## THE KOBE COLLECTION: NEWLY DIGITIZED JAPANESE HISTORICAL SURFACE MARINE METEOROLOGICAL OBSERVATIONS

Teruko Manabe, Japan Meteorological Agency, Tokyo, Japan

#### 1. INTRODUCTION

The Kobe Marine Observatory (formerly the Imperial Marine Observatory), a field office of the Japan Meteorological Agency (JMA), collected and stored surface marine meteorological observations reported by ships in log sheets over the period from 1890 to 1961. This data set is called the Kobe Collection (Komura and Uwai, 1992; Uwai and Komura, 1992; Manabe, 1999*a*).

In the Collection, the data obtained by merchant ships, fishing boats and research vessels amount to around 6.8 million observations. Figure 1 shows annual numbers of ships and reports of these data. In addition, reports by Japanese Imperial Navy ships, which cover the period from 1903 to 1944, amount to around 5 million. Unfortunately, the annual number of reports by Navy ships were not counted and, therefore, could not be included in Figure 1.

In 1960 and 1961, log sheets of the data obtained by merchant ships, fishing boats and research vessels were copied onto microfilm (364 rolls in total) under the JMA-NOAA (National Oceanic and Atmospheric Administration) joint project. In this project all the data taken after 1933 (about 2.7 million in 185 rolls of microfilm) were digitized. These digitized data were already included in the Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1 (Slutz *et al.*, 1985). However, until recently, pre-1933 data and Navy data had not been digitized.

### DIGITIZATION PROJECTS SINCE 1995

At present, COADS is one of the most complete marine meteorological data sets. However, even COADS does not contain a large amount of observations made before the 1950s, especially during the two World Wars (Woodruff *et al.*, 1987). Furthermore, since most data are from US and European ships, which mainly have ship routes in the Atlantic Ocean, there are fewer data for the Pacific than the Atlantic. The Kobe Collection covers the period of the First World War, and the main ship routes of Japanese vessels are in the Pacific. Therefore, digitization of the Kobe Collection was one of the most urgent and important projects amongst the many anticipated projects related to data archeology and rescue.

In the fiscal year (FY) 1995, JMA began to digitize the pre-1933 surface marine observations made by merchant ships in the Kobe Collection, with with cooperation from the Japan Weather Association (JWA) and the Nippon Foundation. So far, the series of efforts can be divided into three phases: Phase I is digitization in FY



Figure 1—Yearly distribution of reports and ships in the Kobe Collection (excluding Navy data from 1890 to 1960). For the period from 1890 to 1900 annual bars were estimated by dividing by the number of years per period since only a total figure was available. 1995/96 and its quality check in FY 1997; Phase II is digitization in FY 1997/98 and its quality check in FY 1998/99; and Phase III is digitization in FY 1999/2000 and its quality check planned for 2001. In Phase I, a total of 1 045 682 reports were digitized and quality checked, and these data are now available on CD-ROM (1998 edition). In Phase II, a total of 571 472 observations were digitized in 1998/99 and quality checked, and these observations will soon be available on CD-ROM (2000 edition). Phase III was carried out from 1999 to 2000, and it is estimated that a total of around 600 000 observations will be digitized.

Because of a particular shortage of data during the First World War in the currently available COADS, the JMA devoted its efforts to digitizing the reports for the period from 1915 to 1917, which is dealt with in Phase I.

#### 3. THE DATA SOURCE AND REMARKS ABOUT THE DIGITIZATION

Figure 2 shows an example of one of the oldest log sheets, which was reported in 1890. The observational elements in the first format include: date and time (local) of the observation; ship location; wind direction and Beaufort force; air pressure; temperature indicated by a thermometer attached to a barometer; dry-bulb and wet-bulb temperatures (Fahrenheit); cloud; present weather; direction and height of wind waves; sea surface temperature (SST); specific gravity of sea surface water; direction and speed of sea surface current; and any remarks. The format of log sheets was changed several times and the biggest change was made in 1923. Following this change, columns were added to describe the type of barometer and its instrumental correction value, as well as columns for new observational elements such as the direction and height of swell. Figure 3 shows an example of a log sheet of 1930.

During the data period (1890-1933), the number of observations taken each day varied according to the ship. Before 1923, most ships made observations six times a day at 0200, 0600, 1000, 1400, 1800, 2200, and after 1923 most ships made them four times a day at 0000, 0600, 1200, 1800. Some ships made observations four times a day at 0000, 0400, 1200, 1600, while some made them three times a day at 0400, 1200, 2000 or 0800, 1200, 2000 and others twice a day at 0000, 1200. All times were local.

Often, the ship's exact location (latitude and longitude) was logged only once a day, especially in early data. In Phase I digitization, reports without latitude/longitude information were not keyed. No reports were digitized for ships which navigated relatively close to Japan and made observations mostly at port (ships were requested to make observations even at port in the early days) and/or reported latitude/longitude once a day or less (e.g. every three or seven days).

						17	× M	海 ETEO	ROLD	E BGAL	氣 肥肉	RT R	N DBISER	MED VED	15 I	IEA,			
-	۰.,	Car		-191	,		*14	me	his	A		36	2 p			enter,	2	 61	a Honsteller
F		-	-	-		1	1	-	-		-	_		5		2.	-		
-	-	-	-	8 - A	-	1. a	100	ь.		1.1	1.0	-	Tube-		a -	-	in a	 -	MALELS.
1	12.4	-		31W 1.W	4. 7.4	# 24n	14.27	44.43		1.1		0	17	50	19-14	1444			the hary
10	10000	11.34	-	RE	12		200	4248		-		#	6113	100	12	1000			Remark Coming By Clifford
"	4.11	Nº4	opri a	-	6	:5	Citra C	4440		Ξ	+ 10	44	14.		4.4.	404.1			this clean by an
10	1000	11.17	-7.1	100	4.10	: 22	144	404		+	diger -		0.4	1	2444	10			transaction in many a
10	4	47	147.00	air V TV	44.1.	115	「「	42.8				4411	uke.	411	4743				time along day have
~	100	29.75	AL IN	97	3644	17.4	* 25	24		1111	10.1	14. 10	1	-15	4.1.1	57.22			Washing mult
17	1000	11-12	107,17	1	5	15	-125	414		- 19-1	17	1.82	40	1.	1-41	4.20			Shen hall
4	Arm. 4.0 4.1	2.4	20	****	1	1244	The Real	7.14			-	43	e	111		12			and date in the second
17	Farmer 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ng m	(The second	11 1 11 11	Reve large	"PRATES	34. 194.94	122422		4-11-12	17. 18	auchan	- Action - 2	112-1122	· N. 1/33	124 420			tall clouds that a south

Figure 2—Sample of one of the oldest log sheets (1890). Yamashiro Maru left Yokohama on 9 May heading for Honolulu. The observational elements are: date and time (local): ship position (latitude and longitude); wind (direction and force); barometer (and attached thermometer reading); thermometer readings (dry-bulb and wet-bulb); clouds; weather; waves (direction and height); sea water (temperature and density); and sea surface current (direction and speed).

Viery	lage à	2. Ra	ito		a.				拆	Ŀ	9.	泉	89	古		-	.EM	ATT
16.0	ting.	sú.	10	er Sau	1.4	-	in B							a .3	-	Enie	e.	an 25
111				÷	**	-	11 2 2	n 11	10			×	-		8	* *	*	The Distance Digitizity - Man was
			مر <i>طو</i> ر.	-	0%	1	-		20-6	4.44	2							
11 8	1.5Ca EL-	n.	112	J	47.04	ŻZ.	$\overline{h'}$	ē2	00	1	e.	10	F	35	A.	4.4	2	The sector - Mar & sector
글 //	놼왌	낦	sinch 240	ᄼ	26	82	2	쉱	R4	20	17	2	3	- ideal Chi	4	-44.80	술	Two differences at last
	U.C. Pro	1	- CAL	-		54	8	셨		-10	13	-5-	5	Cal	1.	1	-	
110 A	· 24 82-	10	44.8	2	10.48	25	6	-	C1	14	12	. 2.	-E	A63		1000		an an to see -
10	N - S	7-	freed.	1	416	i.	였	ir	25	49.	位	-		diand	-	Edad.	1	and the second sec
** <u>}</u>	- as Ur .		-	-	1000	W.	24	22		2	17	~	1	-			1	
	10.00	0.1	1		-	10	臣	22			μe.	1.6.	A	1	1		1	
10 4	10. 190	삵		-44	4,9,10	i lete	in	i.			1	1	-1		-	1	4	
1	-1× 16	M.	÷	+	- 11	17	1	÷	-	48.4	19	2	1.		4		*	a colores
10 A	1 50	1	-	-	27.74	100	1	1		2.15	Te.	10	4	and an	1		1	
and S	it It -	. 0	-	+	25.15	17	4	-		100	经		1	1	1			10.000
- 4	Cia Cru	÷Ē.,	Later	-	1965	1	80	4.0			1¢	- <u>5</u>		2	-	2		-8
22	100 200	0-ei Ar -	alers.	12	10.12	100	22	5.	20	-16	1	- 10	1	-	1			8 /
1	## 45°	5	0.01	-	11.11	14	12	16		- 18	1.	20	·** .	- di M-	1		14	/A-30"-
言父	-41' Ir.	61	1	-	are	1	1			14	12					1		
	115 62		See.	2	20 18	100	61	i.	24	100	1.	. £		See.	2	100	3	A correst
-	11. 110	x.	- <u>h</u>	1	29.22	11.	문	1		1.	+-				1		-1	<u>in</u> 1
**	a tra		1	Ż	13.16	12	Te	11	84	A	Ż	3	4	1.20	1	1	•	8
100 J	10 20	25	52	2	11.14	12	#	12	14	-K-	12	- 2-	3	412	1	1000	- 2	E Discord & lachard
10	11. 121	41		12	20	Į.	22	62	09.	14	Ľ.	1	. 2		Ľ.	1		All the Roigship a start
100 Lat	-21 - 25-	5	-	÷	100	傍	22	5	dh.	N	10	4	1		-		之	18 Cash 1
- 4	100	4	Also	2	あた	18	12	16	11	-0	2	4		sin!	-		1	
2	101 011			5	11.19	1	1	$\overline{G}$	2	25	12		8	- 1		4	5	
24	TR Ink		- Ar. 44	1	34	仍	55	2.	- 64	1.00	2	4	5	100		*	2	10-pt
7.4	35 145	14	1	4	22.22	12	15	1	41	1.00	14	- 1-	1	1	-	-		1
an 1	1.11 1.11	100		1	100	12	12	5	11	Lan.e	6	- 40	5			-	1	dial"
	1.00	14		2	1917	14	34	17	15	H	-	- #	1		-		8	
-	17 13	121	5	-	116	16	15	12		el.	- 10-			1		20	1Ž	
-		- 7.4		17	1.46	t.		1.41	101-10.0	1.00	A.F.	·	44	1 100	11.5	-8414	e-/	Rissor 1 day

This is the main reason why a lot of data, especially those prior to 1901, were not entered into the final digitized data set. Also, many reports in the early days were not keyed because it was difficult to read the handwriting. No reports from the the log sheets of 1891, 1895, 1896 and 1897 were digitized in the Phase I project.

The digitization procedures are as follows. First, the microfilms to be digitized were selected and printed out to produce hard-copies of the original log sheets. Besides the period from 1915 to 1917, microfilms were selected evenly for the whole period (1890–1932) so that various types of log sheets can be seen. Secondly, some elements, such as weather and units of temperature and air pressure, were coded before the keying process was initiated. After coding the information from the log sheets, the data on the coded log sheets were keyed to make an interim-file format designed to include nearly all the information contained in the log sheets. The interim-file format retains the original values (e.g. 32-point scale directions, Beaufort numbers, temperatures in Fahrenheit, weather and visibility). Finally, the digitized data in the interim file were converted into the International Maritime Meteorological Tape format, Version 1 (IMMT-1) (WMO, 1990), which does not retain the original values, but it can be easily handled and is widely distributed.

Then, quality controls were conducted on the digitized data in accordance with the minimum quality control standards of the WMO Marine Climatological Summaries Scheme (MCSS) (WMO, 1990). Each ship's track was examined by checking the ship's speed and land/sea information in the global ocean. Furthermore, in the North Pacific, air temperature, SST and dew point temperature were compared with JMA climatology (JMA, 1993*a*). No data comparisons with climatology were made in other ocean basins. Approximately 5 per cent of all the checked data were manually corrected. The errors appeared to be mainly caused by misinterpretations of the handwritten logs.

Specific remarks on several observational elements can be summarized as follows.

#### 3.1 SHIP TIME AND LOCATION

Reports without location were not keyed. The date and location were checked by examining each ship's track and were manually corrected by referring to the original log sheets, where possible.

via the Panama Canal as part of a round the world trip. Observational elements are: date (date and time); location (latitude and longitude); wind (direction and force); barometer (corrected reading and attached thermometer reading); temperature(air-temperature, sea surface temperature); clouds; weather; visibility; wind waves (direction and height); swell (direction and height). There are five questions at the bottom of the log sheet: (1) Is the barometer mercury or aneroid? (2) Date of last check of the barometer? (3) Value of instrumental correction? (4) Height of barometer? (5) Was the instrumental correction

applied?

Figure 3— Log sheet of 1930. Buenos Aires Maru was en route from Galveston to Los Angeles SEA SURFACE TEMPERATURE

3.2

It has often been pointed out that instrumental bias is quite influential when dealing with historical SST observations (e.g. Folland and Parker, 1995). It was assumed that all the observations were obtained using buckets since the data are pre-1933 and engine room intake measurements appeared for the fist time in the *Guide to Weather Observations for Ships*, 1956 edition (JMA, 1956). However, the material (e.g. canvas, wood or rubber) of each bucket cannot be identified because, unfortunately, there is no remaining documentation on the SST measurement method used on each ship.

The readings were written in Fahrenheit or Celsius in the log sheets, and observations in Fahrenheit were converted into Celsius (Appendix A).

3.3 AIR TEMPERATURE AND DEW POINT TEMPERATURE Ships were requested to report dry-bulb and wet-bulb temperatures. The readings were made in either Fahrenheit or Celsius, and all the observations were converted into Celsius in the IMMT format (Appendix A).

Provided that both dry-bulb and wet-bulb temperatures were reported and the reported values were consistent, the dew point temperature was calculated when converting data into the final IMMT format (Appendix B).

3.4 Air pressure was measured using aneroid or mercury barometers. The readings AIR PRESSURE were reported in inchHg or mmHg and all observations were converted into hPa in the final format (Appendix C). When mercury barometers were used, the reading of the attached thermometer was also written in the log sheets. It is assumed that temperature, gravity and scale corrections were applied before reporting, in accordance with the *Manual on the Marine Meteorological Observation* (Imperial Marine Observatory, 1921).

In the conversion process from the interim-file format to the final IMMT format, height corrections were added when a height was written in the log sheet (Appendix C).

3.5 Wind direction was reported and keyed in 32-point scale and was then converted AIR PRESSURE into 36-point scale according to Table 1. Wind speed estimated visually was reported based on the old Beaufort scale as shown in the *Manual on the Marine Meteorological Observation* (Imperial Marine Observatory, 1921). Each scale was converted into knot according to Table 2.

3.6 W WIND WAVE AND SWELL in

Wave direction and wave height were reported. Direction was reported and keyed in 32-point scale and then converted into 36-point scale (Table 1). The height of wind waves and swell height were reported according to the JMA wind wave scale and the JMA swell scale, respectively. The wind wave scale and swell scale were converted into units of 0.5 m according to Tables 3 and 4, respectively.

# Table 1—Conversion of 32-point scale to 36-point scale.

32	-point scale	3	6-point scale	3.	2-point scale	36-point scale		
Code	Description	Code	Description	Code	Description	Code	Description	
00	Calm	00	Calm	17	S by W	19	185°-194°	
01	N by E	01	5°-14°	18	SSW	20	195°-204°	
02	NNE	02	15°-24°	19	SW by S	21	205°-214°	
03	NE by N	03	25°-34°	20	SW	23	225°-234°	
04	NE	05	45°-54°	21	SW by W	24	235°-244°	
05	NE by E	06	55°-64°	22	WSW	25	245°-254°	
06	ENE	07	65°-74°	23	W by S	26	255°-264°	
07	E by N	08	75°-84°	24	W	27	265°-274°	
08	E	09	85°-94°	25	W by N	28	275°-284°	
09	E by S	10	95°-104°	26	WNW	29	285°-294°	
10	ESE	11	105°-114°	27	NW by W	30	295°-304°	
11	SE by E	12	115°-124°	28	NW	32	315°-324°	
12	SE	14	135°-144°	29	NW by N	33	325°-334°	
13	SE by S	15	145°-154°	30	NNW	34	335°-344°	
14	SSE	16	155°-164°	31	N by W	35	345°-354°	
15	S by E	17	165°-174°	32	N	36	355°-4°	
16	S	18	175°-184°	99	Unknown	99	Unknown	

SECTION 1	1 —	MARINE	DATABASE	ENHANCEMENTS
-----------	-----	--------	----------	--------------

Table 2—Conversion of Beaufort		Beaufort scale		knot
scale into knot	Beaufort	Description	Wind speed	Wind speed
Scale into mot.	number	term	(metres per second)	(knot)
	00	Calm	=<0.3	00
	01	Light air	0.3 - 1.5	02
	02	Slight breeze	1.6 - 3.3	05
	03	Gentle breeze	3.4 - 5.4	08
	04	Moderate breeze	5.5 - 7.9	13
	05	Fresh breeze	8.0 - 10.7	18
	06	Strong breeze	10.8 - 13.8	24
	07	High wind	13.9 - 17.1	30
	08	Gale	17.2 - 20.7	37
	09	Strong gale	20.8 - 24.4	44
	10	Whole gale	24.5 - 28.4	51
	11	Storm	285-335	59
	12	Hurricane	33.6=<	68
Table 3—Conversion of wave	IMA wind wave	IMA wind wave scale	Equivalent wave	IMMT
height in JMA wind wave scale into units used in IMMT.	scale number	Description	height (feet)	Wave height (units of 0.5 metre)
	0	Dead calm	0	00
	1	Very smooth	<1	00
	2	Smooth	1-2 (1 < <2)	01
	3	Slight	2-3	02
	4	Moderate	3-5	02
	5	Pather rough	55	04
	5	Rather Tough	0 1 2	04
	0	Liab	12 20	10
	/		12-20	10
	8	very nign	20-40	18
	9	Phenomenal	40≤	24
Table 4—Conversion of wave height in JMA swell scale into units used in IMMT	JMA swell scale number	JMA swell scale Description	Equivalent height (feet)	IMMT Height of swell (units of 0.5 metre)
	0	Nia avvall	0	
	0	NO SWEII		00
		Slight swell	1-3 (1=< <3)	01
	2	Moderate swell	3-5	02
	3	Rather rough	5-8	04
	4	Rough swell	8-12	06
	5	Heavy swell	12-20	10
	6	Very heavy swell	20-40	18
	7	Abnormal swell	40=<	24
able 5—Conversion from clouds	Clouds in tenth	Clouds in Oktas		
in tenth into Obtas		0		
in tenth into Oktas.	0	0		
in tenth into Oktas.	0 1	1		
in tenth into Oktas.	0 1 2	1 2		
in tenth into Oktas.	0 1 2 3	1 2 2		
in tenth into Oktas.	0 1 2 3 4	1 2 2 3		
in tenth into Oktas.	0 1 2 3 4 5	1 2 2 3 4		
in tenth into Oktas.	0 1 2 3 4 5 6	1 2 2 3 4 5		
in tenth into Oktas.	0 1 2 3 4 5 6 7	1 2 2 3 4 5 6		
in tenth into Oktas.	0 1 2 3 4 5 6 7 8	1 2 2 3 4 5 6 6		
in tenth into Oktas.	0 1 2 3 4 5 6 7 8 9	1 2 2 3 4 5 6 6 7		
in tenth into Oktas.	0 1 2 3 4 5 6 7 8 9 10	1 2 2 3 4 5 6 6 7 8		

3.7 In accordance with the conversion table based on present and old manualsCLOUD (WMO, 1995; Imperial Marine Observatory, 1921), total cloud amount in tenths was converted into Oktas according to Table 5.

3.8 WEATHER AND VISIBILITY These elements are not included in the data on the CD-ROM (1998 edition).

#### 3.9 SHIP IDENTIFIER

### TEMPORAL AND SPATIAL DISTRIBUTION OF THE DIGITIZED DATA

Since we did not have the 'call signs' used by ships during the data collection period, each ship was allocated a 'ship number'. The 'Ship number' has five digits; the first two digits correspond to the last two numbers of the year (e.g. 1910 corresponds to 10xxx); the last three digits correspond to the ship's position in the alphabetical listing of the names of every ship that reported data during that particular year. 'Ship numbers' and 'Ship names' are catalogued in the *Guide Book of the Japanese Marine Surface Data* (United States Weather Bureau and Japan Meteorological Agency, 1960).

Figure 4 shows the geographical distribution of the data available on the CD-ROM (1998 edition). The data are mainly distributed in the North Pacific, especially along the main ship routes: Japan-northern America, Japan-Hawaii-California, and so forth. For all the digitized reports, 82.8 per cent, 11.5 per cent and 5.7 per cent are in the Pacific, Indian and Atlantic Oceans, respectively (each basin corresponds to the information contained in Figure 3.2 of Slutz *et al.*, 1985).

Figures 5 (a) and 5 (b) show the yearly distribution of the data of the Kobe Collection, excluding the Japanese Imperial Navy data, and COADS in the global ocean and the Pacific Ocean, where a large part of the Kobe Collection exists, respectively. In the Pacific Ocean, the presently available COADS, shown by dotted and light hatched areas, has a significant jump in the amount of data between 1932 and 1933. This is because COADS already includes the Kobe Collection data (1933–1961) which were digitized until 1961. The newly digitized Kobe Collection data (1890–1932) significantly increased the amount of available data, especially in the Pacific Ocean. It was also discovered that the amount of data covering the First World War was greatly increased by adding the newly digitized Kobe Collection data to COADS.

To show the effectiveness of the newly digitized data, a preliminary analysis was carried out (Manabe, 1999*b*). By using the newly digitized data along with the presently available COADS,  $2^{\circ} \times 2^{\circ}$  monthly, seasonal and annual SST anomalies were calculated. In the North Pacific, data coverage (per cent of the number of grid boxes with data) increased from 5-40 per cent to 20-60 per cent from 1910 to 1933 compared with those made from COADS alone. Thanks to the increase of grid boxes with data, it was possible to apply empirical orthogonal function (EOF) analysis to data from before the Second World War as well as that from after the war. This showed that the Pacific decadal oscillation founded by Tanimoto *et al.*, 1993 can be observed back to the beginning of this century.

#### 5. FUTURE PLAN ON DIGITIZATION

During the series of digitization projects supported by the Nippon Foundation, more than one million marine meteorological observations taken for the period from 1890 to 1932 in the global ocean, especially in the North Pacific, have been



Figure 4—Geographical distribution of reports for the whole data period (from 1890 to 1932) available on the CD-ROM (1998 edition). Each 2°× 2° box is shaded according to the number of reports. The total number of reports appearing on this figure is 1 045 682. SECTION 1 — MARINE DATABASE ENHANCEMENTS



Figure 5—Yearly distribution of the reports in the Kobe Collection, excluding the Japanese Imperial Navy data, and COADS in the global ocean (a) and in the Pacific (b). The CD-ROM (1998 edition) contains the Kobe digitized in 1995/96, which is shown in black. The presently available COADS is shown by "COADS, excluding the Kobe Collection' and "Kobe digitized in 1960/61". Because the data digitized in FY 97/98 have not been quality checked, the number of these data have not been fixed.

Digitization projects

5.1

newly made available. The biggest contribution of the digitization projects has been to increase the usable marine meteorological observational data in the North Pacific for the period around the First World War. It is expected that the newly digitized Kobe Collection will be widely used.

Following the publication of the CD-ROM (1998 edition) in 1999, a CD-ROM (2000 edition), which contains a total of 5 71 472 observations digitized in Phase II, will be published at the end of FY 2001. Phase III digitization was carried out in FY 1999 and FY 2000. It is expected that around 600 000 observations will be made available in Phase III. However, over 1 million records still needed to be digitized at the end of FY 2000. In cooperation with JWA, the JMA is making efforts to continue to digitize as many data as possible in the period following 2000. To make more historical data available, it is expected that Phase IV of this series of digitization projects will start in 2001.

Also, JMA is exploring the possibility of increasing the number of observations in the data set by including many of the 'no-position' reports, for which a ship's position can be estimated by interpolation. This could substantially increase the number of observations in some regions.

5.2 F DATA FORMAT FOR HISTORICAL MARINE DATA f

For the distribution of the digitized Kobe Collection in Phase I, the IMMT format was adopted since it is easy to handle. However, because this format is designed for the storage and exchange of contemporary marine data (from 1961 onwards) in WMO, it is not well suited to historical data. For example, there are no columns for the thermometer attached to the barometer, the original Beaufort number, or the original units of data. Furthermore, as regards wind and wave direction, a 36-point scale is used in the IMMT, whereas a 32-point scale was often used in historical data, and the conversion from 32-point to 36-point scale could cause problems. It would be very helpful if a data format could be agreed upon that is well suited to historical and modern data and is easy to handle.

In WMO, it is recognized that while efforts have intensified to digitize the additional historical ship data that exist in many national log book collections, such as the Kobe Collection, there is no effective internationally agreed format for the exchange of keyed historical data. Efforts are being made to develop an International Marine Meteorological Archive (IMMA) format that is well suited to historical and modern data and is easy to handle (WMO, 2000).

Once the series of digitization projects are complete, all the digitized data are planned to be made available in a new format (IMMA, if possible) which is more suitable for historical data than the IMMT.

5.3 With regard to the Japanese Imperial Navy data which covers the period from 1903 to
 JAPANESE IMPERIAL NAVY
 DATA
 DATA
 1944, according to a preliminary investigation on the data, it seems that only about 10
 per cent of all the reports (about 5 million reports) include location information.
 However, considering that these data cover the data sparse period, which includes the
 two World Wars, digitizing the navy data would be a valuable exercise. Thus, JMA is
 trying to find a way to rescue these data which will make it possible to interpolate
 missing location data so that as many data as possible will become available.

#### ACKNOWLEDGEMENT

K. Nishiyama and JMA staff assisted in many ways. The author wishes to express her thanks to all these people, especially O. Shimada, A. Shoji, M. Kaneda, Y. Yokote and A. Wada.

The digitization projects have been supported by the Nippon Foundation, with the cooperation of JWA. R. Yamamoto, K. Hanawa, K. Kutsuwada and Y. Matsuyama were instrumental in the promotion of the project.

With the kind permission of the American Meteorological Society, portions of this manuscript have been reprinted from the "Digitized Kobe Collection, Phase I: Historical Surface Marine Meteorological Observations in the Archive of the Japan Meteorological Agency", which appeared in the Bulletin of the American Meteorological Society, No.80.

APPENDICES CONVERSION ALGORITHMS A. Conversion of temperature Temperatures observed in Fahrenheit were converted into Celsius using the following equation:

#### Tc=(Tf-32)/1.8

where Tc is the temperature in Celsius, and Tf is the temperature in Fahrenheit.

B. Calculation of dew point temperature When dry-bulb and wet-bulb temperatures and air pressure (station pressure) were known, dew point temperature was calculated in accordance with the *Guide to Surface Meteorological Observations* (JMA, 1993*b*). First, saturation vapour pressure at the wet-bulb temperature was obtained using the Goff-Gratch formulae described in the WMO *Technical Regulations* (WMO, 1988). Secondly, vapour pressure at the observed wet-bulb and dry-bulb temperatures and air pressure were calculated using the following Sprung formula using the obtained saturate vapour pressure:

#### e= E'- P(T-Tw)A/755

where A=0.5 when the dry-bulb is not iced, A=0.44 when the dry-bulb is iced, e is vapour pressure in hPa, E' is saturated vapour pressure in hPa at the wet-bulb temperature, T is the dry-bulb temperature in Celsius, and Tw is the wet-bulb temperature in Celsius.

Finally, dew-point temperature is extracted from a table on the JMA *Surface Meteorological Tables* (JMA, 1959 (see 1986 edition for amendments)) produced on the basis of the above-mentioned Goff-Gratch formulae.

C. Conversion of air pressure

Air pressure was observed either in mmHg or inchHg. In the interim-file, barometer readings in mmHg or inchHg remain. The interim-file format has a column to show the units of air pressure (mmHg or inchHg). Air pressure in mmHg or inchHg was converted into hPa and a height correction was also made during the conversion process. Thus, air pressure at sea level in hPa was written in the IMMT format. However, if a height was not written in the log sheet, the pressure values that were simply converted into hPa were written in the IMMT format without the height correction. Readings were converted into hPa in accordance with the following equations and then rounded off to the closest hPa:

when air pressure was observed in inchHg:

 $P=b \times 33.8639+1013.25 \times (exp(1.17972 \times 10-4 \times H)-1)$ 

when air pressure was observed in mmHg:

 $P=b\times 1.33322+1013.25 \times (exp(1.17972 \times 10-4 \times H)-1)$ 

where P is air pressure at sea level (hPa), b is the barometer reading, and H is height of the instrument (metres).

- REFERENCES Folland, C.K. and D.E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. *Quart. J. Roy. Meteor. Soc.*, 121, 319-367.
  - Imperial Marine Observatory, 1921: Manual on the Marine Meteorological Observation (in Japanese). Imperial Marine Observatory, 82 pp.
  - Japan Meteorological Agency, 1956: *Guide to Weather Observations for Ships* (in Japanese), Marine Department, Japan Meteorological Agency, 119 pp.
  - Japan Meteorological Agency, 1959: *JMA Surface Meteorological Tables* (in Japanese), Japan Meteorological Agency, 144 pp. (eighth amendment was made in 1986)
  - Japan Meteorological Agency, 1993a: Marine Climatological Charts of the North Pacific Ocean for a Thirty-Year Period (1961–1990). Marine Department, Japan Meteorological Agency, 191 pp.
  - Japan Meteorological Agency, 1993*b*: *Guide to Surface Meteorological Observations* (in Japanese), Japan Meteorological Agency, 167 pp. (fourth amendment was made in 1999)
  - Komura, K. and T. Uwai, 1992: The collection of historical ships' data in Kobe Marine Observatory. *Bull. Kobe. Mar. Obs.*, 211, 19-29.
  - Manabe, T., 1999a: Digitization of the Kobe Collection: The historical surface marine meteorological data collected by the Kobe Marine Observatory, Japan Meteorological Agency. *Proceedings of the International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata* (Toledo, Spain). H.F. Diaz and S.D. Woodruff, Eds., WMO-Technical Document No. 957.
  - Manabe, T., 1999*b*: The Digitized Kobe Collection, Phase I: Historical surface marine meteorological observations in the archive of the Japan Meteorological Agency. *Bull. Amer. Meteor. Soc.*, 80, 2703-2715.
  - Sultz, R.J, S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms, 1985: *Comprehensive Ocean-Atmosphere Data Set (COADS);* Release 1. NOAA Environmental Research Laboratories, Climate Research Program, Boulder, Colorado, 268 pp. (NTIS PB86-105723)
  - Tanimoto, Y., N. Iwasaka, K. Hanawa and Y. Toba, 1993: Characteristic variations of sea surface temperature with multiple time scales in the North Pacific. *J. Climate*, 6, 1153-1160.
  - United States Weather Bureau and Japan Meteorological Agency, 1960: *Guide Book* of the Japanese Marine Surface Data, United States Weather Bureau and Japan Meteorological Agency, 527 pp.
  - Uwai, T. and K. Komura, 1992: The collection of historical ships' data in Kobe Marine Observatory. *Proceedings of the International COADS Workshop* (Boulder), H.F. Diaz, K. Wolter and S.D. Woodruff, (eds.), NOAA Environmental Research Laboratories.
  - Woodruff, S.D, R.J. Slutz, R.L. Jenne and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. Bull. Amer. Meteor. Soc., 68, 1239-1250.
  - World Meteorological Organization, 1988: *Technical Regulations*, Volume I, General meteorological standards and recommended practices, WMO-No.49.
  - World Meteorological Organization, 1990: Manual on Marine Meteorological Services, Volume I, Global aspects, WMO-No.558.
  - World Meteorological Organization, 1995: *Manual on Codes* International Codes, Volume I.1, Part A: Alphanumeric codes, WMO-No.306.
  - World Meteorological Organization, 2000: Final Report, Subgroup on Marine Climatology, Eighth Session, JCOMM Meeting Report No. 2.