



# ARKTIEF

## - Bericht zur Datenaufarbeitung –

### 1. Einleitung

Ziel des Auftrages war es, die Parasounddaten von drei ARKTIEF-Reisen zu bearbeiten. Die Daten sollten wie in unserem Angebot beschrieben mehrere Stufen der Datenbearbeitung durchlaufen, so dass im Ergebnis:

- die Daten in lesbarer digitaler Form vorliegen (d.h., problemlos mit Software speziell für PS3-Dateien - SENT – zu bearbeiten sind)
- die Daten in Standard Seismik Formaten vorliegen (d. h. SEG-Y)
- zu den Daten Metainformationen vorliegen und abrufbar sind (d. h. Import in PANGAEA)
- zu den Daten Übersichtsdarstellungen ebenfalls über PANGAEA verfügbar gemacht werden können (GIF Dateien)

Mit einigem Mehraufwand konnten die geforderten Leistungen geliefert werden. Im Folgenden werden aufgetretene Probleme, sowie deren Lösung beschrieben. Des Weiteren sind für alle bearbeiteten Fahrten die Fehlerstatistiken etc. gegeben.

### 2. Zur Datengrundlage (ps3)

Die Parasounddaten lagen auf CD vor. Die Expeditionen ARKXVII und ARKXVIII waren schon direkt auf dem Schiff auf CD gebrannt worden, die Daten der Expedition ARKXVI waren von Tape kopiert. Es ist zu bemerken, dass die älteren Daten eine deutlich bessere Navigation und auch weniger korrupte Zeiteinträge in den Headern aufweisen. Einzelheiten dazu finden sich in den Dokumentationen zu den einzelnen Expeditionen weiter unten. Bedingt durch die Abspeicherung der originären Daten gibt es höhere Datenausfälle bei den neueren Daten. Das liegt daran, dass nicht mehr direkt die Spuren auf Tape geschrieben werden, sondern schon bei der Datenerfassung PS3-Dateien erzeugt werden. Kommt es während des Betriebes der Anlage zu einem Absturz, wird die jeweils geöffnete Datei nicht abgeschlossen und schon erfasste Daten gehen verloren. Dieses äußert sich in PS3-Dateien der Größe 0. Bei den älteren Daten existieren keine leeren Dateien. Dafür traten bei den älteren CDs andere Fehler auf:

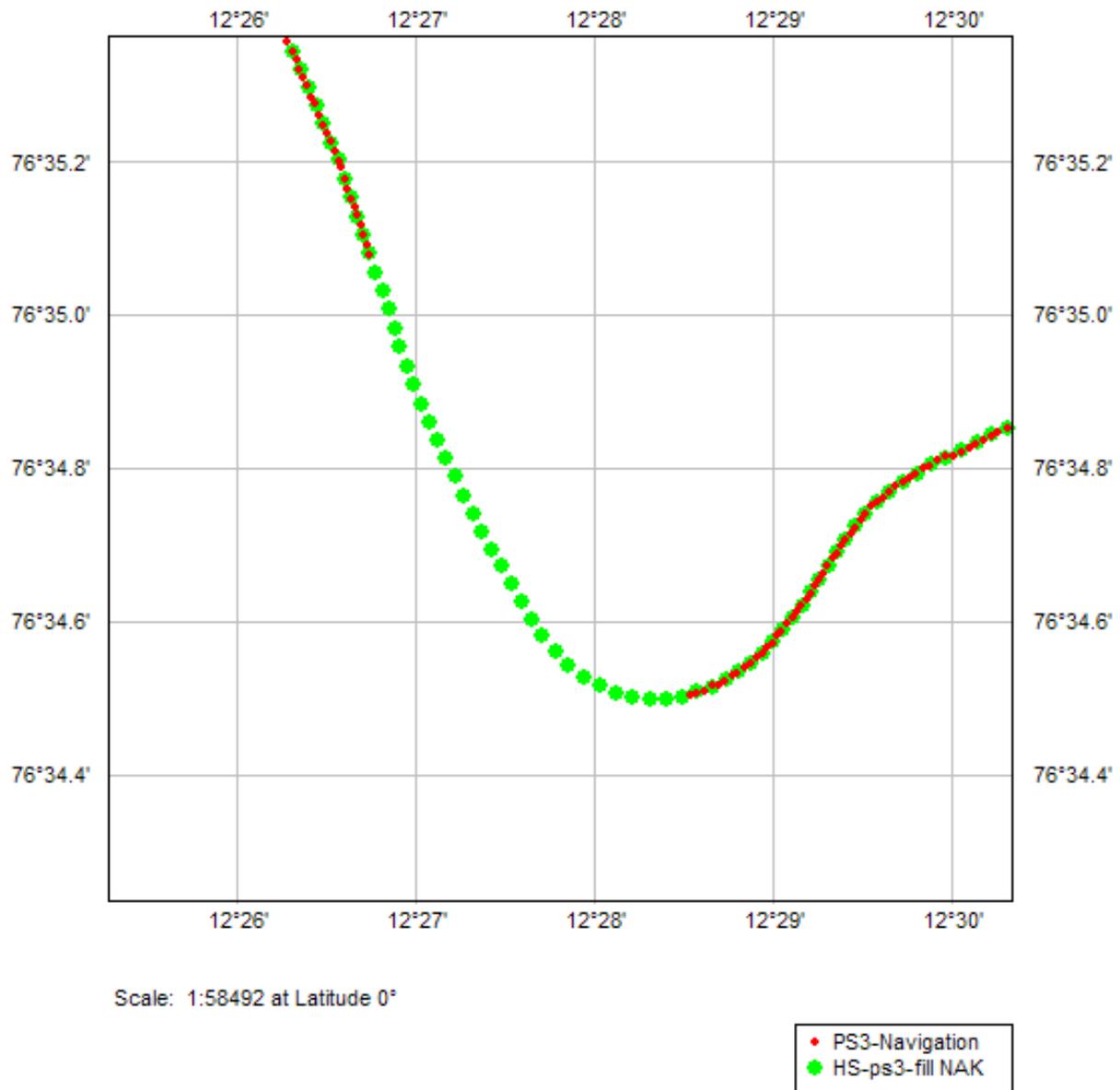
- CDs waren teilweise nicht lesbar (etwa jede achte, d.h. 15% der Daten), die Daten mussten erneut von Tape gelesen werden
- Bei jedem Tape ist der erste erzeugte Dateiname falsch. Diese mussten umbenannt werden
- Es kommen – im Gegensatz zu den neueren Daten, wo wir dieses nicht beobachtet haben – häufig leere Spuren vor. Die Dateigröße ist normal, jedoch steht in allen Spuren ‚0‘. Dieses haben wir jeweils in den für Pangaea erzeugten Metafilen kommentiert.

### 3. Zur Datengrundlage (Navigation)

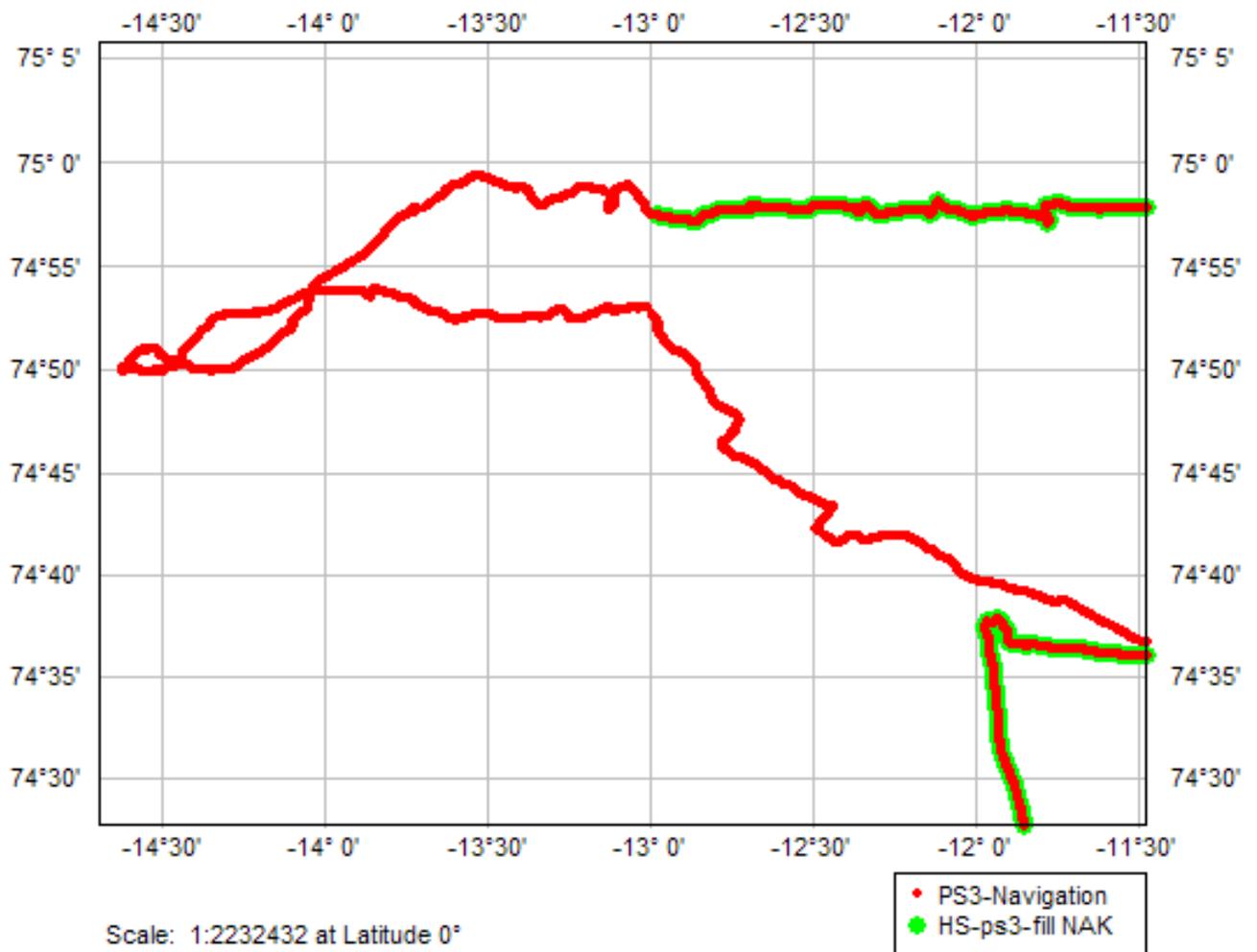
Für die Navigationskorrektur der Daten war vorgesehen, die als beste Navigationsdaten gehandelten ‚NAK-Dateien‘ der Bathymetrie zu nutzen. Leider haben wir festgestellt, dass diese nicht immer vollständig sind. Das liegt teilweise an technischen Störungen, teilweise wird aber auch die Hydrosweep-Speicherung ausgeschaltet, wenn sich das Schiff länger auf Station befindet. Des Weiteren beginnt die Navigationsspeicherung von Hydrosweep nicht immer gleichzeitig mit Parasound, auch die Endzeiten können leicht unterschiedlich sein. Zusammengefasst ist es so, dass nicht immer der gesamte Zeitrahmen, in dem Parasounddaten vorliegen, auch mit korrigierter Navigation überdeckt wird. Andererseits kommt es durchaus vor, dass das 5s-Telegramm nicht über den gesamten Zeitraum an Parasound geliefert wurde, so dass auch in den Parasound-Headern nicht die vollständige Navigation gespeichert ist. Zusätzlich zu den Navigationsdaten von Parasound und Hydrosweep sind unter <http://podas.awi-bremerhaven.de> die Schiffspositionen vieler der vergangenen Fahrten abgespeichert. Die Migration der Altdaten (PODEV-System) ist noch nicht abgeschlossen, so dass die Navigation bei den älteren Fahrten (betrifft auch schon die Reise ARKXVI) nicht ohne weiteres verfügbar ist. Zusammenfassend eine Auflistung der vorhandenen Navigationsdaten mit Problembeschreibung:



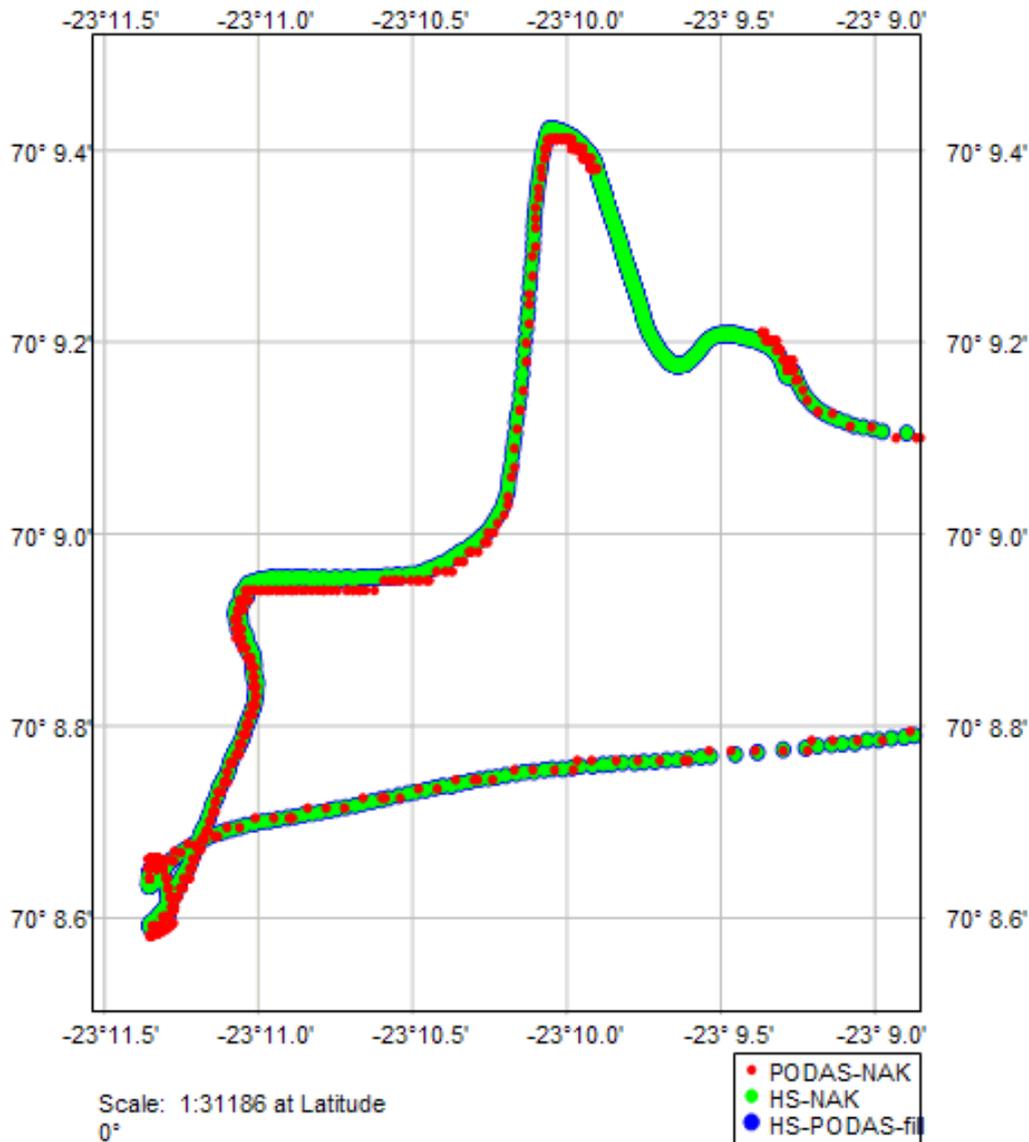
- **Parasoundnavigation:** Im Prinzip immer verfügbar, wenn Parasound läuft, da diese direkt in die Header geschrieben wird. Probleme treten auf, wenn das 5s-Telegramm nicht zur Verfügung stehen oder falsche Koordinaten in die Header geschrieben werden (Abbildung 1).
- **Hydrosweepnavigation:** Immer verfügbar, wenn Hydrosweep-Speicherung läuft. Lücken treten auf, wenn technische Fehler auftreten oder während Stationen abgeschaltet wird (Abbildung 2).
- **PODAS-Navigation:** Sollte eigentlich immer vorhanden sein. Allerdings gibt es auch hier zuweilen technische Ausfälle. Bei alten Fahrten können Probleme wegen noch nicht abgeschlossener / geprüfter Datenmigration vom Vorgängersystem PODEV auftreten (Abbildung 3).



**Abbildung 1: Ausschnitt aus der Navigation von ARKXVI. Rote Punkte sind aus den PS3-Headern. In diesem Fall kam über einen Zeitraum von ca. 5 Minuten keine neue Koordinate in die Header (vermutlich fehlendes 5s-Telegramm).**



**Abbildung 2: Ausschnitt aus der Navigation ARKXVI. In diesem Fall fehlen über ca. 22 Stunden Hydrosweep-Informationen (technischer Ausfall?).**



**Abbildung 3: Ausschnitt aus der Navigation ARKXVIII/1a. In diesem Beispiel fehlen über ca. 90 Minuten PODAS-Daten (technischer Ausfall?)**

Nach Rücksprache mit den Beteiligten am AWI wurde beschlossen, folgende Hierarchie für die Navigationserstellung einzuführen: Wenn Hydrosweepdaten vorhanden, werden diese genommen. Etwaige Lücken werden mit PODAS-Daten aufgefüllt. Sollten auch die PODAS-Daten nicht vorhanden sein, so wird die Navigation aus den Parasound-Headern extrahiert. Für die PODAS- sowie für die Parasound-Headerdaten haben wir Programme geschrieben, um Navigationsdateien im ‚NAK-Format‘ zu erstellen. Mit der beschriebenen Vorgehensweise ist es zwar immer noch möglich, aber sehr unwahrscheinlich, dass noch Lücken übrig bleiben. Bei den bearbeiteten ARKTIEF-Daten konnten so alle Lücken gefüllt werden. Mit dem erstellten Navigationsdatensatz wird dann die Navigationskorrektur in den originalen Parasounddaten (PS3-Dateien) vorgenommen. Da sich die Erstellung der Navigationsdaten im Einzelnen doch als recht aufwendig erwiesen hat und jeweils spezielle Lösungen gefunden werden mussten, haben wir dieses zu den entsprechenden Fahrten dokumentiert.



## 4. Zum Jahreswechsel / Year2000

Die Parasounddaten haben für alle (von uns untersuchten) Header im Jahr 2000 falsche Einträge in den Headern. Statt ,00' steht hier ,1900'. Dieses Problem macht eine Weiterverarbeitung (u. A. auch mit SENT) unmöglich. Daher haben wir in einer Vorkorrektur ein Programm angewendet, welches die falschen Headereinträge (Zahl und Format) korrigiert. Des Weiteren zu beachten, dass die Dateinamen (PS3) nicht eindeutig sind, da die Jahreszahl nicht im Dateinamen steht. Wir behandeln alle Daten nach Jahren bzw. Reisen getrennt und auch im Vorschlag zur Datenarchivierung müssen die Jahre getrennt aufbewahrt werden.

## 5. Grafik-Darstellungen mit SENT

### a) Allgemeines und gewählte Parameter

Das Programm SENT ist sehr gut geeignet zum Darstellen der PS3-Files. Für SENT haben wir zusammen mit Gerd Kuhn einen Konsens über vernünftige Einstellungen für standardisierte Profil-Plots erarbeitet, die in einer Muster-Parameter-Datei "ARKTIEF\_defaults.sed" abgelegt sind (zu beziehen über [usbeck@fielax.de](mailto:usbeck@fielax.de)). Bei dieser Gelegenheit stellten wir fest, dass es Inkompatibilitäten zwischen der älteren (Dez.2000) und neuesten (Update-Version, 04/2002) des Programms SENT gibt. Dies zeigt sich folgendermaßen:

- die ältere Version liest zwar Parameter-Dateien ("\*.SED"), die mit der neueren Version erstellt worden sind, ohne Fehlermeldung ein, übernimmt aber nicht alle Einstellungen, sondern verwendet irgendwelche Defaults. Vorsicht! Dies fällt im ersten Moment kaum auf und führte gelegentlich zu sehr großen Verwirrungen!
- die Update-Version bringt eine Fehlermeldung und kann die SED-Dateien, die mit älteren Versionen erzeugt wurden, überhaupt nicht einlesen.

Man sollte **alle älteren Versionen von SENT grundsätzlich NICHT mehr verwenden**, sondern nur noch die neueste Version (Stand 02.04.2002). Trotzdem sollte man sich nach jedem Einlesen der Parameter-Datei vergewissern, dass auch wirklich alle gewünschten Optionen eingestellt sind.

### b) Fehler in der Distanz-Berechnung und Darstellung

SENT berechnet aus den Headern Lon und Lat die Distanzen zwischen den Spuren und plottet die Spuren nach diesen Distanzen auf der horizontalen Achse. Leider ist die Berechnung der Distanz in SENT falsch und es werden zu große Distanzen ermittelt (Faktor ~ 1.6 bei 62 Grad nördl. Breite)). Der Fehler ist vermutlich breitenabhängig ( $\cos(\text{Lat}) = \text{vergessen?}$ ).

Damit ist die km-Annotierung der horizontalen Achse in der Grafik Darstellung falsch und sollte weggelassen werden. Die Tickmarks an der Achse sind dann auch keine Kilometer-Abstände; diese sollte aber in der Grafik belassen werden. Falls das Profil nicht zu lang ist, können diese Tickmarks immerhin als Markierungen für relativ gleiche Abstände genommen werden. Bei langen Profilen, die viele Breitengrade überfahren, ist dies nicht mehr gegeben.

Beispiel (vergl. Abbildung 4):

- Profil PS02\_07011837, Zeit von 18:37 bis 19:46 (69 min)
- wahre Distanz 18.451 km, von SENT berechnet und dargestellt 30.375 km (Lat ~ 64.86°, Lon ~ -33.28°)

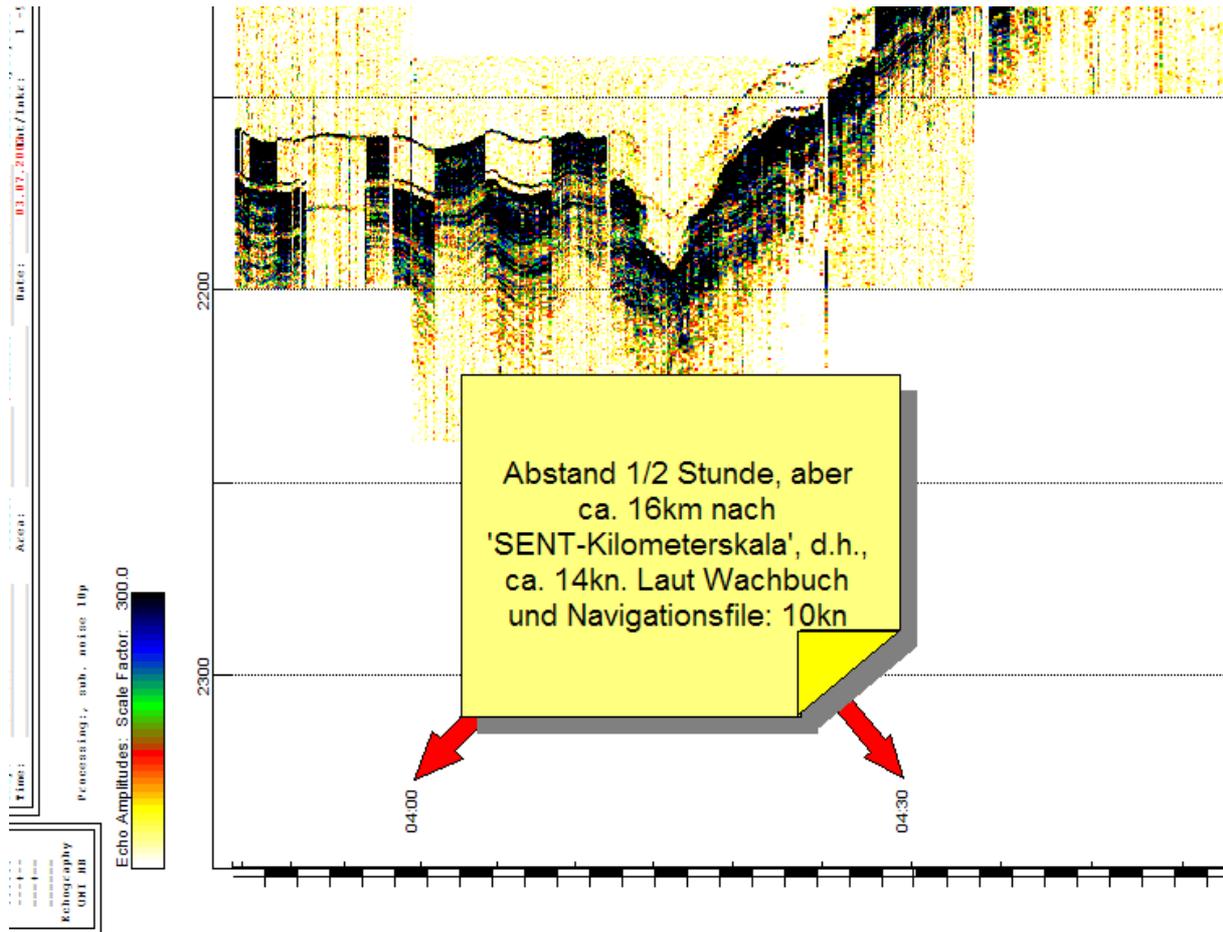


Abbildung 4: Ausschnitt aus dem überprüften Profil.

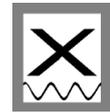
### c) Update 08/2003 - Programm SENT berichtigt

Seit dem 08.08.2003 ist dieser Fehler in SENT behoben. Dr. Hanno von Lom (Hanno von Lom [110i@uni-bremen.de](mailto:110i@uni-bremen.de)) hat den fehlenden Kosinus in der Distanzberechnung eingefügt. Über ihn oder über FIELAX ([info@fielax.de](mailto:info@fielax.de)) kann die aktuelle Version des Programms bezogen werden.

## 6. Zur Metainformation

In Absprache mit AWI-Geologie und AWI-Rechenzentrum enthalten die Metadaten (Tabellen für PANGAEA) folgende Spalten:

	Name	Beispiel	Format	Kommentar
1	Latitude [°N]	74.99766	f8.5	Geogr. Breite in Dezimalgrad
2	Longitude [°E]	2.29006	f8.5	Geogr. Länge in Dezimalgrad
3	Date/Time	27.06.2001 00:00:00	dd.mm.yyyy hh:mm:ss	Datum mit Uhrzeit
4	Speed [kn]	0.1	f3.1	Geschwindigkeit in Knoten
5	Course [°]	254.2	f3.1	Kurs in Grad
6	HS-Depth [m]	3370	i5	Hydrosweep-Tiefe
7	PS-Depth [m]	3364.4	f7.1	Parasound-Tiefe
8	NAV-Depth [m]	3366	i5	Navigationslot-Tiefe
9	Path (ps3)	<a href="http://hs.pangaea.de/para/ps59/ps59_06301307_s.ps3.tar">http://hs.pangaea.de/para/ps59/ps59_06301307_s.ps3.tar</a>	char[80]	Pfad zu den ps3 Dateien. Diese liegen als Tape Archive (tar) vor, sind auch unter WinZip zu entpacken.



10	Path (sgy)	<a href="http://hs.pangaea.de/para/ps59/ps59_06301307_s.sgy">http://hs.pangaea.de/para/ps59/ps59_06301307_s.sgy</a>	<i>char[80]</i>	Pfad zu den SEG-Y Dateien
11	Path (image)	<a href="http://hs.pangaea.de/para/ps59/ps59_06301307_s_01.gif">http://hs.pangaea.de/para/ps59/ps59_06301307_s_01.gif</a>	<i>char[80]</i>	Pfad zu den Grafiken (Übersichtsplots). Diese sind bei langen Profilen ggf. unterteilt
12	Last file	06301438.ps3	<i>mmddhmm.ps3</i>	Letzte Datei zum Track
13	Comments	OFOS, Wait time 20	<i>char[80]</i>	Freier Text, wenn Besonderes im Wachbuch

**Tabelle 1: Beschreibung der Metainformationen**

## 7. ARKXVIII/1a

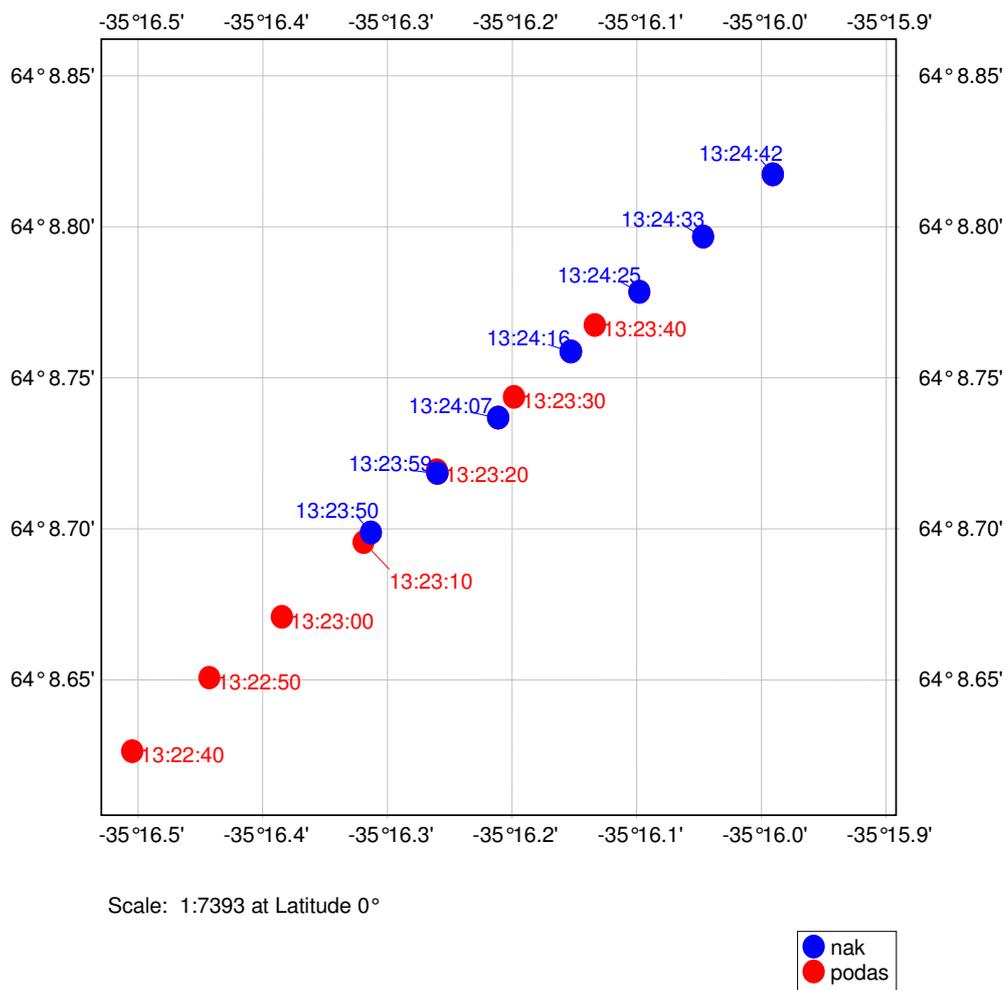
### a) Zur Navigation

Wie schon oben beschrieben, waren die NAK-Files der Bathymetrie nicht ausreichend, um die Parasounddaten zu interpolieren. Die Lücken wurden mit PODAS-Daten aufgefüllt. In einem ersten Versuch wurden die Daten aus PODAS in Lücken über 10 Sekunden aus einem 10s-File aus PODAS in den Hydrosweep-NAK-File eingefügt. Eine genauere Untersuchung der aufgefüllten Tracks ergab leider, dass es auf dieser Reise zwischen Hydrosweep und PODAS offenbar Zeitsynchronisationsfehler ab. Abbildung 5 zeigt deutlich, dass die Überschneidung für den Zeitraum der ARKXVIII/1a z. B. bis zu 40s betragen kann. In diesem Fall gehen wir davon aus, dass dieses durch eine Fehlfunktion in PODAS (NTP-Server?) verursacht wurde. Das Schiff kommt von Südwesten, eingefügt sind die (roten) PODAS-Daten und zwar bis zur Uhrzeit, zu der wieder NAK-Daten vorliegen, also bis incl. 13:23:40, denn ab 13:23:50 (südwestlichster blauer Punkt) gibt es wieder Daten. Wie man aber leicht erkennt, ist das Schiff mit PODAS schon viel weiter im Nordwesten als mit ‚NAK‘. Das interpretieren wir so, dass entweder PODAS' Uhr 40s nach-, oder Hydrosweeps Uhr 40s vorgeht. Es wäre ggf. möglich, die Daten PODAS / HS durch einen static shift konsistent zu machen, dieses wird aber zumindest von uns weder empfohlen noch durchgeführt.

In anderen Vergleichen (andere Fahrten) lagen die Zeitüberschneidungen im unteren Sekundenbereich.

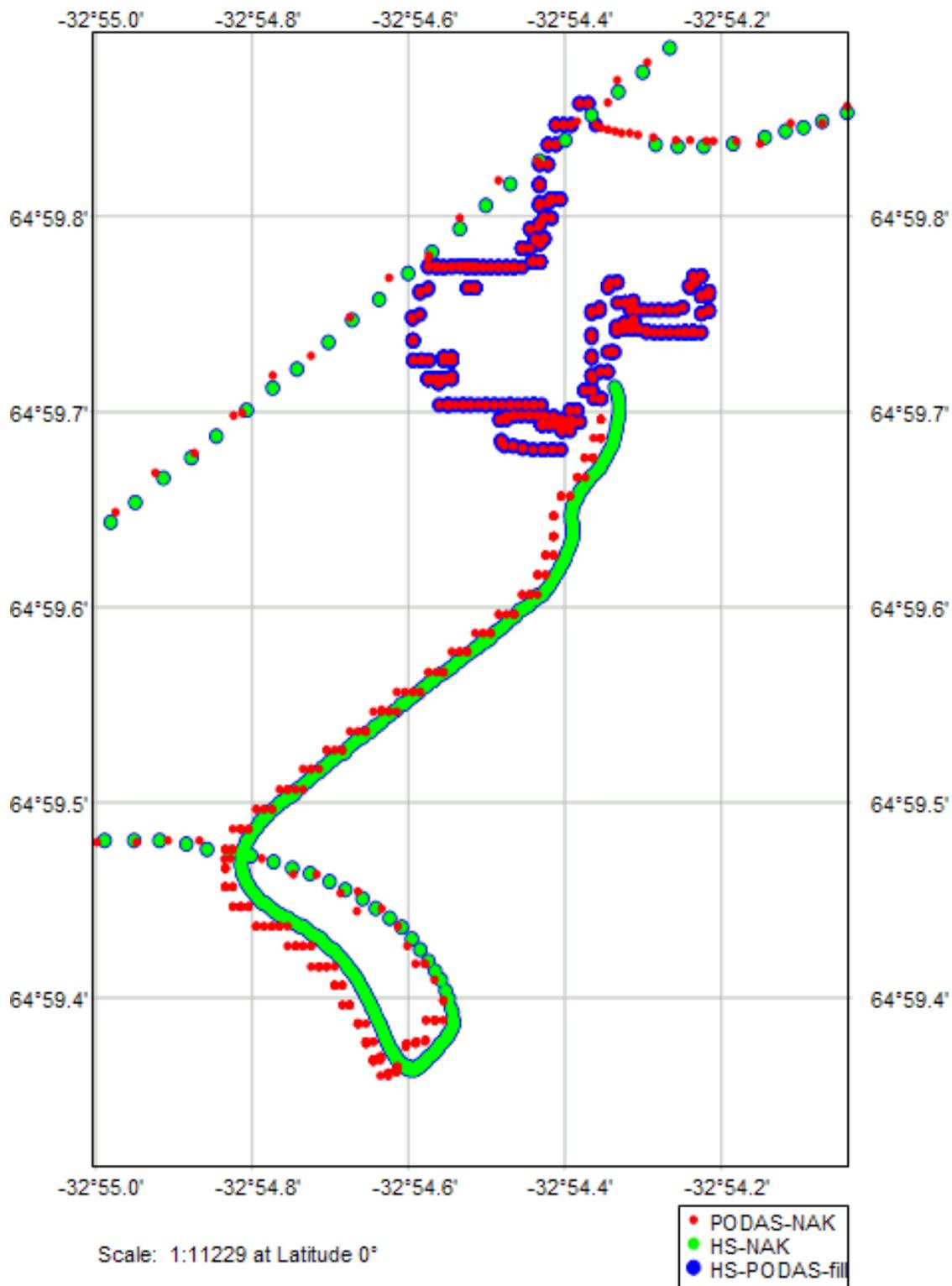
Schon bei den Interpolationsalgorithmen haben wir uns darauf festgelegt, dass zumindest die Originalzeiten der Daten erhalten bleiben sollen und diese ist die gemeinsame Referenz bildet.

Da man bei dieser Größenordnung (~1 Minute) bei Lücken ~2 Minuten noch einen Interpolationsfehler in der Position macht, der größer als der Zeitversatz ist, haben wir entschieden, nur Lücken > 5 Minuten mit PODAS-Daten aufzufüllen.

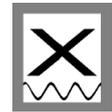


**Abbildung 5: Zeitversatz in den NAK- bzw. PODAS-Daten ARKXVIII/1a**

Bei der Interpolation der Daten fiel noch ein weiterer Fehler auf. In den PODAS-Daten werden unglücklicherweise zu geringe Auflösungen der MINS-Position abgespeichert. Dieses wird veranschaulicht in Abbildung 6: Grün dargestellt ist der Hydrosweep- NAK Kurs, Rote Punkte sind die Daten aus PODAS. Die dahinter liegenden Punkte (Blau) stellen den nach unserer Ergänzung vollständigen Kurs dar, mit dem die Parasounddaten interpoliert werden. Die Treppenstufen in den PODAS-Kursen sind bedauerlich, aber im Nachhinein nicht mehr zu ändern. Wir haben dieses Problem an WERUM weitergeleitet und es wird in der nächsten Version von PODAS bereinigt (siehe dazu auch entsprechende Berichte von Peter Gerchow, FIELAX).



**Abbildung 6: Beispiel, in dem auf Station HS-Nak-Daten fehlen. Auf diesem Bild ist auch zu erkennen, dass die in PODAS abgespeicherten Navigationsdaten die ‚Treppchen‘ enthalten, die durch zu ungenaue Speicherung kommen.**

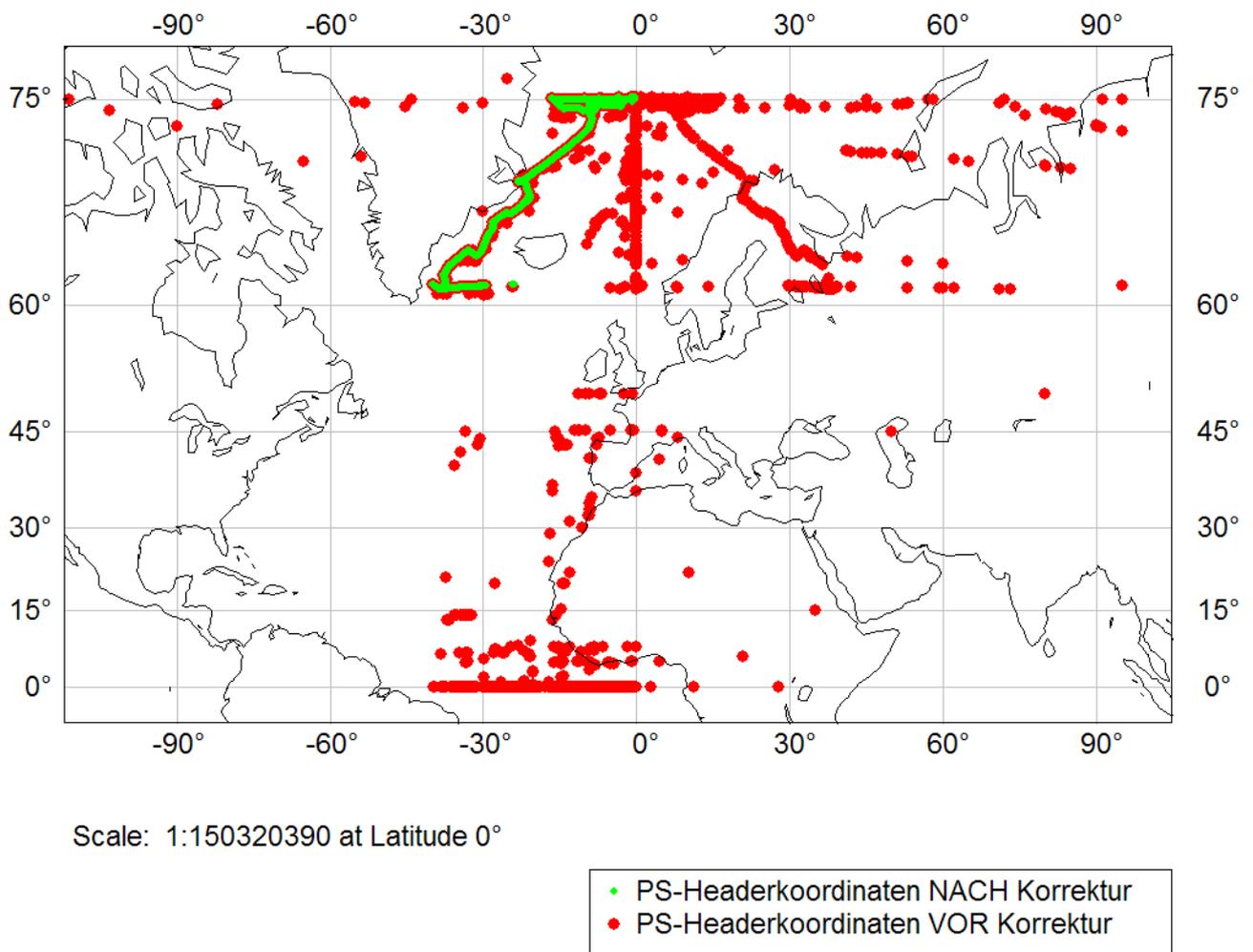


## b) Korrekturen

Datenvolumen ARKXVIII/1a: 10.05GB

Total number of files checked	: 4510
Total number of headers checked	: 739971
Total number of files with invalid date/time	: 296 ( 6.56 percent )
Total number of invalid date/time headers	: 1652 ( 0.22 percent )
Total number of files with wrong lat/longs	: 1528 (33.88 percent)
Total number of wrong lat/long headers	: 20305 (2.74 percent)

Die Prozente sehen nicht so schlimm aus (2.74%), allerdings betrifft es ca. 1/3 der Files und man sieht in der geographischen Darstellung (s. Abbildung 7), dass die falschen Koordinaten **sehr** falsch sein können, so dass die Darstellung der Daten nach ‚Distance‘ nicht sinnvoll sein kann.



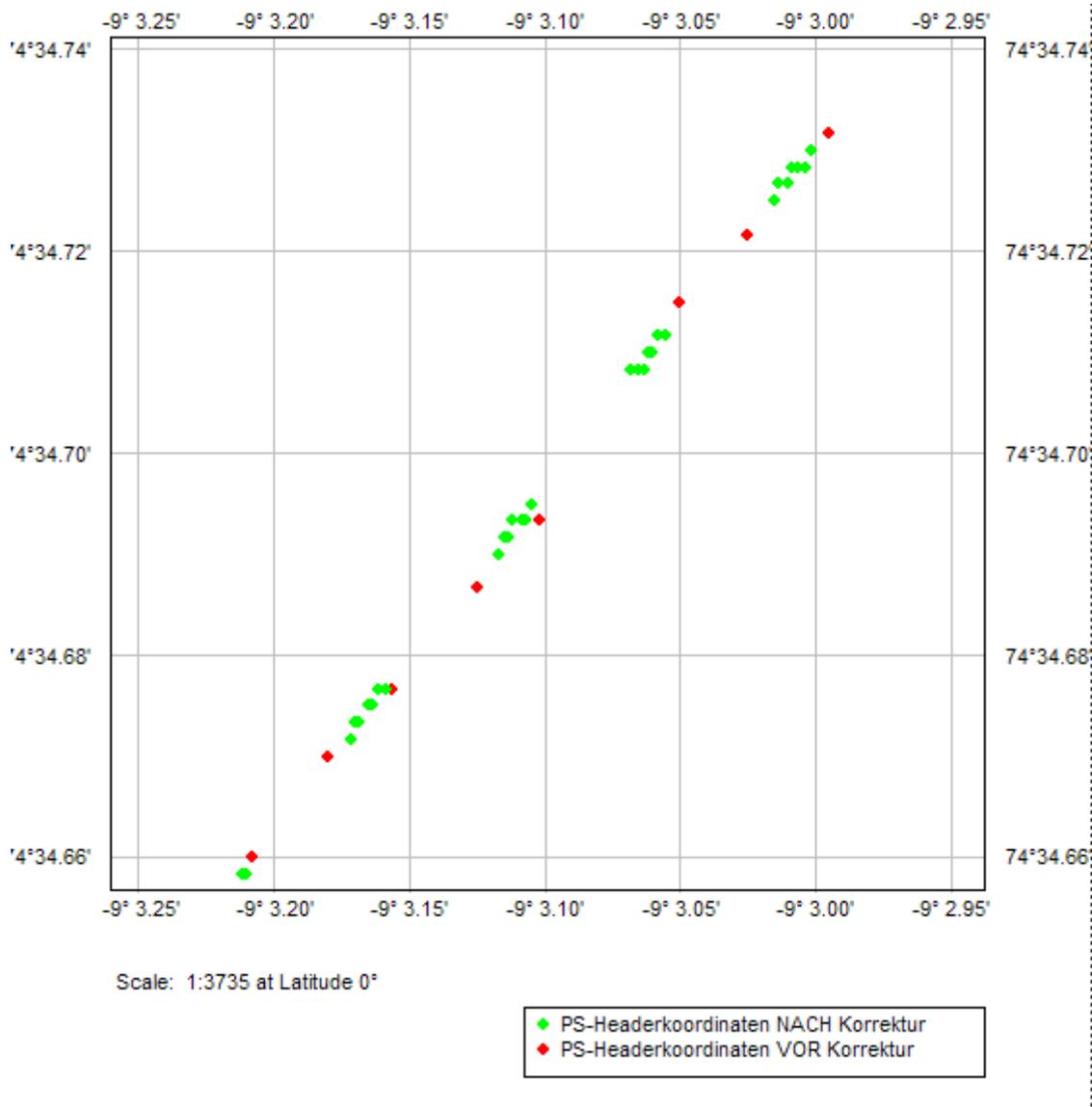
**Abbildung 7: Darstellung der Koordinaten der Reise ARKXVIII/1a aus den Parasond-Headern. Rote Punkte sind die Daten vor der Korrektur, grüne nach der Korrektur.**

Offenbar treten neben scheinbar statistisch verteilten Fehlern hauptsächlich auf:

- 0-Einträge Latitude
- 0-Einträge Longitude
- Vorzeichenfehler Longitude



Da wir nun alle Spuren nach ihrer Zeit richtig in dem NAK-file einordnen, können auch die einzelnen Lotungen (Trains) abgebildet werden. In Abbildung 8 sieht man, dass in den unkorrigierten Daten alle Spuren zu einem Train offenbar auf derselben Koordinate landen (rote Punkte), die interpolierten (grüne Punkte) haben entsprechend ihrer Zeiten unterschiedliche Koordinaten.



**Abbildung 8: Interpolierte Koordinaten in den PS3-Headern, sichtbar nun die individuellen Koordinaten innerhalb der Trains**

### c) *Sonstiges*

Bei der Datenerhebung wurde relativ häufig die Fensterbreite verstellt (100m -> 200m -> 100m), resultierend in unterschiedlich langen Spuren (10640 bzw. 5320 Werte pro Spur). Für Programme wie SENT ist das irrelevant, weil zu jeder Spur der Header eingelesen wird, in dem die Spurlänge steht. Bei den erzeugten ‚segyl‘-Dateien steht ebenfalls zu jeder Spur ein Header mit der Angabe zur Spurlänge, zusätzlich wird aber an den Anfang eines ‚segyl‘-Profiles ein so genannter ‚Reelheader‘ geschrieben, in dem eine Spurlänge angegeben wird. Diese richtet sich nach der Spurlänge in jeweils der ersten zu einem Profil gehörigen Datei (in unserem Fall, Track- oder Stationsverzeichnis). Bei dieser Reise ist es also vorgekommen, dass innerhalb einer ‚segyl‘-Datei die Spurlänge wechselt. Eine ‚faule‘ Software, die nur den Reelheader liest, sich danach aber darauf verlässt, dass die Spuren



innerhalb eines Profils immer gleich lang sind (d.h., nicht bei jeder Spur deren Länge überprüft), wird sich vermutlich beim ersten Wechsel der Spurlänge aufhängen.

## 8. ARKXVII/1

### a) Zur Navigation

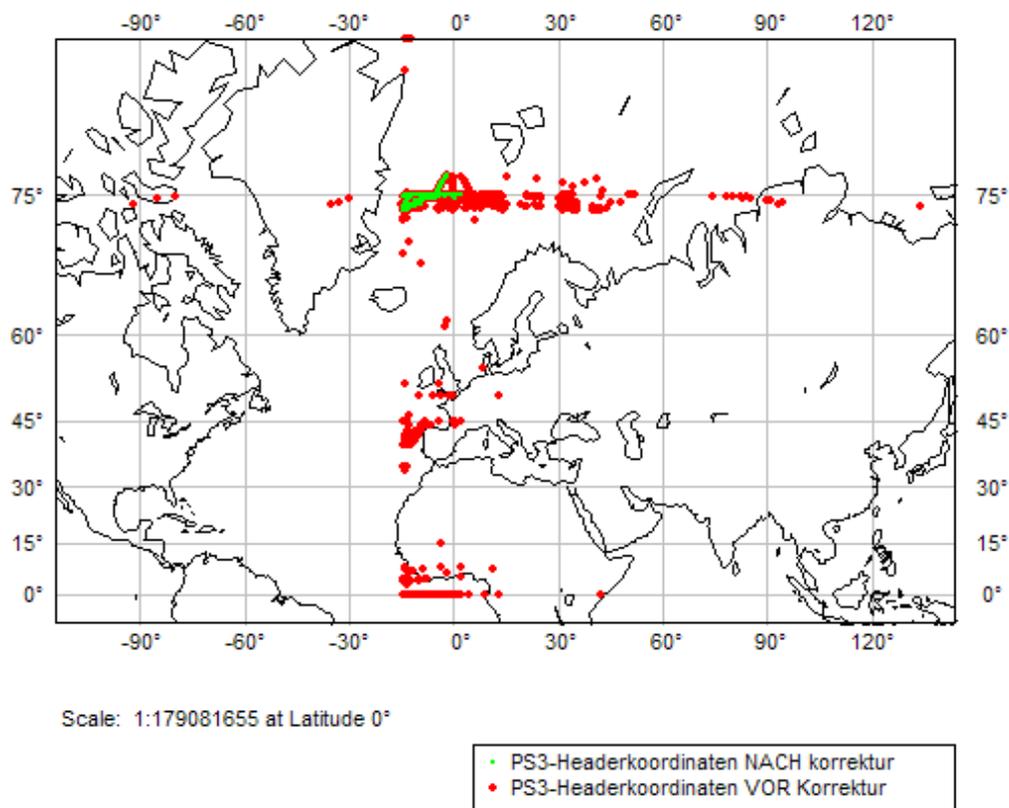
Nach den Erfahrungen aus ARKXVIII konnten die Navigationsdaten relativ schnell zusammengestellt werden. Lücken größer als 5 Minuten wurden mit den von FIELAX erstellten Programmen mit PODAS-Daten aufgefüllt.

### b) Korrekturen

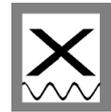
Datenvolumen 5.5GB

Total number of files checked	: 2344
Total number of headers checked	: 275750
Total number of files with invalid date/time	: 165 ( 7.04 percent )
Total number of invalid date/time headers	: 513 ( 0.19 percent )
Total number of files with wrong lat/longs	: 1049 ( 44.75 percent )
Total number of wrong lat/long headers	: 8592 ( 3.12 percent )

Die Prozente sehen wieder nicht so schlimm aus (3.12%), allerdings betrifft es knapp die Hälfte der Files und man sieht in der geographischen Darstellung (s. Abbildung 9), dass die falschen Koordinaten **sehr** falsch sein können, so dass die Darstellung der Daten nach ‚Distance‘ nicht sinnvoll sein kann.



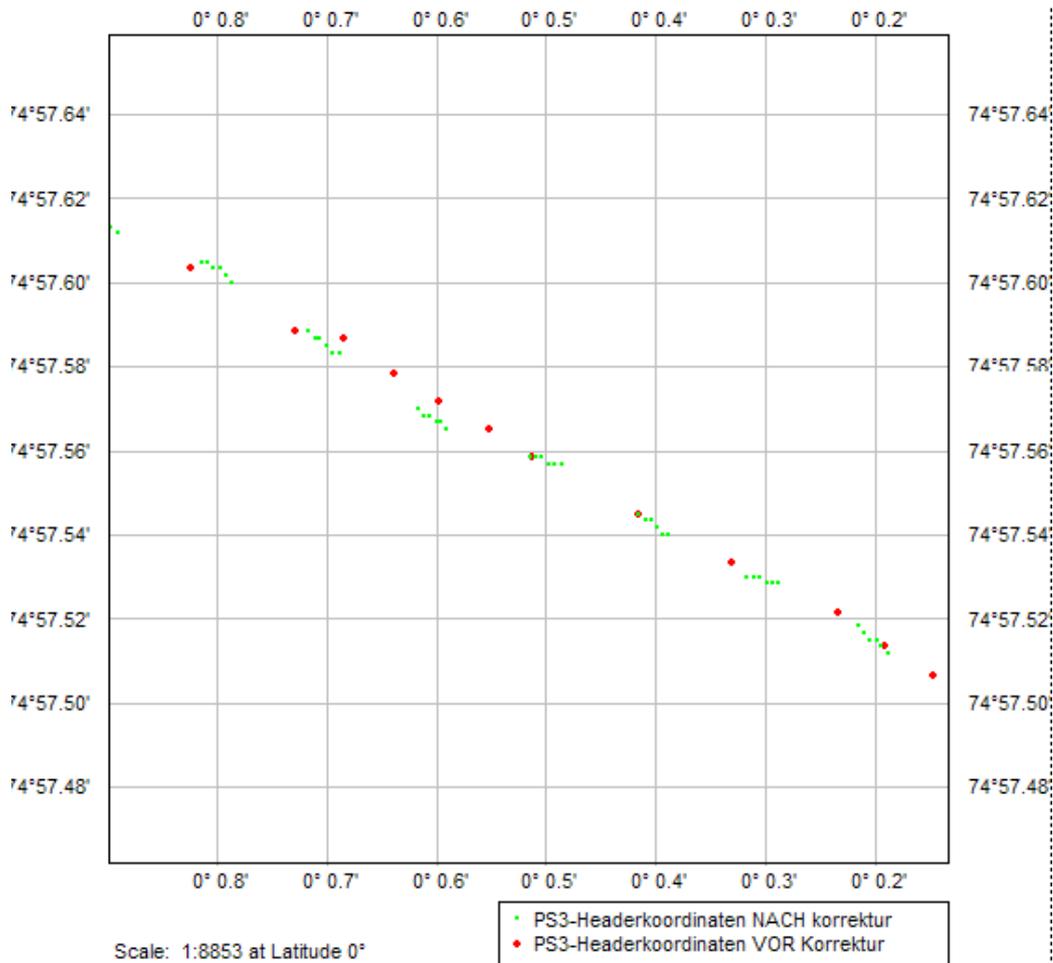
**Abbildung 9: Darstellung der Koordinaten der Reise ARKXVII/1 aus den Parasound-Headern. Rote Punkte sind die Daten vor der Korrektur, grüne nach der Korrektur.**



Offenbar treten neben scheinbar statistisch verteilten Fehlern hauptsächlich auf:

- 0-Einträge Latitude
- 90-Einträge Latitude
- 0-Einträge Longitude
- Vorzeichenfehler Longitude

Da wir nun alle Spuren nach ihrer Zeit richtig in dem NAK-file einordnen, können auch die einzelnen Lotungen (Trains) abgebildet werden. In der nächsten Abbildung sieht man, dass in den unkorrigierten Daten alle Spuren zu einem Train offenbar auf derselben Koordinate landen (rote Punkte), die interpolierten (grüne Punkte) haben entsprechend ihrer Zeiten unterschiedliche Koordinaten.



**Abbildung 10: Interpolierte Koordinaten in den PS3-Headern, sichtbar nun die individuellen Koordinaten innerhalb der Trains ARKXVII/1**

### c) *Sonstiges*

- Die Metainformationen in PODAS enthalten für die gesamte Fahrt keine Informationen über ‚Speed‘ und ‚Course‘ im ‚System‘-Eintrag. Die entsprechenden Daten wurden daraufhin aus PODAS aus den Daten für die MINS extrahiert.
- Es fiel bei der Grafikerstellung mit SENT auf, dass nicht immer, wenn ein Fenster verschoben wird, eine neue Achse gezeichnet wird. Dieses kann u. U. daran liegen, dass die aktuelle Fenstertiefe nicht richtig mitgeführt wurde. Sollte sich dieses als ein häufigerer Fehler herausstellen, müsste man hier nach weiteren Unstimmigkeiten / Fehlerquellen suchen.



## 9. ARKXVI/1

### a) *Zu den Tapes*

Leider mussten wir beim Kopieren der Daten von CD feststellen, dass dieses nicht immer möglich war (CDs nicht lesbar) und außerdem in fast allen Verzeichnissen Dateinamenfehler auftreten. Die Daten wurden uns schon auf CD übergeben. Diese wurden vorher am AWI von den Tapes erzeugt. Hierfür wird ein Programm von V. Spieß (Tape Ver. 46) benutzt, welches zunächst PS3-Dateien auf einer SCSI-Wechselplatte erzeugt. Diese wird an einem anderen Rechner auf das Netz, die Daten dann wiederum auf einem anderen Rechner auf CD gebrannt. Auf den CDs sind jeweils 2-4 Tapes in Unterverzeichnissen mit dem entsprechenden Tapesnamen (z. B. Tape46).

#### - **Schreib-/Lesefehler**

Insgesamt etwa 1.8GB von insgesamt 11.5GB waren von den CDs nicht lesbar. Es kamen verschiedene Fehler vor (keine Daten, stattdessen Wiederholung des Strings ‚C:‘ ; ‚RPC‘ File not found; etc.). Das entspricht einer Fehlerrate von etwa 15%. Es mussten daraufhin von den Tapes noch einmal die Daten kopiert werden (insgesamt 9 Tapes, 3 CDs, Aufwand incl. Tapebeschaffung, CD-Brennen, Einspielen der Daten: 9 Stunden), was bis auf den Dateinamenfehlern (s. u.) Fehler problemlos lief.

#### - **Dateinamenfehler**

Die Dateinamen werden offenbar dynamisch nach der Konvention ‚mmyyhhmm.ps3‘ erzeugt (gelesen aus den Spurheadern?). Beim Einlesen eines neuen Tapes wird sehr zuverlässig und reproduzierbar für den ersten Dateinamen ein falsches Datum generiert. Die Uhrzeit passt sehr gut zu den nachfolgenden Dateinamen. Daraufhin haben wir nicht nur bei den neu erstellten CDs die Datumsangaben geändert, sondern alle schon kopierten (bei FIELAX auf Platte zur Weiterbearbeitung) Dateien entsprechend umbenannt (Aufwand ca. 6 Stunden). Im Folgenden eine Auflistung der geänderten Dateinamen:

Alter Dateiname	Neuer Dateiname	Tape-Nr.
03280225	07060225	3
03280745	07060745	5
03281650	07061650	6
03290203	07070203	8
03290540	07070540	10
03290835	07070835	11
03292138	07072138	12
03301025	07081025	13
03310117	07090117	14
03311943	07091943	15
04010608	07100608	16
04012218	07102218	17
04020847	07110847	18
04022144	07112144	19
04030921	07120921	20
04040017	07130017	21
04041331	07131331	22
04050016	07140016	23
04050356	07140356	24
04050651	07140651	25
04050931	07140931	26
04051125	07141125	27
04051756	07141756	28
04060040	07150040	29
04061650	07151650	31
04070443	07160443	32
04071447	07161447	33
04080348	07170348	35
04081227	07171227	36
04081659	07171659	37
04082335	07172335	38
04090623	07180623	39
04091136	07181136	40
04091657	07181657	41
04092355	07182355	42



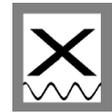
01311605	07191605	44
07190000	07200000	45
02010907	07200907	47
02011911	07201911	48
02041720	07231720	59
02070521	07260521	61
02070744	07260744	62
02071001	07261001	63
02071847	07261847	64
02072105	07262105	65
02072247	07262247	66
02080029	07270029	67
02080257	07270257	68
02080440	07270440	69
02080758	07270758	70
02081144	07271144	71
02081642	07271642	72
02082125	07272125	73
02082328	07272328	74
02090126	07280126	75
02090653	07280653	77
02090847	07280847	78
02091220	07281220	79
02091514	07281514	80
02100025	07290025	82
02100514	07290514	83
02101009	07291009	84
02101232	07291232	85
02101413	07291413	86
02101623	07291623	87
02102204	07292204	89

## b) Zur Navigation

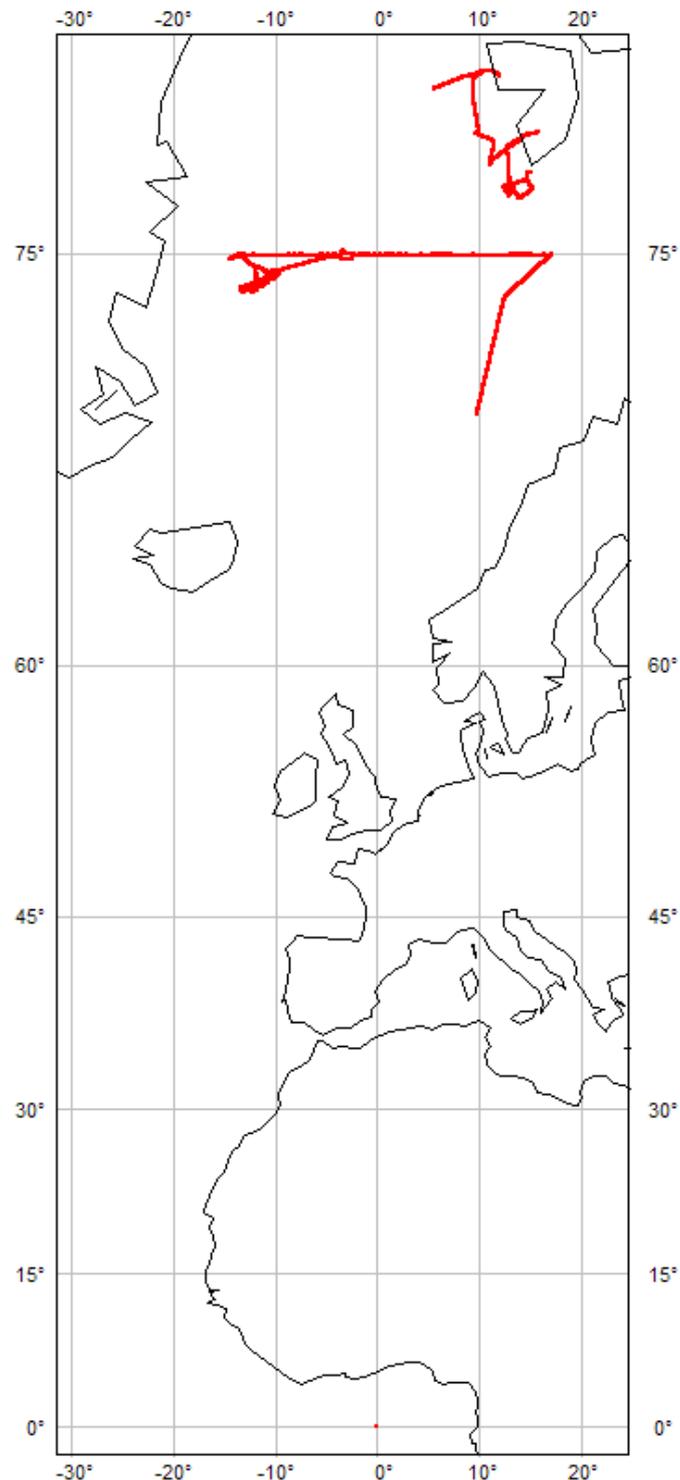
Die Navigation war sehr schwer zu beschaffen. Die Daten in der Hydrosweep-NAK Datei waren stark lückenhaft, die Beschaffung der älteren Daten via PODAS nicht trivial, da diese noch nicht in der aktuellen Datenbank vorliegen. Die Konvertierung und besonders die Überprüfung der Altdaten hat noch nicht stattgefunden und wir haben noch Fehler gefunden, die an Peter Gerchow und WERUM weitergeleitet wurden. Offensichtlich ist es so, dass die Parasoundheader und die Hydrosweepheader konsistent sind (identische Telegramme?), die in PODEV / PODAS abgespeicherten Navigationsdaten weisen aber dazu einen Zeitversatz von genau einem Tag auf. Nach unseren Recherchen liegt das daran, dass die Telegramme zu der Zeit noch von dem NACOS-Vorgänger-System ANP generiert wurden, dieses hatte einen Tag Zeitfehler (Jahr 2000 / Schalttag vergessen). Damit wird verständlich, warum die Daten in dem Wachbuch nicht zu den Zeitinformationen in den PS3-Dateien passen. Beispiel: Erste Station laut Wachbuch am 05.07.2000 09:00, erste ‚kleine‘ Datei (d. h. Wait Time 60) 07060912.ps3, d. h. am 06.07.2000 um 09:12!!! Problematisch ist dieses natürlich, denn wenn man von den Hydrosweepdaten ausgeht, so wird man nach den dort gefundenen Zeiten und Koordinaten in den Parasounddaten suchen. Wenn man aber z. B. von den Stationsbüchern (z. B. gezogenen Sedimentkernen) ausgeht, so wird man nach den dort vorgefundenen Zeiten / Koordinaten suchen. Eigentlich gibt es zwei Möglichkeiten, diesem Dilemma zu entgehen. Entweder man korrigiert alle Hydrosweep- und Parasounddaten (erfordert großen Zeitaufwand) oder man vermerkt an einer offensichtlichen Stelle, dass dieser Zeitversatz existiert. Wir haben im Wachbuch vermerkt, dass es diese Diskrepanz gibt. Für die Metadaten in Pangaea könnte man in einer Kommentarspalte darauf aufmerksam zu machen.

Wegen der Verwechslungsgefahren wurde entschieden, diesen Zeitfehler zu korrigieren. Die Parasounddaten haben also nunmehr korrekte Zeiten, was man beim Vergleich mit Hydrosweepdaten beachten muss, die noch nicht korrigiert wurden.

Für diese Reise haben wir die Daten aus den originalen PS3-Headern angeguckt, um ggf. diese Daten zum Auffüllen der NAK-Dateien zu nutzen. Die Daten sahen sehr gut aus, so dass wir die gesamte Navigationsinformation aus den Headern extrahiert haben, um diese zu prüfen. Es hat sich gezeigt, dass offenbar deutlich weniger, als auch andere Fehler in der Navigation auftauchen. Beim Vergleich von Abbildung 7, Abbildung 9 und Abbildung 11 sieht man deutlich, dass bei der Reise aus dem Jahr 2000 (ARKXVI, Abbildung 11) nur der Fehler ‚(0,0)-Koordinate‘ auftritt, aber nicht die Spiegelungen und ((0,Lon) und (Lat,0)) sowie ‚zufällige‘ Koordinaten‘. Dieses bestätigt unsere frühere



Vermutung, dass es bei der neuen Konfiguration der Datenspeicherung häufiger zu Hard- und Softwarekonflikten kommt. Im Übrigen liegen bei dieser Reise die Headerkoordinaten ziemlich genau auf den Hydrosweep-NAK-Daten, so dass wir uns entschlossen haben, die fehlenden Koordinaten mit einem NAK-File zu interpolieren, den wir aus den Parasound-Headern erzeugt haben nachdem die fehlerhaften Daten (0,0) entfernt wurden.



**Abbildung 11: Darstellung der Koordinaten der Reise ARKXXVI/1 aus den Parasound-Headern. Rote Punkte sind die Daten vor der Korrektur.**



## c) Korrekturen

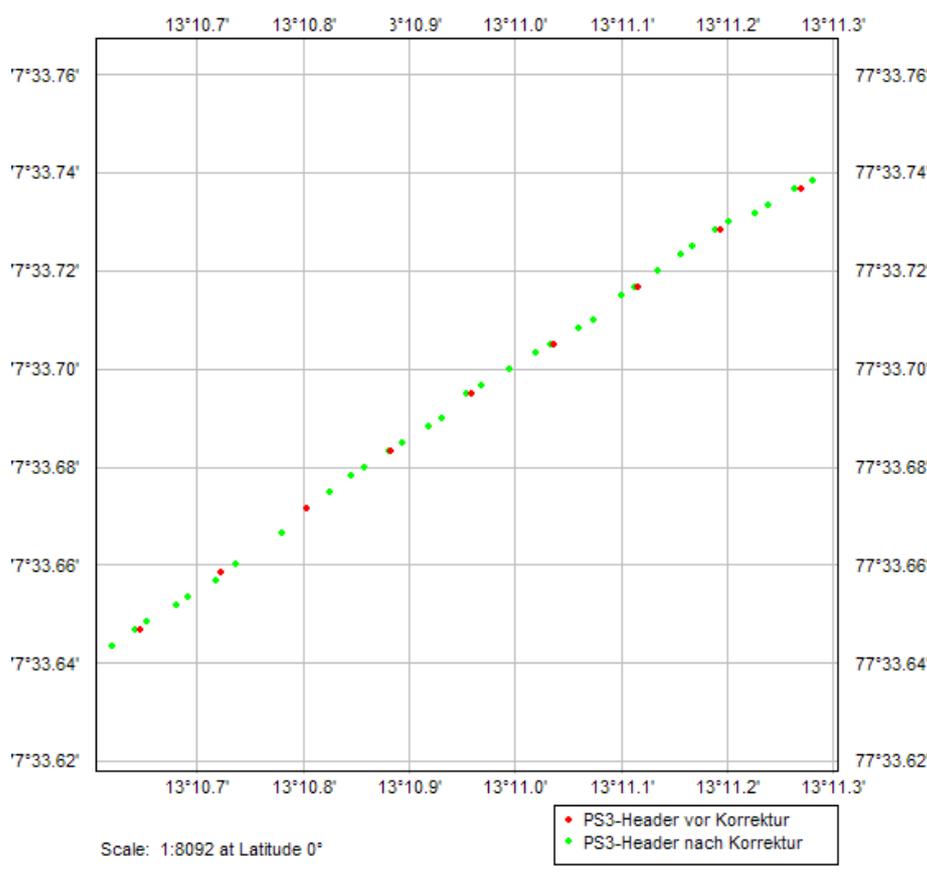
Datenvolumen 11.5GB

Total number of files checked	: 5829
Total number of headers checked	: 389710
Total number of files with invalid date/time	: 50 ( 0.86 percent )
Total number of invalid date/time headers	: 110 ( 0.03 percent )
Total number of files with wrong lat/longs	: 6 ( 0.10 percent )
Total number of wrong lat/long headers	: 82 ( 0.02 percent )

Nach einem zweiten Durchlauf (Spuren mit korrupten Zeiten entfernt) reduziert sich die Anzahl der falschen Koordinaten gar auf nur 32 Spuren (0.01%). Das heißt, dass die Header von den alten Daten (von Tape) sehr viel besser sind, als die neuen (Wechselplatte), siehe auch Bemerkungen in 7.b).

Offenbar tritt hier nur der Fehler (0,0) in (Lat,Lon) auf (s.o.).

Da wir nun alle Spuren nach ihrer Zeit richtig in dem NAK-file einordnen, können auch die einzelnen Lotungen (Trains) abgebildet werden. In der nächsten Abbildung sieht man, dass in den unkorrigierten Daten alle Spuren zu einem Train offenbar auf derselben Koordinate landen (rote Punkte), die interpolierten (grüne Punkte) haben entsprechend ihrer Zeiten unterschiedliche Koordinaten.

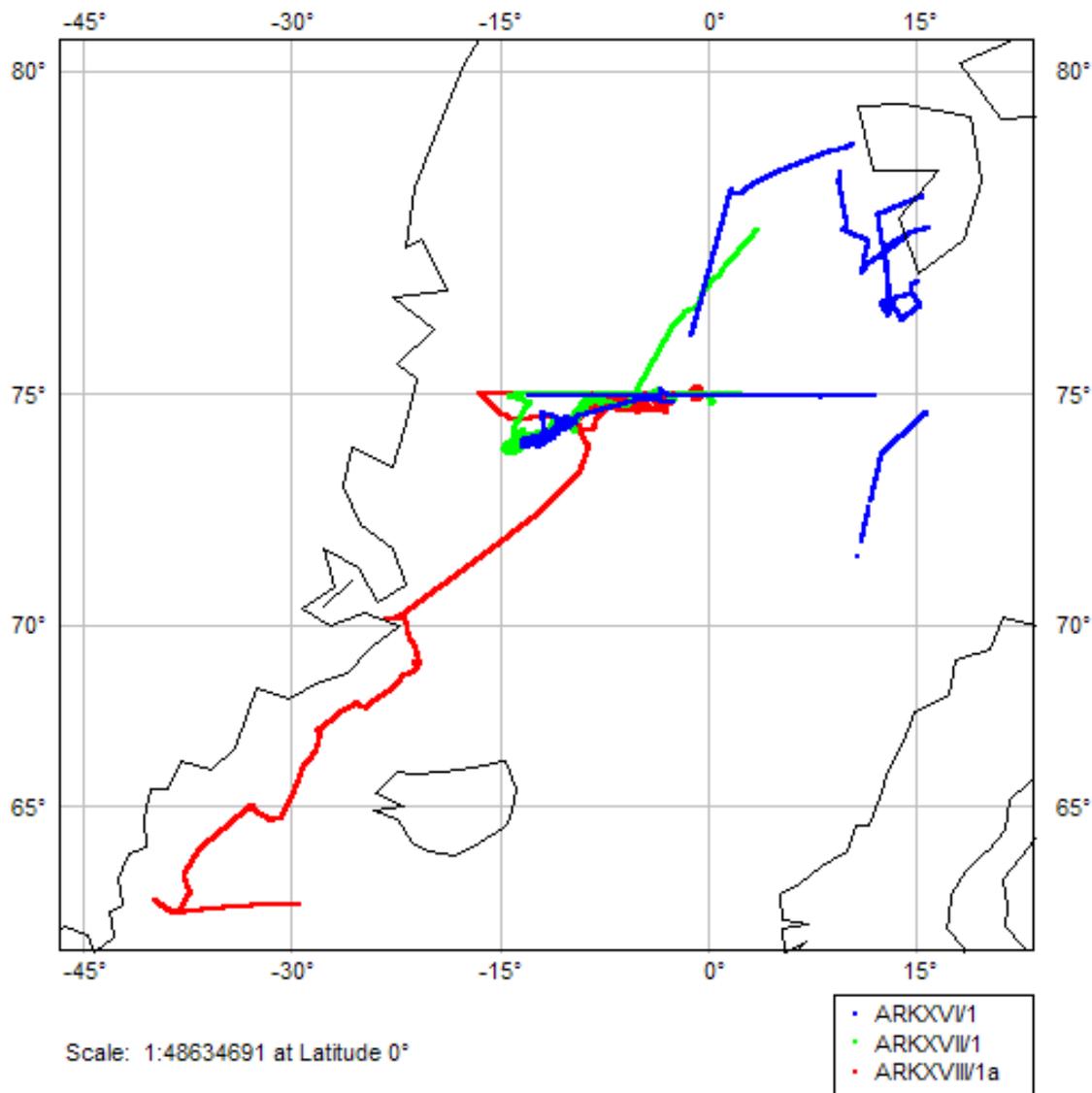


**Abbildung 12: Interpolierte Koordinaten in den PS3-Headern, sichtbar nun die individuellen Koordinaten innerhalb der Trains ARKXVI**



## 10. Zusammenfassung

Insgesamt wurden mit den drei Expeditionen ca. 27GB Parasounddaten vollständig prozessiert. Als Übersicht zeigen Abbildung 13 und Abbildung 14 die Tracks der einzelnen Fahrten.



**Abbildung 13: Navigation aller drei bearbeiteter Fahrten**

Die PS3-Dateien wurden bereinigt (korrupte Spuren entfernt) und die Navigation wie oben beschrieben korrigiert. Die Dateien wurden nach Tracks und Stationen sortiert und in Unterverzeichnissen abgelegt. Aus diesen Verzeichnissen wurden jeweils ‚SEG-Y-Profile‘ erzeugt. Für alle Expeditionen wurden je zwei Metafiles erstellt, der in Pangaea einsehbar ist. Diese enthalten in 2- sowie in 10-Minuten Intervallen neben allen Informationen, die auf Polarstern auch als ‚Table‘ ausgedruckt werden, die Pfadinformationen der erzeugten und am AWI zentral abgelegten Files.

Außerdem sind aus den Wachbüchern wichtige Kommentare, sowie Auffälligkeiten aus der Prozessierung (z. B. ‚empty traces‘) in einer Kommentarspalte abgelegt. Die Metadaten sind in PanMap importierbar und lassen den Nutzer ‚auf Mausklick‘ alle Informationen zukommen (vergl. Abbildung 15).

Ebenfalls für jede Fahrt wurde ein ‚Processing Log‘ im ‚pdf‘-Format, erstellt, in dem wesentliche Beobachtungen und Prozessierschritte beschrieben werden. Diese Dateien sind ebenfalls in Pangaea verfügbar.

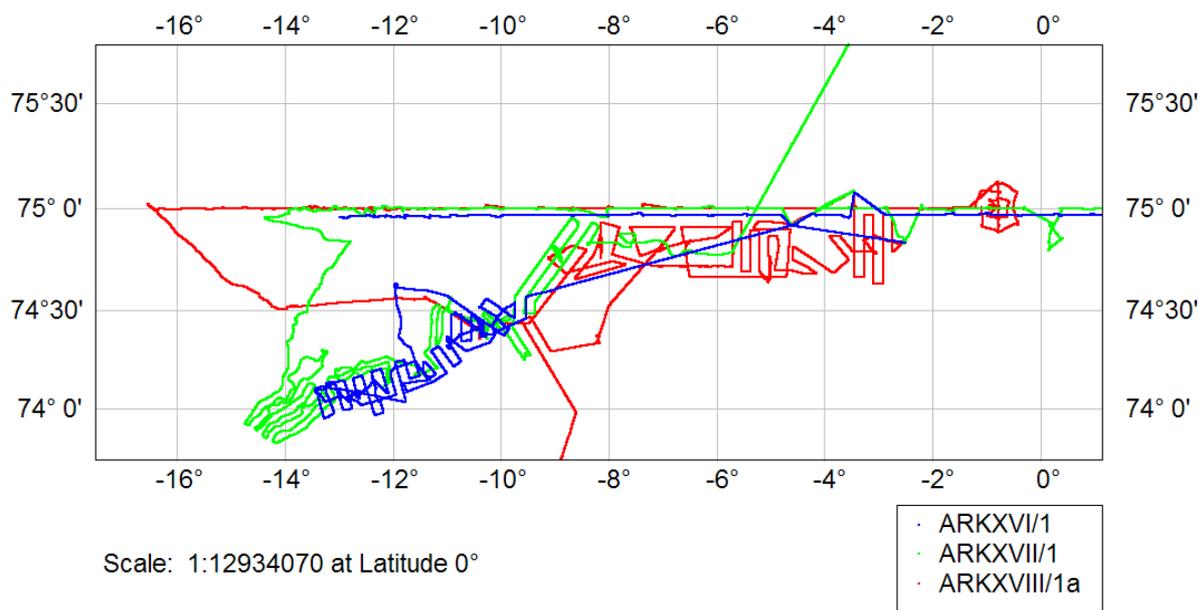


Abbildung 14: Navigation aller bearbeiteten Daten. Ausschnitt, der das ARKTIEF-Gebiet in höherer Auflösung zeigt.

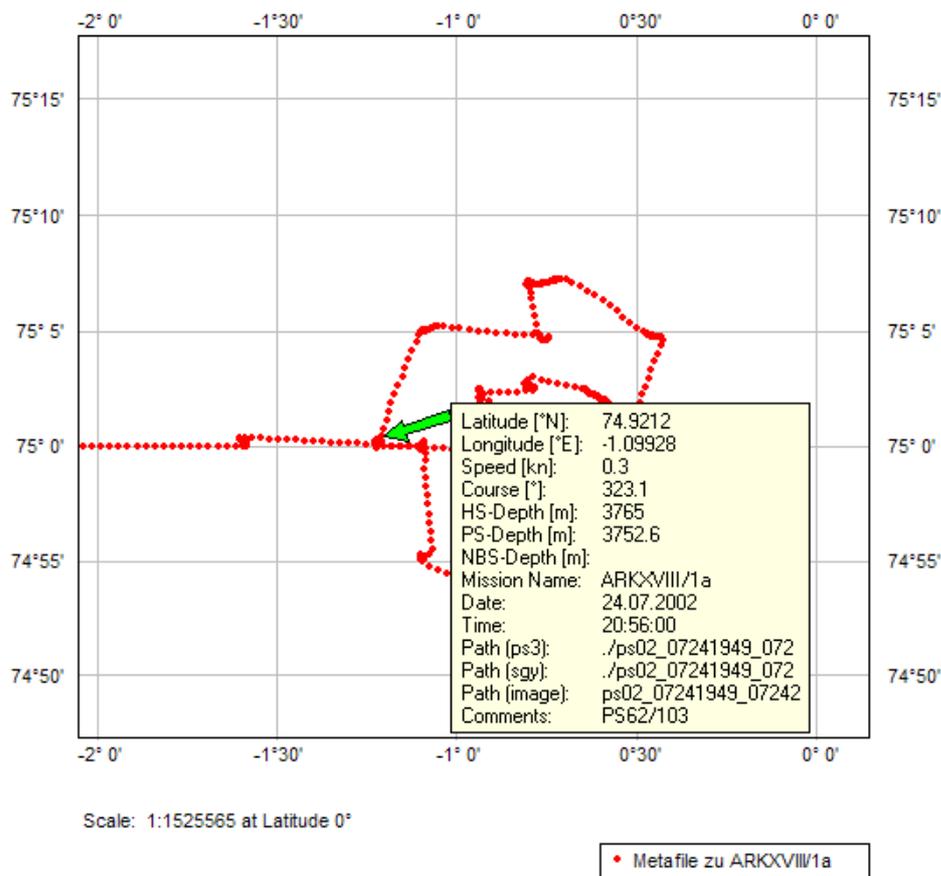


Abbildung 15: Vollständige Trackinformation im 2-Minuten Intervall, Beispiel ARKXVIII/1a

Die Daten liegen nun in einer Form vor, die es erlaubt, sowohl mit spezieller Parasound-Software (SENT), als auch mit klassischer geophysikalischer Software (SEGY-basierend) zu arbeiten, ohne auf irgendwelche Probleme – zumindest technischer Art – zu stoßen.



## ARK-XVI processing log

### a) Original data

Original Parasound data were held on magnetic tapes and copied on CD-ROM. Some of the CD-ROMs were not readable and had to be read from the original media. When reading the tapes, the filenames are generated dynamically and the first file on each tape is assigned to a wrong filename which was corrected.

### b) Navigation

The navigation in the Parasound headers was interpolated using Hydrosweep ,nak'-files and – where no Hydrosweep information was available – PODAS navigation information. In these data, also some gaps were present. Where applicable, Parasound navigation information (extracted from the headers) was edited and also used for interpolation.

### c) Corrections

Data volume 11.5GB

Total number of files checked	: 5829
Total number of headers checked	: 389710
Total number of files with invalid date/time	: 50 ( 0.86 percent )
Total number of invalid date/time headers	: 110 ( 0.03 percent )
Total number of files with wrong lat/longs	: 6 ( 0.10 percent )
Total number of wrong lat/long headers	: 82 ( 0.02 percent )

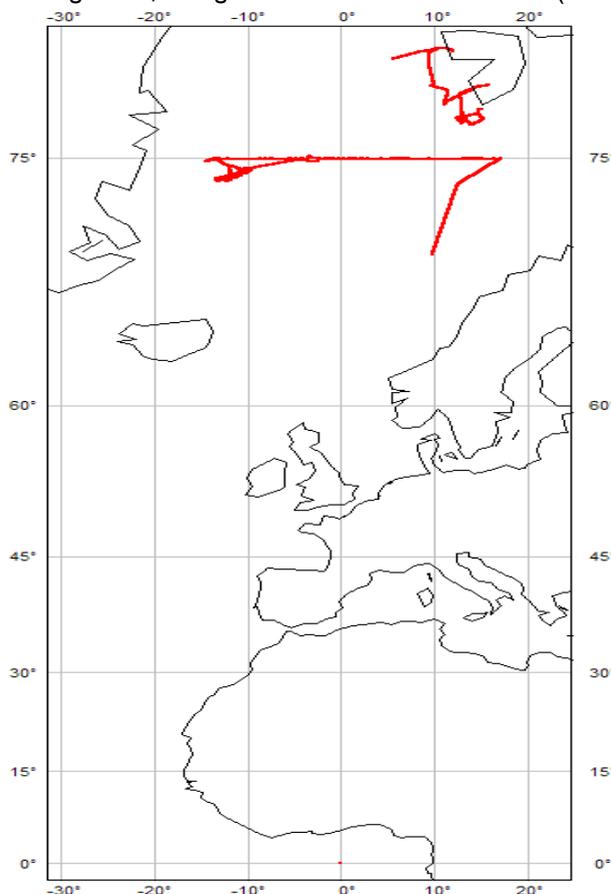
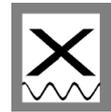


Figure 1: Map showing the full track before (red points) correction and interpolation of navigation data. The only obvious errors are (0,0)-coordinates.



#### **d) Miscellaneous**

- The graphic presentations of the Parasound data are plotted along ‚distance‘. The scale is not correct as the calculation of ‚distance‘ in the program used (SENT) is erroneous (missing scaling dependent on latitude). In August 2003, this bug was fixed by Dr. von Lom, Bremen University ([l10i@uni-bremen.de](mailto:l10i@uni-bremen.de)).
- All Navigation information in the Parasound headers as well as in the Hydrosweep data contained a ‚year-2000-bug‘. The dates are all one day in advance because the original navigation system ‚forgot‘ the 29<sup>th</sup> of February 2000. All Headers have been corrected.



## ARK-XVII processing log

### a) Original data

Original Parasound data were held on CD-ROM (ps3-files).

### b) Navigation

The navigation in the Parasound headers was interpolated using Hydrosweep ,nak'-files and – where no Hydrosweep information was available – PODAS navigation information.

### c) Corrections

Data volume 5.5GB

Total number of files checked	: 2344
Total number of headers checked	: 275750
Total number of files with invalid date/time	: 165 ( 7.04 percent )
Total number of invalid date/time headers	: 513 ( 0.19 percent )
Total number of files with wrong lat/longs	: 1049 ( 44.75 percent )
Total number of wrong lat/long headers	: 8592 ( 3.12 percent )

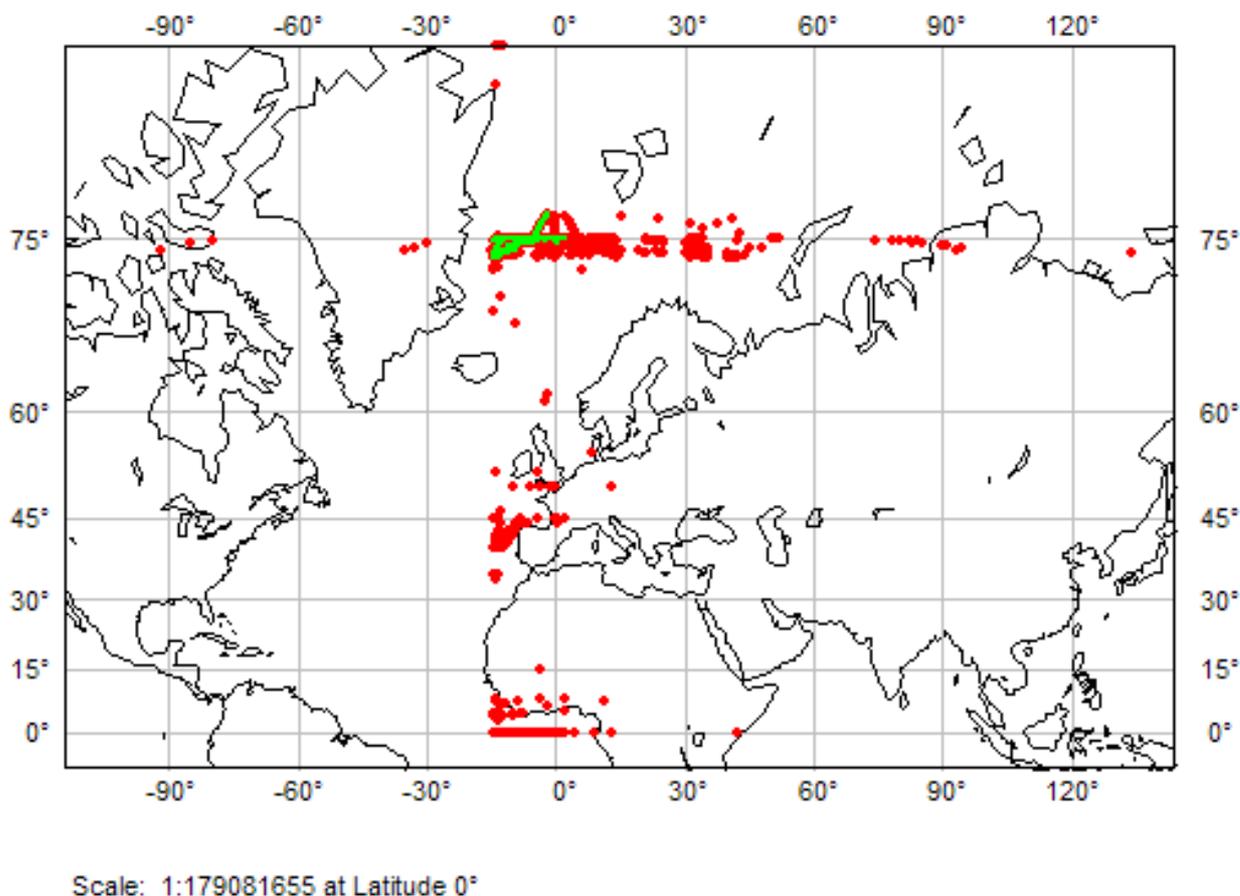
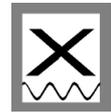


Figure 1: Map showing the full track before (red points) and after (green points) correction and interpolation of navigation data.



#### **d) Miscellaneous**

- The graphic presentations of the Parasound data are plotted along ‚distance‘. The scale is not correct as the calculation of ‚distance‘ in the program used (SENT) is erroneous (missing scaling dependent on latitude). In August 2003, this bug was fixed by Dr. von Lom, Bremen University ([l10i@uni-bremen.de](mailto:l10i@uni-bremen.de)).
- The meta-information in the PODAS-system contained no ‚Speed‘ und ‚Course‘ values withing the ‚System‘-subset. The respective data were extracted from the ‚MINS‘-subset.
- When processing the data with ‚SENT‘, not always when appropriate (i.e. when the data acquisition window was moved) a new axis was drawn. It might be caused by the fact that the actual window depth may be not always set properly in an acquisition subroutine.



## ARK-XVIII processing log

### a) Original data

Original Parasound data were held on CD-ROM (ps3-files).

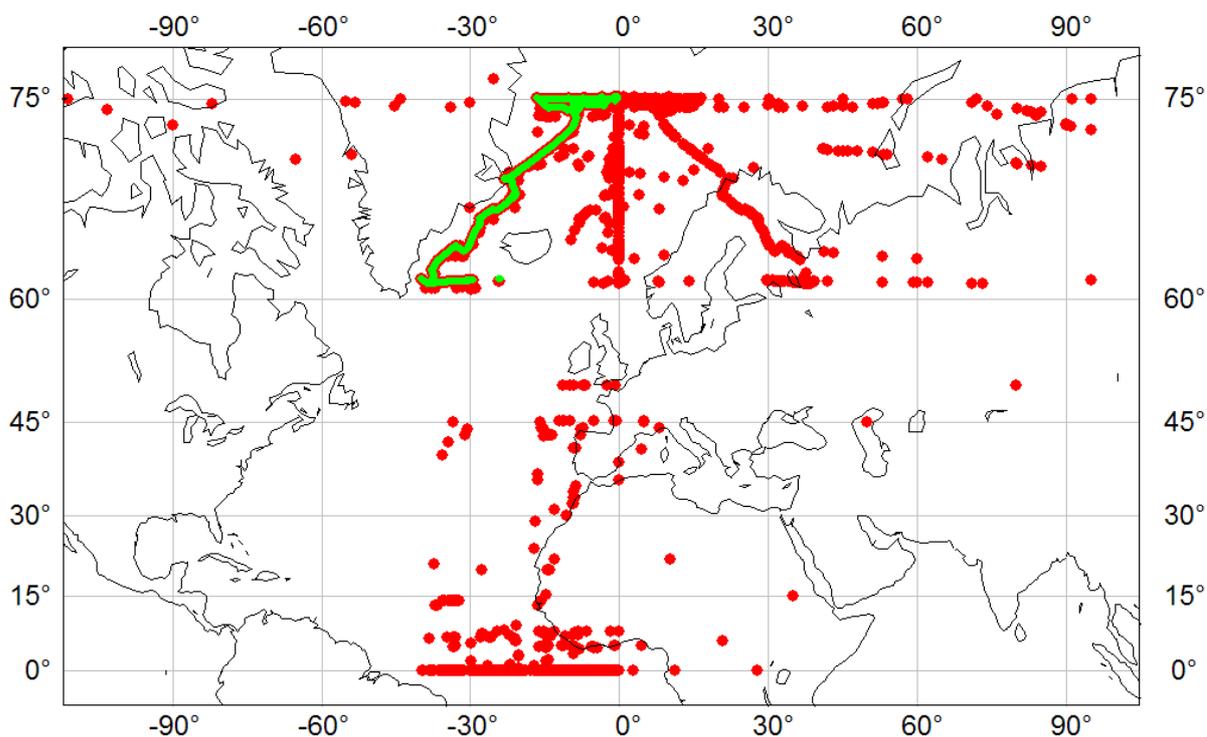
### b) Navigation

The navigation in the Parasound headers was interpolated using Hydrosweep ,nak'-files and – where no Hydrosweep information was available – PODAS navigation information.

### c) Corrections

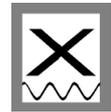
Data volume 10.05GB

Total number of files checked	: 4510
Total number of headers checked	: 739971
Total number of files with invalid date/time	: 296 ( 6.56 percent )
Total number of invalid date/time headers	: 1652 ( 0.22 percent )
Total number of files with wrong lat/longs	: 1528 (33.88 percent)
Total number of wrong lat/long headers	: 20305 (2.74 percent)



Scale: 1:150320390 at Latitude 0°

**Figure 1: Map showing the full track before (red points) and after (green points) correction and interpolation of navigation data.**



#### **d) Miscellaneous**

- The graphic presentations of the Parasound data are plotted along 'distance'. The scale is not correct as the calculation of 'distance' in the program used (SENT) is erroneous (missing scaling dependent on latitude). In August 2003, this bug was fixed by Dr. von Lom, Bremen University ([l10i@uni-bremen.de](mailto:l10i@uni-bremen.de)).
- During this cruise, the width of the data acquisition window was changed a lot between 100m and 200m resulting in different trace lengths (10640 resp. 5320 values per trace). For processing in SENT, this is irrelevant but the 'reel header' in the 'seg3'-files contains the trace length as well. In this case, the trace length of the first converted 'ps3'-file is taken, regardless whether this is true for the whole profile. This might cause problems when using programs which do not check for the trace length in each trace but 'trust' in the reel header.