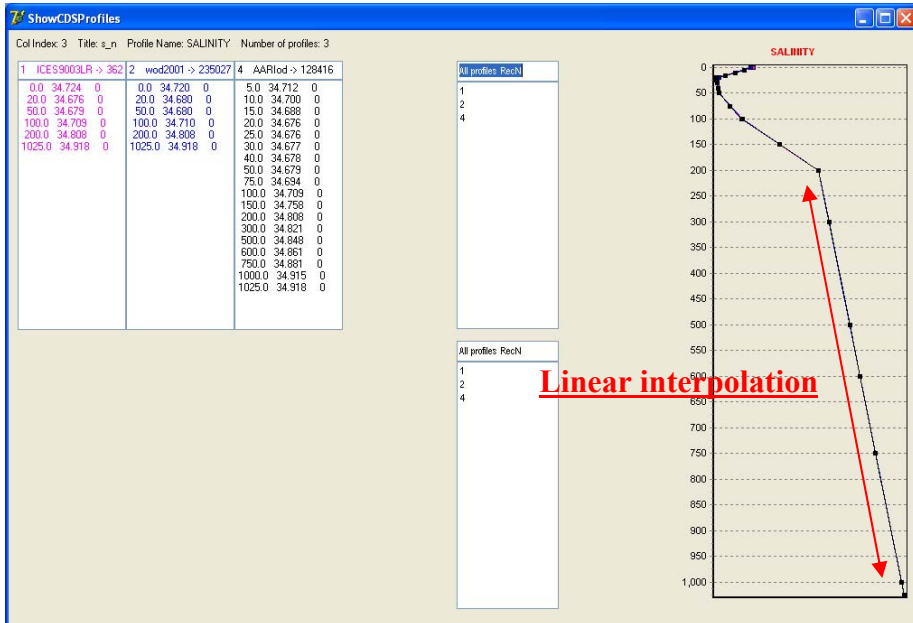
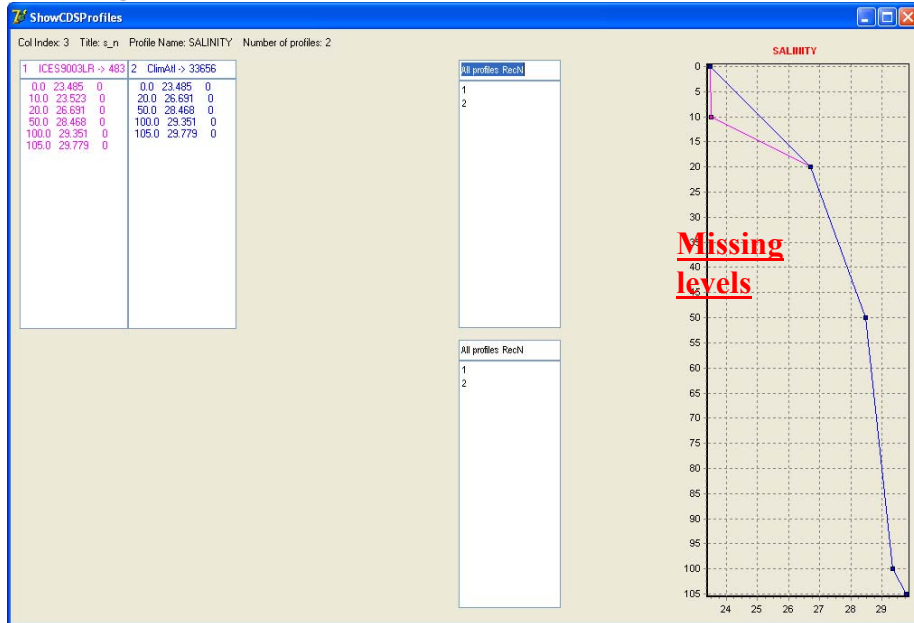
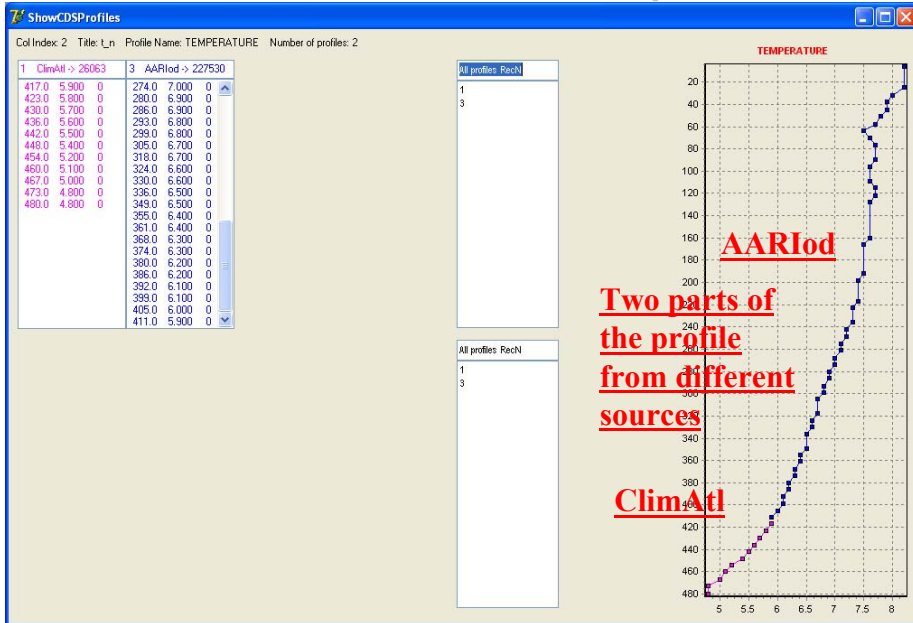
parent station' record number

RecN	StAccess	t_st	s_st	o2_st	si_st	p_st	n_st	al_st	ch_st	ph_st
1	0	1	1							
2	0	2	2							

mean value

RecN	StAccess	t_pmd	s_pmd	o2_pmd	si_pmd	p_pmd	n_pmd	al_pmd	ch_pmd	ph_pmd
1	0	4.4186	35.061							
2	0	5.2876	35.072							

Duplicates analysis (3)



Vertical Interpolation Module

Vertical interpolation can be the source for large errors in the datasets!

Many historical profiles have been interpolated forever!

For data homogeneity the common methods and limitations have to be applied for whole dataset

Values within upper 5 meters regarded as surface

IntAtStandardLevels [populate interpolated database]

SLev	inside and outside int. limits	
0	5.00	200.00
5	6.42	201.14
10	7.84	202.29
20	10.69	204.57
30	13.53	206.86
50	19.21	211.43
75	26.32	217.14
100	33.43	222.86
125	40.54	228.57
150	47.64	234.29
200	61.86	245.71
250	76.07	257.14
300	90.29	268.57
400	118.71	291.43
500	147.14	314.29
600	175.57	337.14
700	204.00	360.00
800	232.43	382.86
900	260.86	405.71
1000	289.29	428.57
1100	317.71	451.43
1200	346.14	474.29
1300	374.57	497.14
1400	403.00	520.00
1500	431.43	542.86
1750	502.50	600.00
2000	573.57	657.14
2250	644.64	714.29
2500	715.71	771.43
3000	857.86	885.71
3500	1000.00	1000.00
4000	1142.14	1114.29
4500	1284.29	1228.57
5000	1426.43	1342.86
5500	1568.57	1457.14

Buttons: Create Target Database, Write to ODB?, Show profiles?

Set skip flagged values:

- [1] accuracy lost
- [2] suspicious value
- [4] erroneous value
- [64] range error
- [128] stability error
- [1024] 3 SD
- [2048] 4 SD
- [4096] 5 SD

Memo1

Improvement?
Using of 3D objective analysis?

Flagged low-quality measurements are illuminated prior interpolation.

Instead stepwise 'inside' and 'outside' depth limitation criterion used in OCL datasets limits were computed from regression

An additional limitation was set on 3 and 4 points interpolation methods: if value computed by Lagrange or Reiniger and Ross algorithms deviated more then 25% from linear interpolated value – linear interpolation is substituted.

Depth interval	Acceptable depth difference	
	Inside values	Outside values
0-5	5	
10-200	50	
250-800	100	
900-1750	200	
2000-5500	1000	
0-400		200
500-1200		400
1300-5500		1000

SD control module for monthly data

SDControlByLevels

Levels: 0, 10, 20, 30, 50, 75, 100, 125, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1750, 2000, 2250, 2500, 3000, 3500, 4000, 4500, 5000, 5500

Parameter: ALKALINITY, CHLOROPHYLL, NITRATE, OXYGEN, PH, PHOSPHATE, SALINITY, SILICATE, TEMPERATURE

Month: January, February, March, April, May, June, July, August, September, October, November, December

UpdateSelection

Visualization: Current station, Sample stations, Histogram, Normal Distribution, Sample Points, Log

SD Analysis

Search [temporal spatial] [Years] [km]

Initial: 10 50
step: 10 50
limit: 60 300

SD control limit: 3
Normal distribution st. # control limit: 100

[1] accuracy lost
 [2] suspicious value
 [4] erroneous value
 [64] range error
 [128] stability error
 [1024] outside 3sd
 [2048] outside 4sd
 [4096] outside 5sd

p_value: 0.20, 0.15, 0.10, 0.05, 0.01

P_ALKALINITY->754
 P_CHLOROPHYLL->236
 P_NITRATE->3022
 P_OXYGEN->5872
 P_PH->1948
 P_PHOSPHATE->4116
 P_SALINITY->33199
 P_SILICATE->3130
 P_TEMPERATURE->36800
 0->5: 681
 40->62.5: 74207

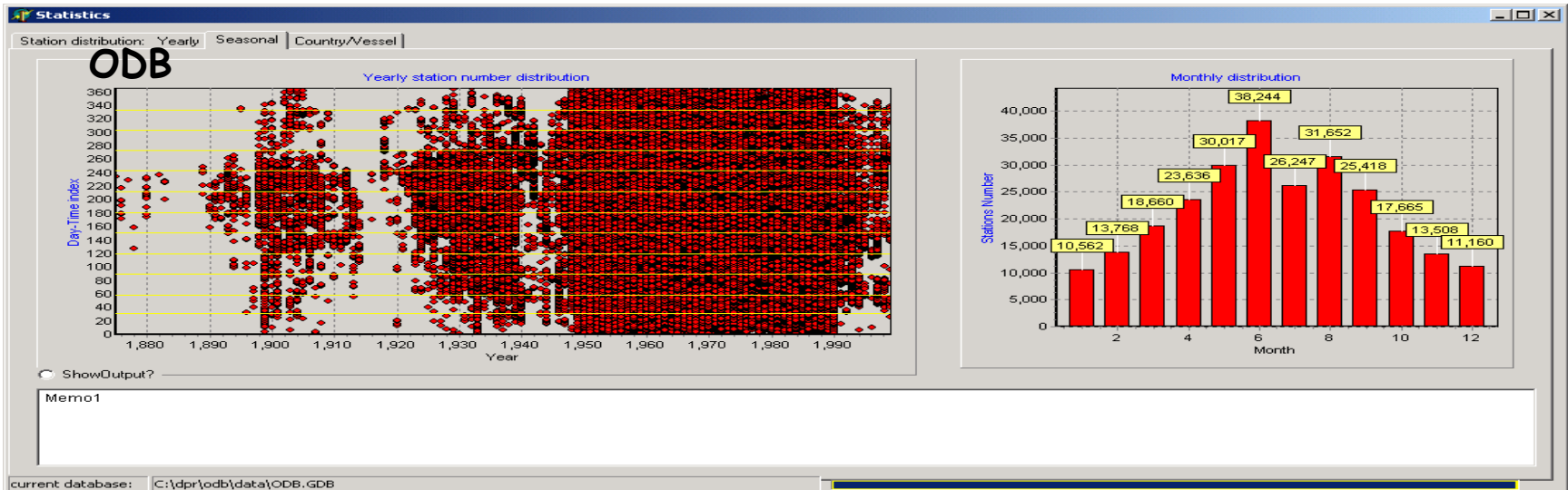
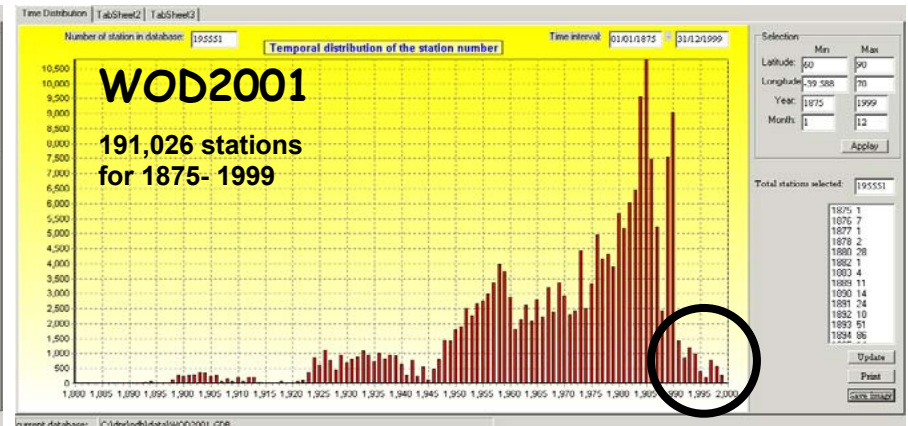
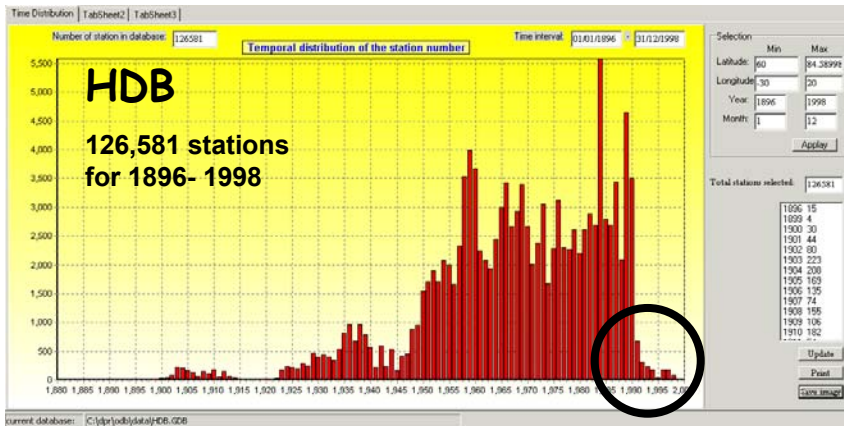
Count SD QControl flags

QFlag	Sample<=100	Sample>100
1024:	Edit9	Edit12
2048:	Edit10	Edit13
4096:	Edit11	Edit14

btnStart

Standard deviation control allows to set flags on data with value outside 3, 4 or 5 standard deviations. The algorithm starts from sample definition around each selected station. Initial criterions for selection are 50km in space around the station and +/- 10 years in time. On each step the normality of parameter distribution is estimated for sample (Hi-criterion and Chalmogorov-Smirnov criterion for small sample less then 35). If the normality condition is not complied at first step the sample size enlarged by 50km in dimension and 10 years in time, until reaching limit set as 300 km. If normal distribution not found – value not analyzed for standard deviations. If sample size exceeds 100 normality check does not applied.

Temporal distributions of the oceanographic stations number in the HDB (AARI, without the Barents Sea) and WOD2001 (OCL/NODC) databases show lack of the data for 1990s (top). Merged databases (ODB) demonstrate major gaps in data temporal and monthly distributions (bottom)



NS database content: comparison with OCL products

DB:	NS	WOD2001	WOD2005
Period:	10.04.1988- 14.04.2006	23.08.1875- 4.12.1999	23.08.1875- 26.07.2005
St# Initial	2,494,094		
St# QC	995,074		
St# DC Current	404,808	191,026	307,565
		-213,782	-97,273

NS database content: countries

	Country	Stations #
1	USSR/RUSSIA	132,081
2	NORWAY	123,740
3	UNKNOWN	68,678
4	UNITED KINDOM	20,408
5	ICELAND	19,808
6	UNITED STATES	15,946
7	GERMANY	11,157
8	DENMARK	5,125
9	POLAND	2,268
10	JAPAN	1,560
11	NETHERLAND	1,416
12	CANADA	1,203

NS database content: reference period for anomalies computing

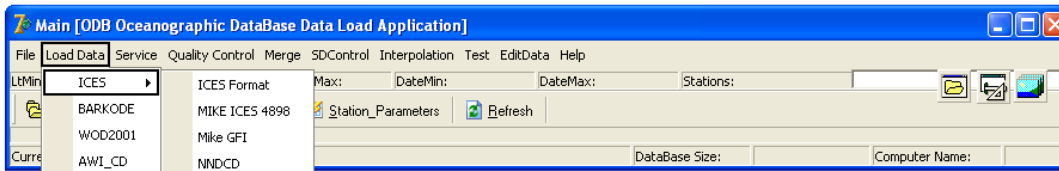
Year	0	1	2	3	4	5	6	7	8	9
1900	267	481	551	378	459	280	285	89	182	130
1910	221	76	172	306	62	0	7	0	69	15
1920	45	84	96	387	894	687	1458	824	457	942
1930	790	886	986	1298	1226	922	1568	1330	1776	1584
1940	1088	450	1280	857	1032	559	901	869	2024	1578
1950	2238	2629	3954	2987	4158	4382	4652	5528	7796	9335
1960	7530	6699	6037	7084	5761	5489	5481	5852	5326	6815
1970	6703	5635	6147	7946	4950	5751	8317	8115	7064	6261
1980	8675	8062	8503	10010	15222	13433	10796	12756	8780	155248
1990	14994	6312	6070	6660	6556	5280	4342	5730	5454	5209
2000	4268	5178	4135	1722	867	688	36			

Summary of the data processing

- Developing of the oceanographic database structure (Interbase 7.0 server), metadata composition and quality control flags
- Download data (more than 20 converters) from initial sources to separate databases
- Quality control for all initial databases (9 procedures)
- Merging initial databases for the Nordic Seas region with instruments (high accuracy) and profiles (≥ 3 levels) control
- Automatic and expert control of the duplicates with metadata merging
- Standard deviation flags ($>3,4,5 \sigma$) setting on the data in layers
- Interpolation at the standard levels
- Standard deviation flags ($>3,4,5 \sigma$) setting at the standard levels
- Objective analysis of the monthly fields (with kriging errors estimates)
- Mean and anomalies fields computing

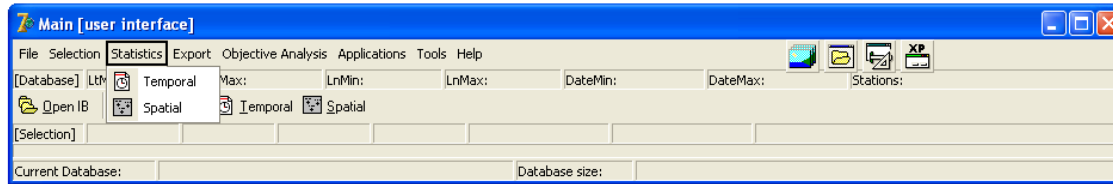
Software

ODB3ALoad



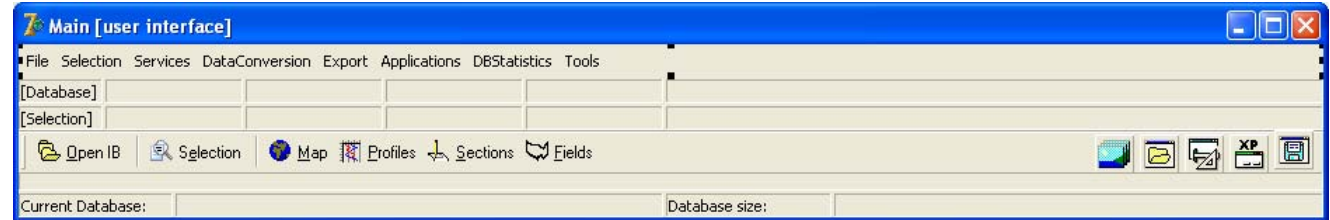
Quality-duplicate controlled, merged initial database

ODB3A



Objectively analyzed databases for requested grid net

ODB3AClimate



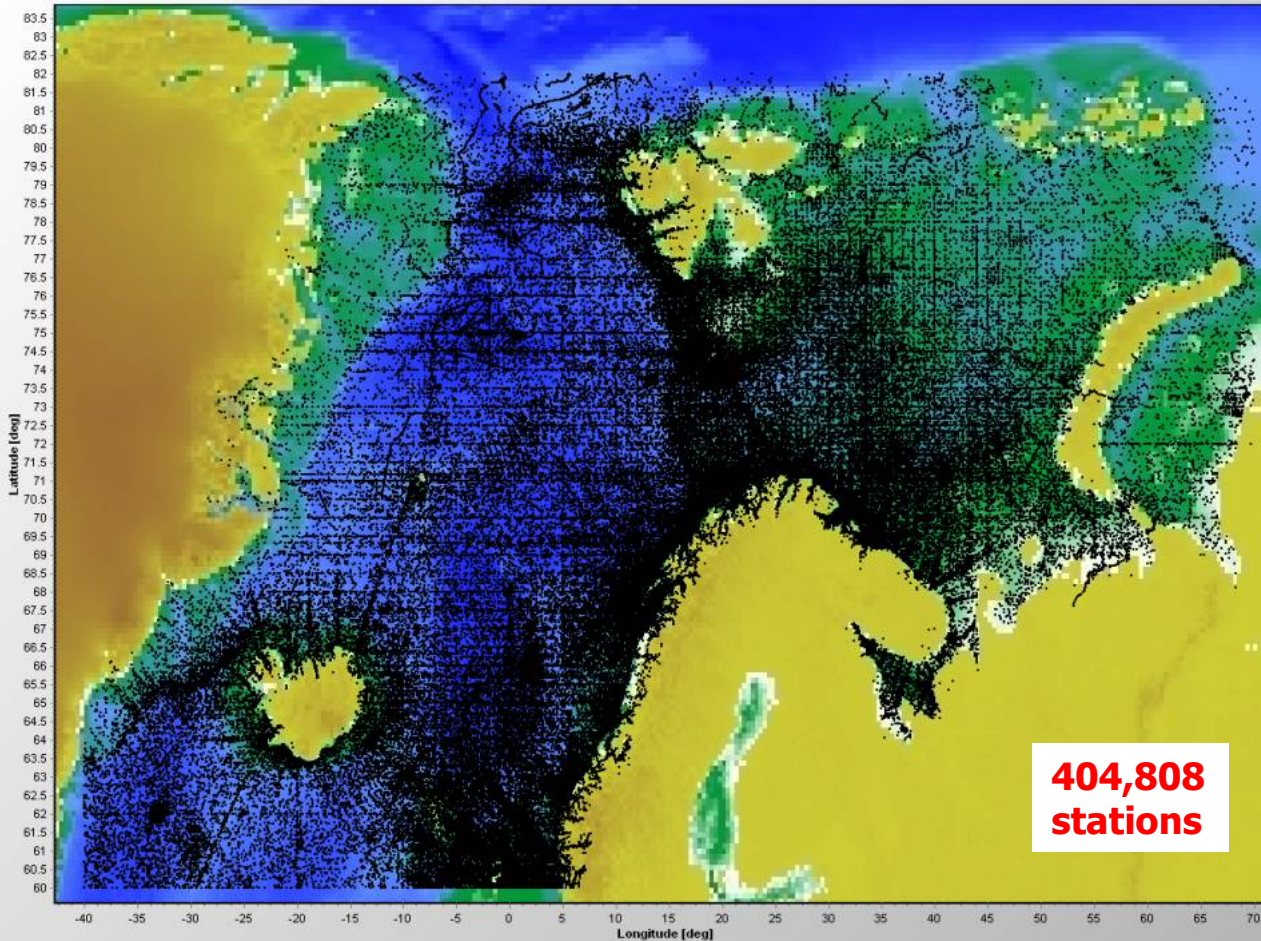
Mean and anomalies fields at standard levels, in layers and on density surfaces

Initial data -> climatology

**Korablev A.A., Pnyushkov A.V., A.V. Smirnov
Oceanographic database compiling for climate
monitoring in the North European Basin. Trudy
AARI. 2006**

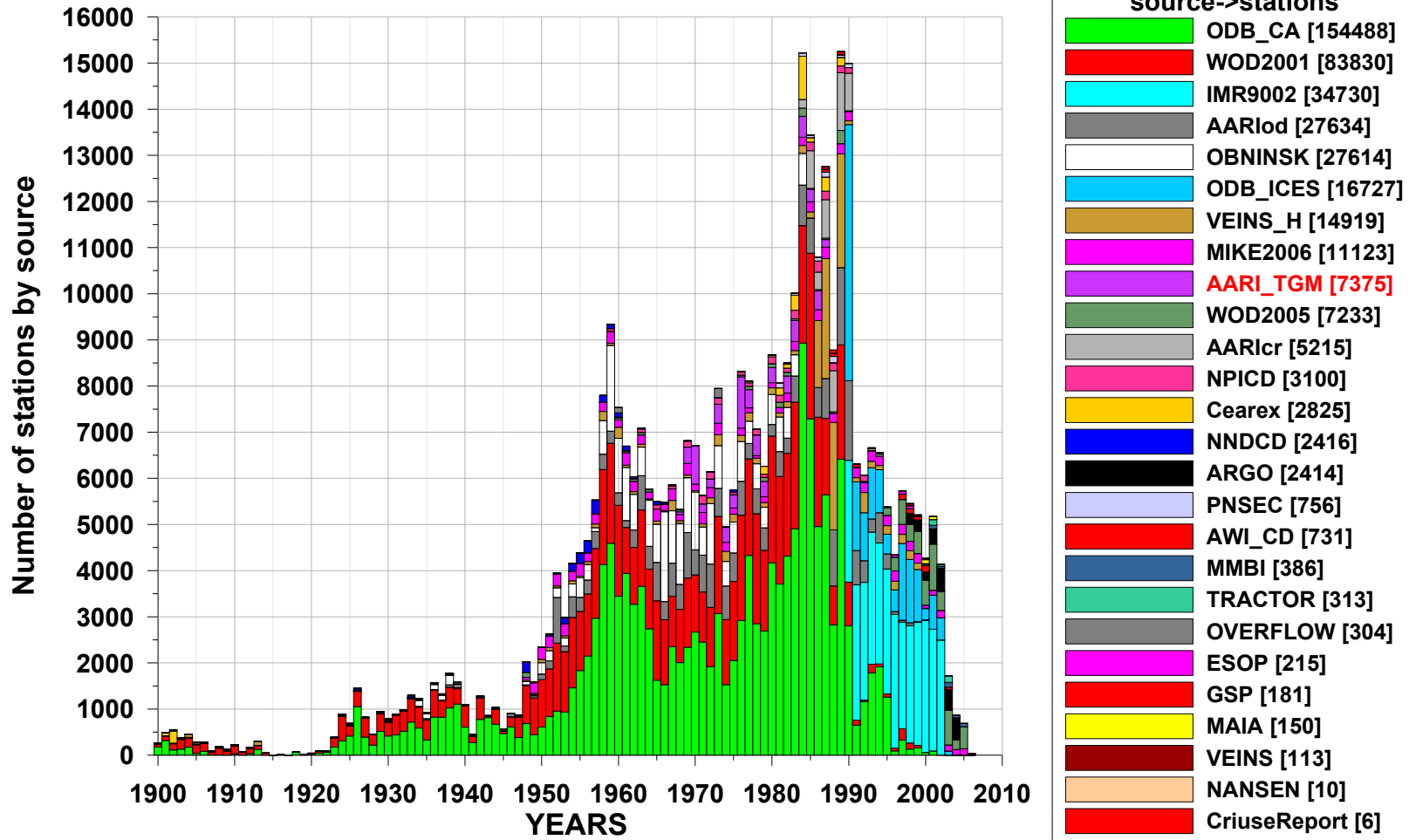
Oceanographic Database for the Nordic Seas (NS)

Area: 60-82°N, 40°W-70°E

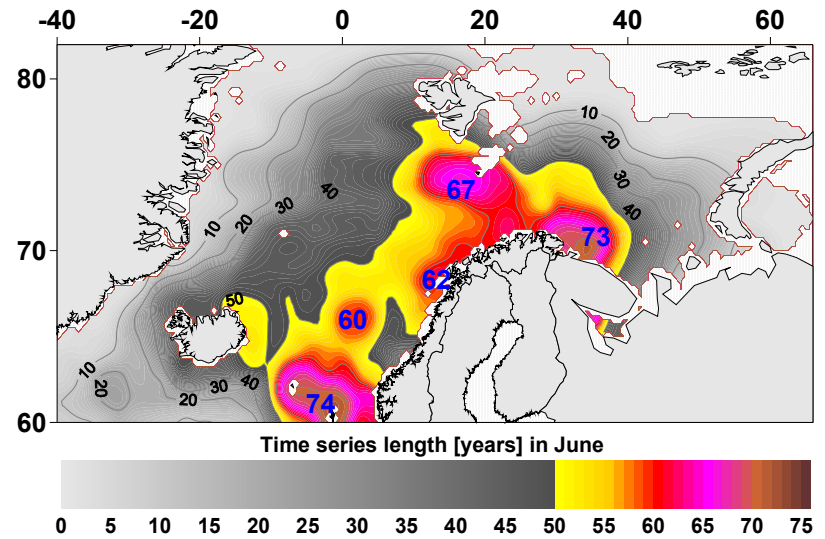
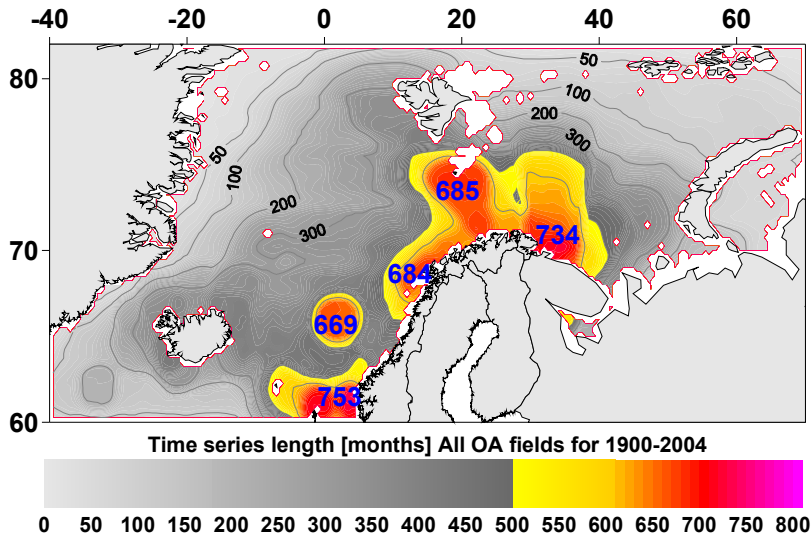


- only profiles with high accuracy (bottle, CTD, unknown instrument)
- profiles with more than three measured levels

NS database content: sources

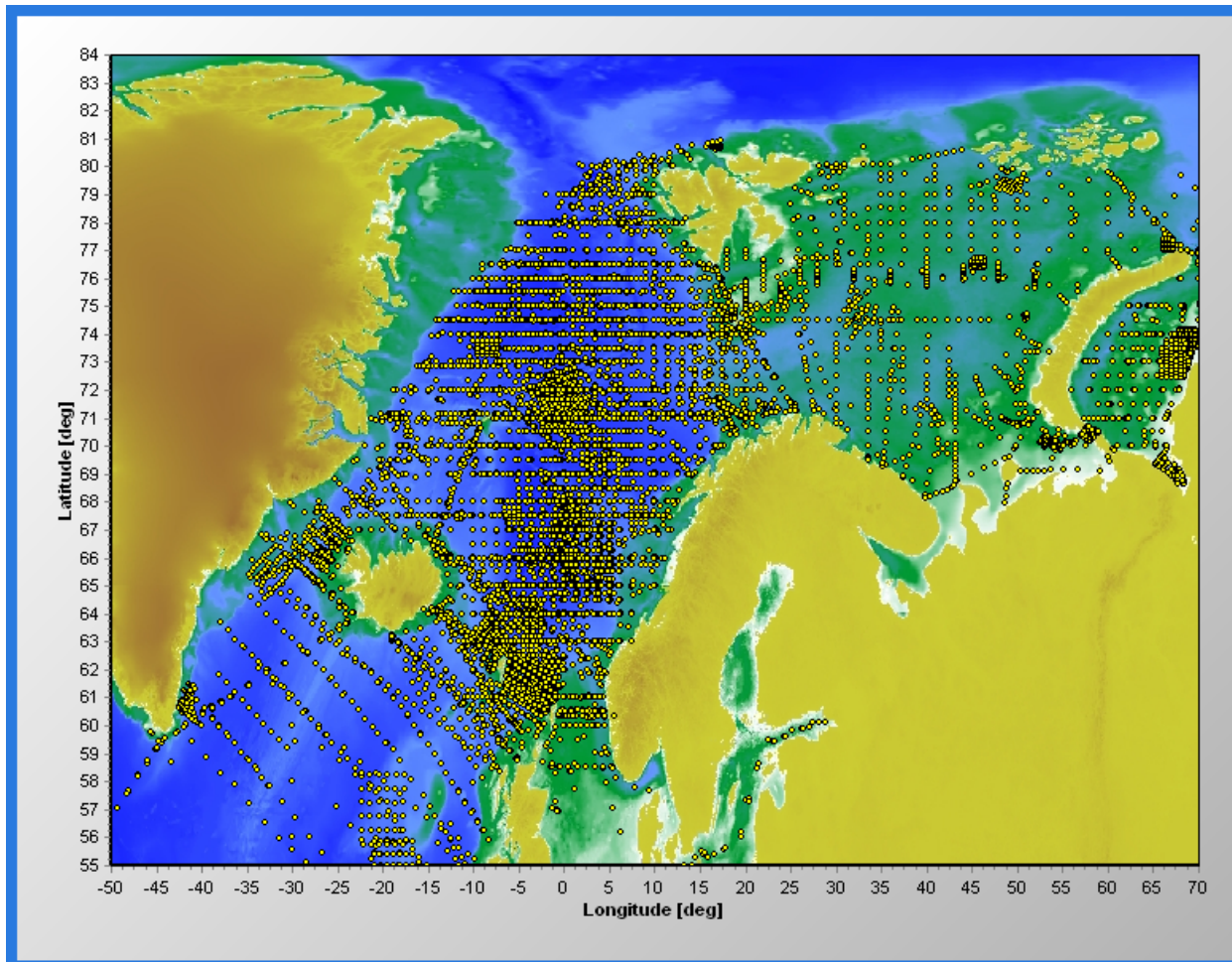


Time series length examples that can be reconstructed from OA fields (0.25x0.5 latitude-longitude grid) for 1900- 2004. Temperature at 50m depth level.



Database Updates: UNESCO/IOC/GADAR project: AARI cruise data (since September 2006).

Repeated measurements at the regular grid net



Period: 1968-93
Cruises: 113
2006->17 cruises
2007->92 cruises
2008->4 cruises
Stations: 20021

Almost 100% stations were updated

Full composition of metadata, profiles and meteorological information

'Professor Vize'
Period: 1968-90
Cruises: 27
Stations: 5088 st.

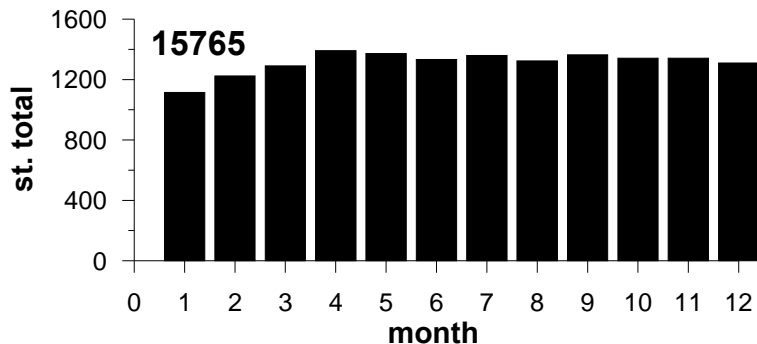
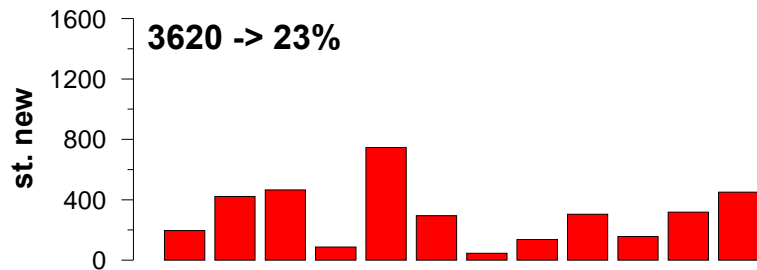
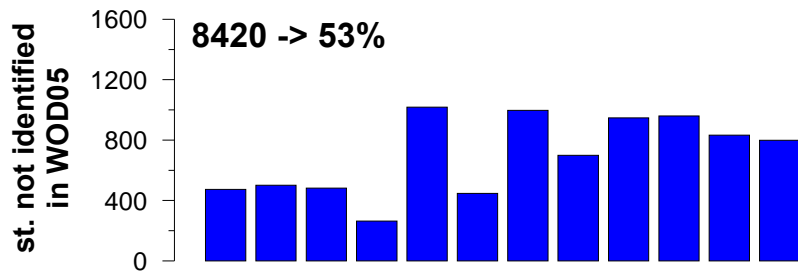
'Professor Zubov'
Period: 1969-89
Cruises: 18
Stations: 4098 st.

'Akademik Shuleykin'
Period: 1982-91
Cruises: 24
Stations: 3922 st.

'Professor Multanovskiy'
Period: 1983-93
Cruises: 27
Stations: 4340 st.

UNESCO/IOC/GODAR project: AARI cruise data. Data controlled in 2007.

AARI cruises controlled during 2007



Considerable improving of the data in WOD05:

- Metadata (coordinates, time, vessel name, depth, cruise & st. numbers)
- Profiles (especially chemistry)
- Meteorological information

Full range of supplementary materials:

- all data in OCL format
- original cruise maps, tables
- measurements details
- correction reports for each cruise

Conclusions

- **Quality-duplicate controlled oceanographic database for the Nordic Seas (NS) region was compiled with more than 400,000 stations for 1990-2006. While the database contains wide range of chemistry profiles, quality control was applied only for temperature, salinity and dissolved oxygen.**
- **Database contains high quality data (if possible to identify) to avoid instrumental and vertical biases and appropriate for temporal and spatial oceanographic characteristics study**
- **Three program applications were developed: (i) for data download, processing and quality/duplicates control (ODB3ALoad), (ii) user interface for data excess and analysis (ODB3A), (iii) processing and analysis of the OA fields (ODB3AClimate)**
- **AARI cruise data in the Nordic Seas were considerably updated in strict correspondence with hydro-meteorological tables (TGM-3M) under the UNESCO/IOC/GODAR project**