

アルゴデータ遅延品質管理用歴史的塩分データセット: Selected Hydrographic Dataset (SeHyD)

小林 大洋*¹

アルゴデータの遅延品質管理用の歴史的塩分データセット, Selected Hydrographic Dataset (SeHyD: version 1.0, 2003年11月末現在)を整備した。SeHyDは太平洋全域(縁辺海を含まず)を対象とし, 品質管理されたデータを格納している。今後も更新作業は続けられるが, 最新のもはJAMSTECのアルゴウェブサイトからダウンロードできる。

SeHyDを用いて遅延品質管理を行った結果は, 非常に良好である。修正された塩分値は, そのほとんどがフロート近傍のCTD観測結果や回収後のセンサー検定結果とほぼ一致し, 塩分の修正誤差は ± 0.01 psuよりも小さい。

ダウンロードサイト(2003年11月末現在; SeHyD version 1.0): http://www.jamstec.go.jp/ARGO/product/SeHyD_1.zip

キーワード: アルゴ計画, 品質管理, 塩分修正

Historical salinity dataset for Argo delayed-mode quality control: Selected Hydrographic Dataset (SeHyD)

Taiyo KOBAYASHI*²

For the salinity correction method by Wong et al. (2003), JAMSTEC/FORSGC have made a new Pacific historical dataset, named as Selected Hydrographic Dataset (SeHyD). SeHyD covers the whole open Pacific and consists of the quality-controlled historical profiles. We will continue to improve SeHyD; its latest version (1.0 as of November 2003) is available from ARGO JAMSTEC web site: http://www.jamstec.go.jp/ARGO/product/SeHyD_1.zip.

The performance of the Wong's salinity correction with SeHyD is very good: Most of the corrected salinities agree with the true values obtained by nearby shipboard CTDs or post-calibrations in laboratory, and their correction errors are smaller than ± 0.01 psu.

Keywords : Argo Project, quality control, salinity correction

* 1 地球観測フロンティア研究システム 気候変動観測研究領域

* 2 Climate Variations Observational Research Program, Frontier Observational Research System for Global Change

1. はじめに

アルゴデータの統一的な遅延品質管理(塩分修正)手法として、Wong et al. (2003)を用いることが合意されている。これは、フロート観測点における気候学的な塩分プロファイルを、最適内挿法を用いて計算し、フロート観測値と気候値とのズレを最小にするように塩分観測値を補正するものである。

遅延品質管理によって、より良いデータを得る、つまり何らかの理由によってバイアスの生じた塩分データをうまく修正するためには、優れた修正手法を用いるだけでなく、(その修正システムに合致した)優れた歴史的データセットを用いる必要がある。Wong et al. (2003)は、(このような手法には良くあることだが)調整が必要なパラメータはいくつか存在するものの、手法自体は極めて合理的にできている。つまり、優れた歴史的データセットを用意できるか否かによって、その修正結果の善し悪しが決まると考えてよい(Kobayashi and Minato 2003)。

海洋科学技術センター(JAMSTEC)と地球観測フロンティア研究システム(FORSGC)では、共同してアルゴデータの遅延品質管理用の歴史的データセットを作成し、その太平洋版を完成させた。このデータセットは、Selected Hydrographic Dataset (SeHyD)と名付けられ、JAMSTEC-Argo Webサイト上にてversion 1.0が公開されている(2003年11月末現在)。以下では、このデータセットの作成手順とその注意点、実際の塩分修正結果などを紹介する。

2. データセット作成手順

2.1. 元データセットの特徴および前処理

遅延品質管理用の歴史的塩分データセットは、先行研究(Kobayashi and Minato 2003)に従い、太平洋全体ではデータ格納数の多いWorld Ocean Databaseの最新版を基本とし、既に多くの観測データが得られている日本近海および北米大陸沿岸域では高度に品質管理されたデータセット modified North Pacific HydroBase (小林ら2001)を用いることとする(図1)。以下では両者の特徴と、データセットの前処理について述べる。

2.1.1. World Ocean Database 2001 (WOD01)

WOD01は、アメリカ連邦海洋気象局の海洋データセンター(National Oceanographic Data Center, NODC)によって編纂された最新のプロファイルデータセットであり、公開されているデータセットの中で、最も多くのプロファイルデータを格納していると考えられる。WOD01に格納されている全観測プロファイルから、以下の基準を満たすものだけを選抜し、サブデータセットを作成する。

- ・格納ディレクトリがOSD(採水観測および低鉛直分解能CTD観測データを集めたもの)もしくはCTD(高鉛直分解能CTD観測データを集めたもの)である。
- ・水温と塩分の観測ペアが存在する。
- ・NODCによって付与された品質管理フラグが全て0である(NODCによる品質管理チェックを全てパスしている)。

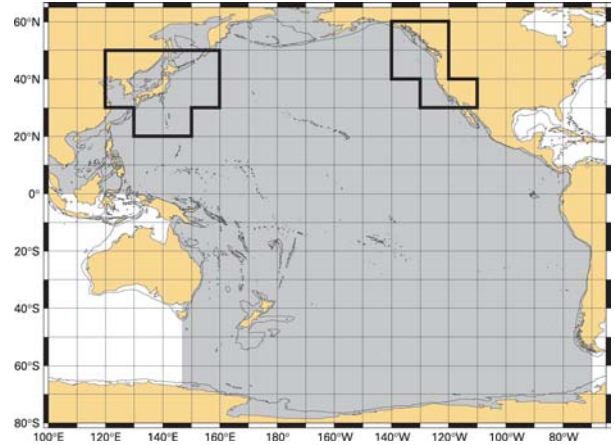


図1 データセットの海域分け。太平洋領域にはハッチをかけている。歴史的塩分データセットの元データとして、黒線で囲まれた部分ではmodHBを、それ以外ではWOD01を用いる。

Fig. 1 The Pacific sector for historical salinity dataset (the hatched area). All profiles originate from WOD01 except for the regions enclosed by thick line from modHB.

- ・観測点の水深が200m以上である(地形データETOPO5に基づく)。

このサブデータセットを改めてWOD01と呼ぶことにし、これ以降はこのサブデータセットからデータが選抜される。

なお、水温・塩分の観測ペアを持つプロファイルが格納されているディレクトリとして、上記の2つのほかにPFL(Profiling float), UOR(Undulating Ocean Recorder), DRB(Drifting Buoy)の3つが存在する。しかし、いずれもデータの正確さの点で疑問があるため、今回は利用しない。なおPFL(およびXCTD観測データ:現在多くの観測がなされているので近い将来データセットに格納されるだろう)は、アルゴ計画の進展に伴い、今後多くのデータが集積・格納されることが予想される。従って、その品質を検討する必要が生じるだろう。

2.1.2. Modified North Pacific HydroBase (modHB)

ModHBは次の2つのデータ群からなる。

1. 北太平洋のHydroBase(Macdonald et al. 2001)に格納されているプロファイルに対して、目視による品質管理を施したもの。全て採水観測もしくは低分解能CTD観測データ。なお、北太平洋HydroBaseはWorld Ocean Atlas 1994を元としている。
2. 2002年4月現在で公開されているWOCE-CTDデータ。品質管理フラグを用いて異常データを削除した後、目視による品質管理を施している。

なお、太平洋領域として、北限をベーリング海峡、南大洋の西端はタスマニア島と南極大陸を結ぶ経度(147°E)、東端はマゼラン海峡(60°W)とし、東西に若干の余裕を持たせた形にする(図1)。なお、インドネシア多島海(現時点では)含まないものとする。なお、インドネシア多島海へ流

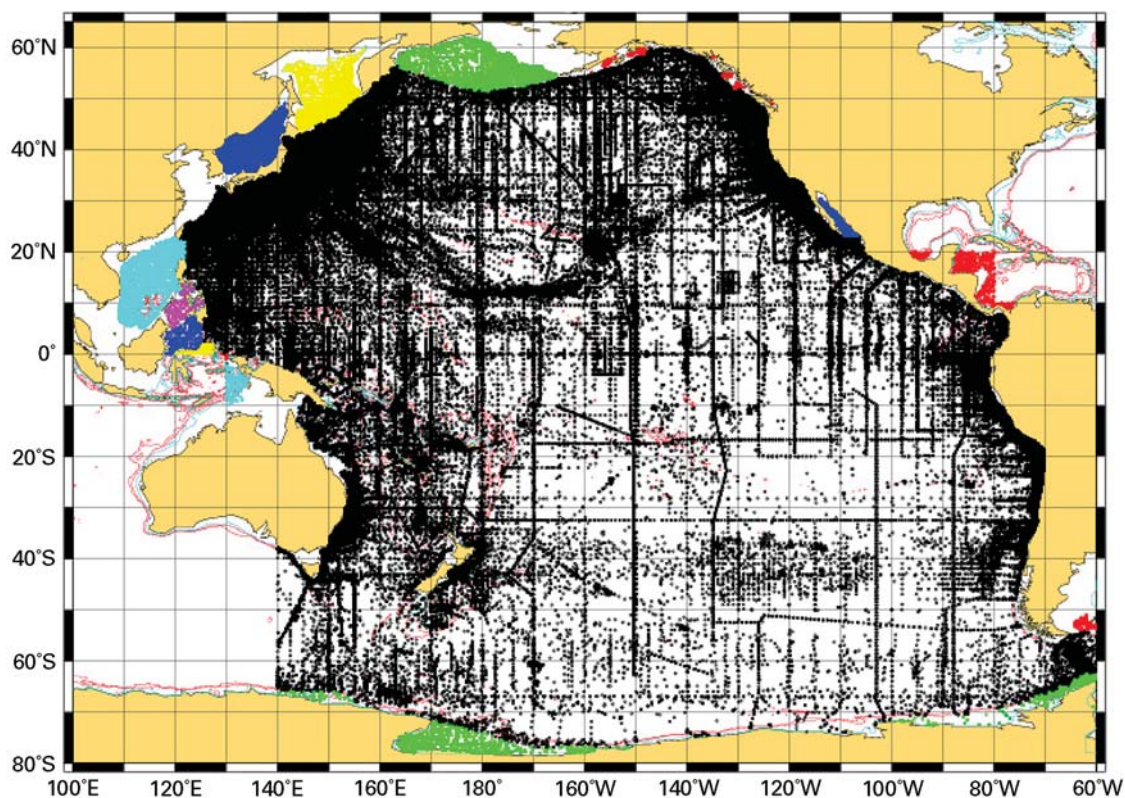


図2 太平洋と縁辺海の分離。点はWOD01に格納されているプロファイルデータの分布を示す。黒で示された部分が太平洋。

Fig. 2 Separations of the open Pacific (black) and the marginal seas (others). All profile data included in WOD01 are also shown by the dots.

入する海水のほとんどは太平洋起源であることから、海域内の水塊は太平洋のものに近い(ただし、海域内での激しい鉛直混合のため、その特徴は大きく変化している)。そのため、大スンダ列島が太平洋とインド洋の境界と見なされることが多い。なお、国際アルゴ計画の領域センター(Regional Center)の担当領域区分に従うと、インドネシア多島海(および北極海の一部)は太平洋に含まれ、40°S以南の海域は南大洋に分類されている。

2.2. 水塊の特徴の把握

Wong et al. (2003) のシステムは、気候学的塩分プロファイルの推定に最適内挿法を用いている。ここで用いられている最適内挿法では、「データ間の変動の類似性は、時空間的な距離に対してガウス型関数で表される」ことを仮定している。逆に言えば、望ましい推定値を得るためには、この仮定を満足するようなデータセットを作る(仮定を満足しないデータは取り除く)必要がある。

2.2.1. 縁辺海の水塊構造の特殊性

太平洋には、日本海・オホーツク海をはじめとする、いくつかの縁辺海が存在する。これらの縁辺海は、多くの場合、島弧や半島・海底地形などによって地形的に太平洋から分離され、さらには海域独自の循環が存在することにより、太

平洋の水塊とはその特徴が異なっている可能性がある。実際、日本海の水塊構造は太平洋のものとは全く異なるし、オホーツク海にも太平洋と異なる水塊構造が存在することが明らかになってきた。これらは、最適内挿法が仮定している条件に明らかに反する。そのため、海底地形を参考にして、太平洋の全ての縁辺海のデータを分離した(図2)。日本海、ベーリング海、南シナ海など、既にフロートが展開されている縁辺海も存在するので、それらの海域の歴史的塩分データセットも順次作成する予定である。

2.2.2. 水温極大層の特徴

データセットに用いるべきプロファイルを選択するに当たり、あらかじめ水温極大層の分布を調べておく必要がある(付録Aを参照)。図3および図4は北太平洋および南太平洋における水温極大層の特徴を示したものである(縁辺海を含む)。北太平洋では、亜寒帯域全域にわたって水温極大層が分布する。西部亜寒帯ではほぼ例外なく水温極大が観測されるが、東部ではその頻度が6割程度まで下がっており、水温逆転層の構造が不明瞭になっていることをうかがわせる。なお、東部亜寒帯の水温極大は、成因の異なる2種類のものからなるとみられ、循環内部の高密度(約27.0 σ_θ)のものは西部亜寒帯から運ばれてきたもの、循環外部の低密度(約26.5-26.7 σ_θ)のものは東部亜寒帯内で形成された

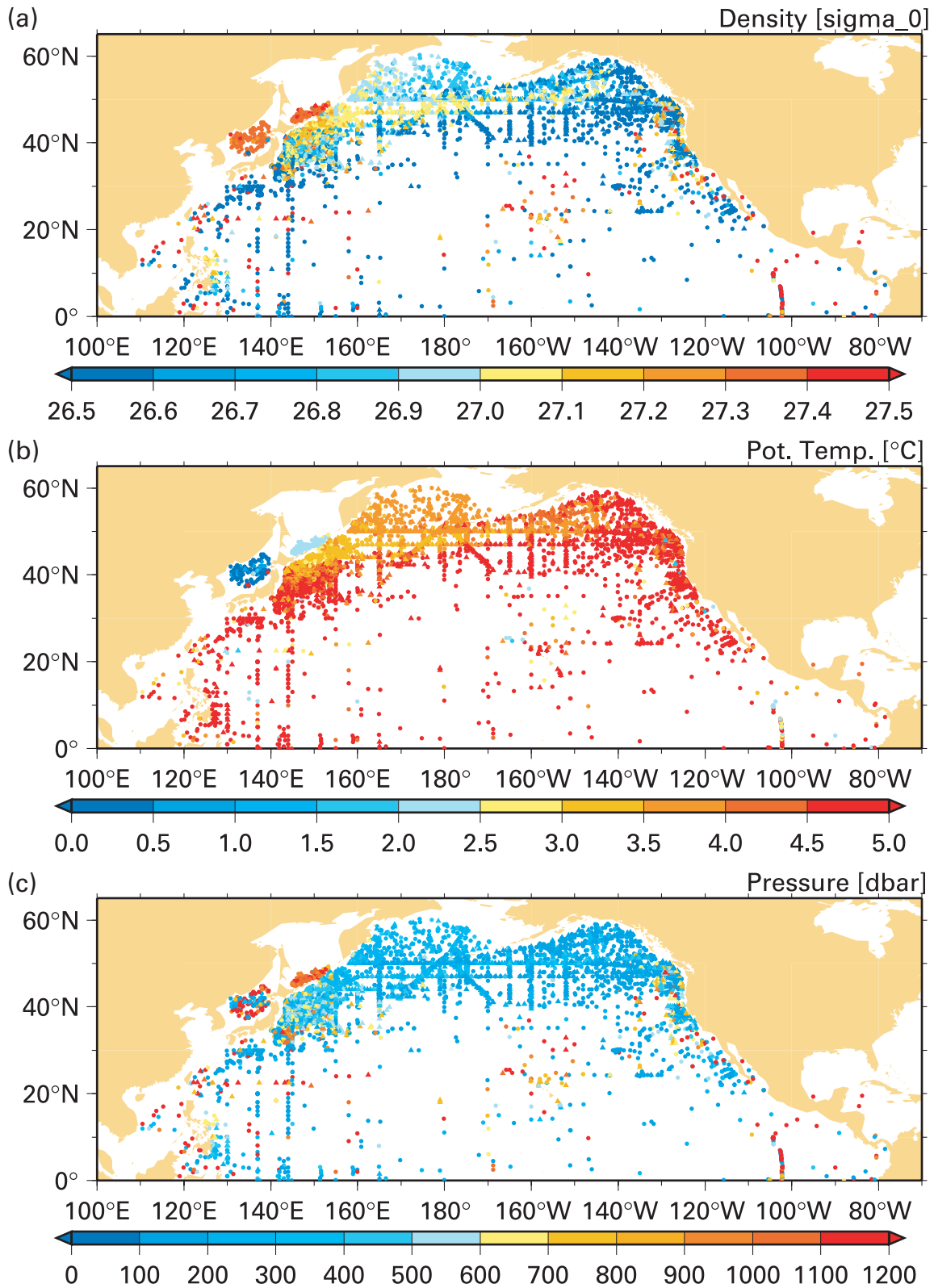


図3 北太平洋における水温極大の(a)ポテンシャル密度、(b)ポテンシャル水温、(c)圧力の分布。2000dbarよりも深層までデータのある観測プロファイルを用いて計算。

Fig. 3 Distributions of (a) potential density, (b) potential temperature, and (c) pressure at temperature maximum in the North Pacific, which is calculated from the profiles the deepest layer of which exceed the depth of 2000dbar.

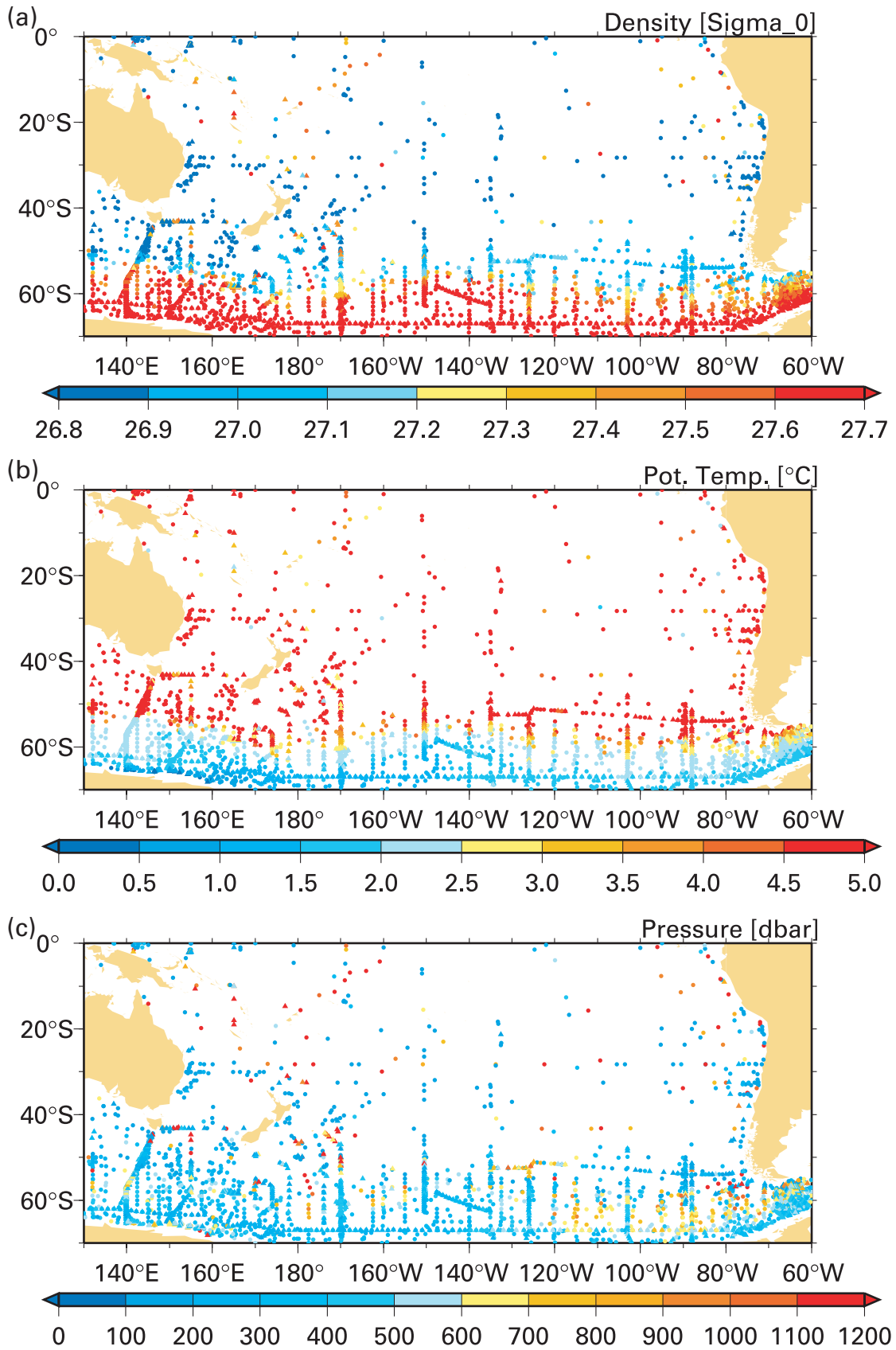


図4 南太平洋における水温極大の分布。図3に同じ。

Fig. 4 Distributions of temperature maximum layers in the South Pacific. Otherwise, as described for Fig. 3.

Profiles through Criteria (WOD01)

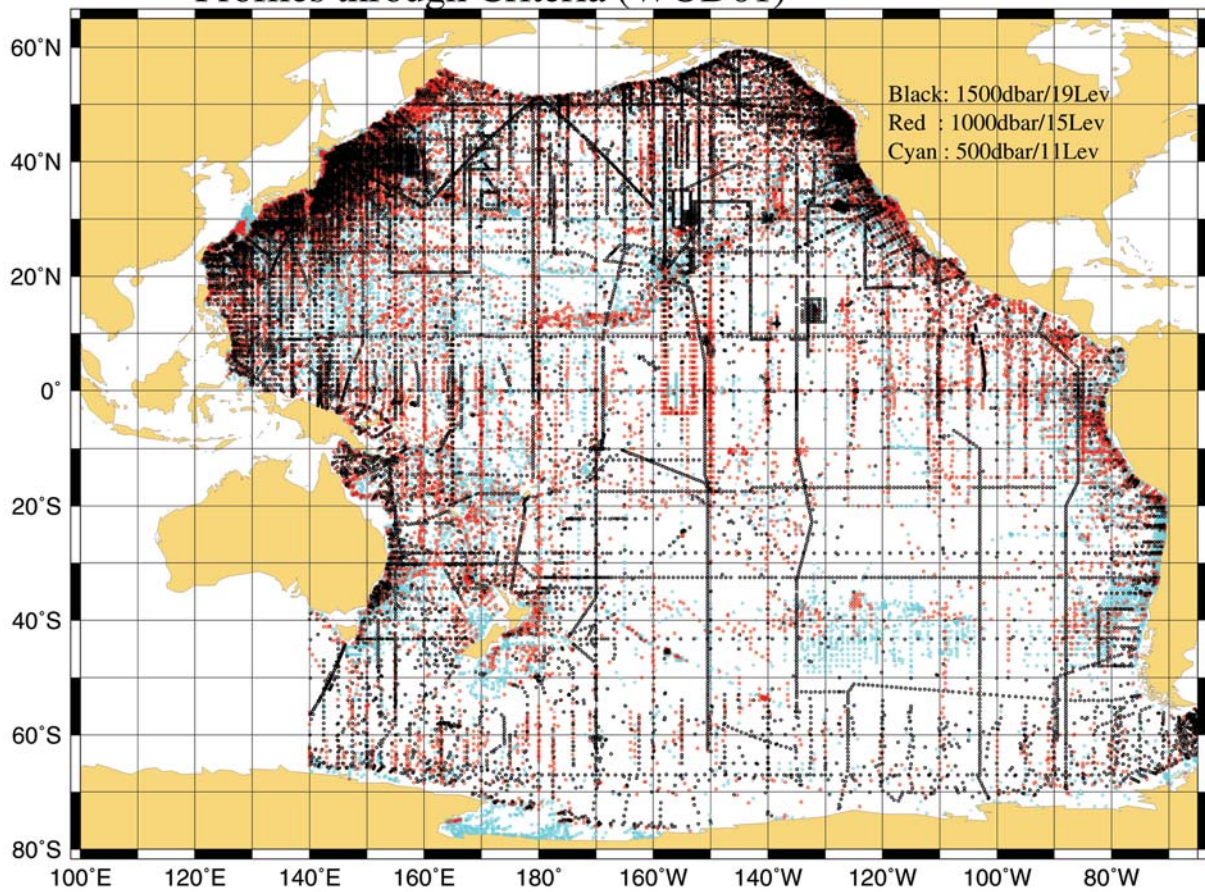


図5 縁辺海を除く太平洋領域における観測データの空間分布。ただし、最大深度が1500dbar、観測層が19層以上のものを黒、1000dbar以上、15層以上を赤、500dbar以上、11層以上を水色で表す。

Fig. 5 Distributions of historical profiles in the open Pacific (WOD01). The profiles of 19 layers or more and the deepest layer exceeding 1500dbar (19 layers / 1500dbar) are shown by black, the profiles of 15 layers / 1000dbar are shown by red, and the profiles of 11 layers / 500dbar are shown by light-blue, respectively.

ものではないかと推測される。北太平洋で最も深層に分布する水温極大はオホーツク海のもので、その深度は900dbar以上である。縁辺海を除くと、西部亜寒帯循環の千島列島沖に分布するものが最も深く、その深度は約800dbarとなっている。なお、World Ocean Atlas 1998には、東部亜寒帯域およびカリフォルニア沿岸域には水温極大が存在せず(たとえばUeno and Yasuda (2000))、分布域の様相はかなり異なっている。

南太平洋の水温極大は、亜南極前線(約45°S)より南の海域に広く分布している。西部のものに比べて、東部の水温極大の方がやや深く分布する傾向を示しており、最も深いものは南米ホーン岬の南西沖に分布する。その深度はおおよそ800dbarで、1000dbarを越えるものもいくつか存在する。

つまり、太平洋では、亜寒帯域の広い領域に水温極大が分布しており、その分布深度は縁辺海を除けば1000dbarよりも浅いと考えるが良い。また、亜熱帯循環域および日本海にも散発的に水温極大が認められる。しかしながら、そのほとんどが1000dbarより深層、もしくは200dbarより表層に分布している。そのため、これらは次のいずれかに分類され

るものと考えられる。

- ・水温の異常データに起因するもの
- ・深層(の水温変化の乏しい層)における観測誤差に起因するもの
- ・混合層(の水温変化の乏しい層)における観測誤差に起因するもの

従って、これらの海域に見られる水温極大には実体はないと考えられる。

2.3. 統一的なデータ選択

図5は縁辺海を除く太平洋領域における観測点の分布を、その選択方法(プロファイルの最大深度および観測層数)を変えて示したものである。最大深度の制約を厳しくすると歴史的データの数は大幅に減少する。遅延品質管理用歴史的データセットに用いるプロファイルとして、その最大深度の基準を水温極大の最大深度(約1000dbar)よりも十分深いもの(たとえば1500dbar)に設定した場合、日本沿岸域と北米大陸沿岸を除けば、ほとんどの海域がデータの空白域になりかねない。そこで、2.1節で準備したデータセットか

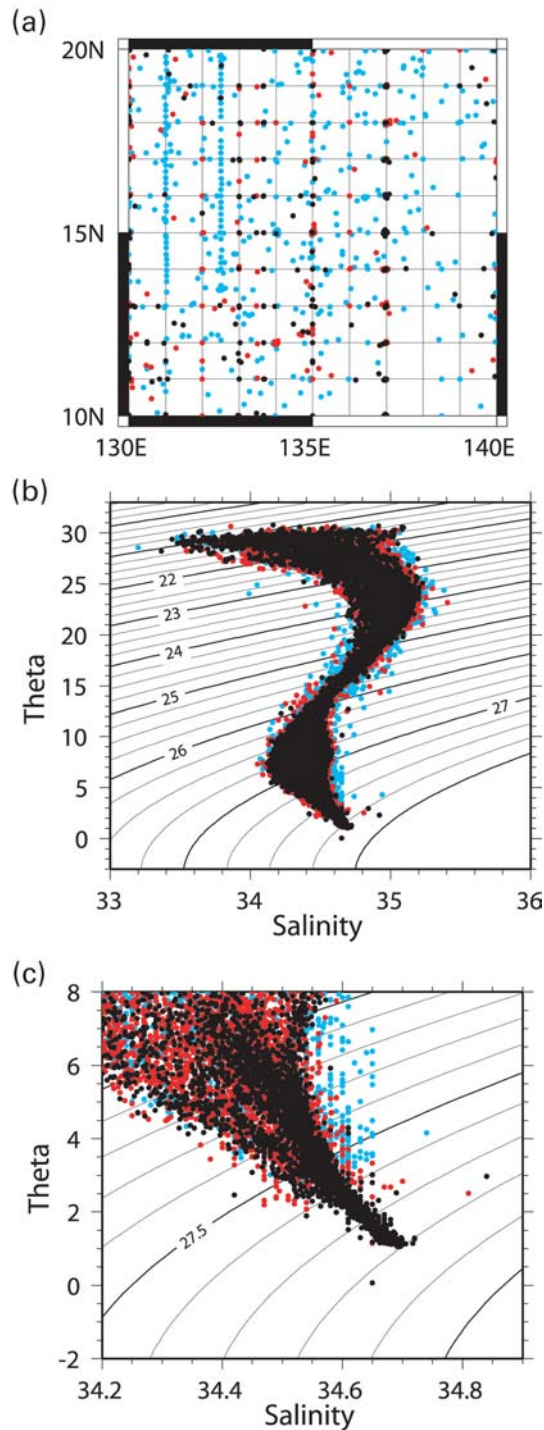


図6 MSN 1113海域 (10-20° N, 130-140° E) における観測プロファイルの (a) 観測点分布, (b) 水温-塩分分布および (c) 深層部の拡大図。ただし, プロファイルの最大深度が1500dbar, 観測層が19層以上のものを黒, 1000dbar以上, 15層以上を赤, 500dbar以上, 11層以上を水色で表す。

Fig. 6 (a) Horizontal distributions, (b) potential temperature - salinity diagram and (c) its enlargement for deep layer of historical profiles in the region of MSN 1113 (10-20° N, 130-140° E). The profiles of 19 layers or more and the deepest layer exceeding 1500dbar (19 layers / 1500dbar) are shown by black, the profiles of 15 layers / 1000dbar are shown by red, and the profiles of 11 layers / 500dbar are shown by light-blue, respectively.

ら, 次の基準を満たしたプロファイルを選択し, 遅延品質管理用歴史的塩分データセットの作成に用いるものとする。

- ・プロファイルの最大深度が1000dbar以上
- ・プロファイルの観測層数が15層以上

なお, 観測層数に課した条件は, 不適切なデータ補間が行われるのを避けるためであり, 1000dbar以浅に海洋観測の標準層が19層設定されていることを念頭に置いている。

SeHyD version 1.0では, データ選択の基準として太平洋全海域で同一のものを適用する。しかし, 以下で示すような理由により, 熱帯や亜熱帯域にはより緩やかな基準(たとえば500dbar/11層)の適用も可能である。

- ・水温逆転層(水温極大)が存在しない。
- ・最大深度が500-1000dbar程度の歴史的観測データが非常に多い。

これにより, データ数が大幅に増加し, より高い精度の品質管理が可能となると期待される。この点については, SeHyDの改訂の際に検討する予定である。

図6は西部熱帯太平洋のMSN 1113海域におけるデータ分布の特徴を示す。データ選択の基準を引き下げることにより, プロファイル数は大幅に増加するが, 同時にデータのバラつきも大きくなっている。これらは, 極めて大きな海洋変動を観測したものというよりも, 不適切な観測に起因していると考えられる。そのため, 2.4節で述べるプロファイルデータの品質管理は一層重要になると予想される。

2.4. プロファイルデータの品質管理

2.1で示したように, WOD01ではNODCの品質フラグを参照したデータ選抜を行うことにより, 品質の極端に悪いデータはあらかじめ排除されている。しかし, Kobayashi and Minato (2003)で示されたように, NODCのチェックをくぐり抜ける低品質データも存在する。遅延品質管理の基準塩分プロファイルを精度良く計算するためには, このような低品質データをあらかじめ削除しておく必要がある。

海洋観測における低品質データの発生原因は, 大まかに分けて次の3種類に分類できるものと考えられる。

1. 採水ミスなど, 観測技術の稚拙さに起因し, 大きなバラつき伴うデータとして現れるもの
2. 測器の劣化や標準海水の不備などに起因し, バイアスを伴って現れるもの。
3. データの読みとりミスや記載・転記ミスなどに起因し, 極めて大きな不連続プロファイルとなって現れるもの

なお, 低品質データの発生原因それぞれが稚拙な観測技術に原因を求められるので, 多くの場合は上記の3つが複合的に絡んでいる。そのため, 低品質データがバイアスを持つものとバラつきが大きいものいづれかに分類できるわけではない。

上記のうち, 原因1に起因する低品質データは, 平均的なプロファイルの周辺で大きなバラつきを示す傾向のものが多い。従って, 統計的な手法を用いた品質管理を施すのが効果的である。しかし, バイアスを伴って現れるもの(原因2)は統計処理に適さない。特に, 海洋観測のようにデー

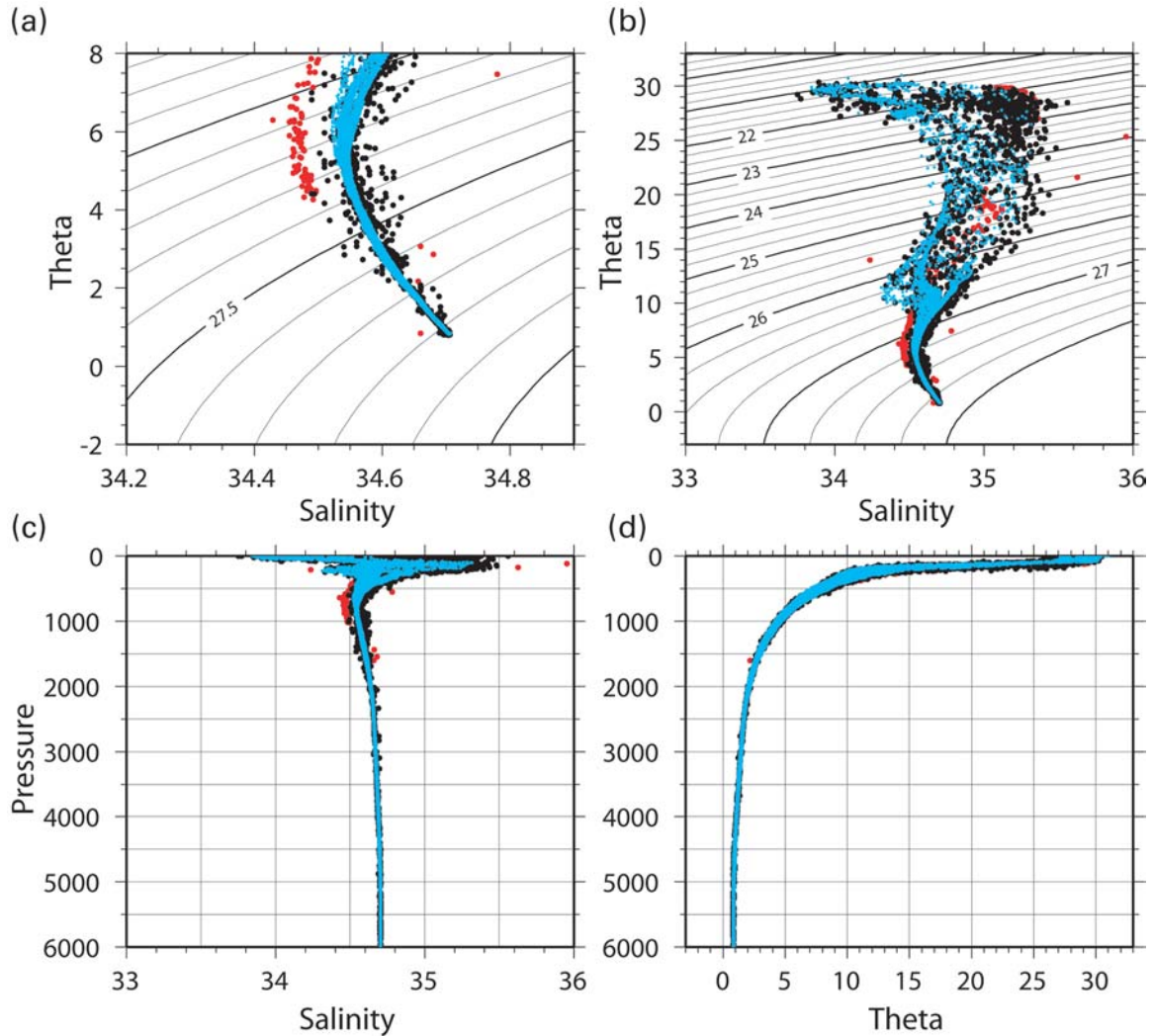


図7 MSN 1017海域 (0-10° N, 170-180° E)にて目視による品質管理を施した結果。(a) 深層および(b) 水塊全体の水温-塩分分布, (c) 塩分-深度構造および(d) 水温-深度構造。赤は削除されたデータを, 黒は選択されたデータを示す。また, 水色のデータはこの海域で得られたWOCE-CTDデータであり, 品質管理用の基準として用いている。

Fig. 7 Quality control by visual inspection for the profiles in the region of MSN 1017 (0-10° N, 170-180° E). Potential temperature - salinity diagrams (a) for deep layer and (b) for the whole water-mass, (c) salinity structure (vs pressure), and (d) potential temperature structure (vs pressure), respectively. The black and red show the data selected and removed, respectively. The light-blue show WOCE-CTD data obtained in the region, which are used as the reference for quality-control.

タ数そのものが少ない場合には, 統計処理では取り除くことができないだけでなく, データ自体が歪んだ統計値を生成して, 誤った処理 (削除すべきでないデータを削除し, 削除すべきデータを削除しない) を行う可能性が高い。原因3に起因する低品質データも, 統計処理の際に(一時的にせよ) 非常に大きなバラつきを示す統計値を与えることになり, 本来行われるべき処理よりも緩い処理が施されることになる。

そこで, あらかじめ目視による品質管理を施してバイアスがかかったデータなどを取り除いた後, 統計的な品質管理を行うことにする。以下では, その具体例を簡単に示す。なお, この品質管理手順はHydroBase作成の手順と同一であり, 将来のPacific HydroBase更新を念頭においている。

2.4.1. 目視による品質管理

図7と図8はMSN 1017海域およびMSN 5214海域にて目視による品質管理を行った例である。この両海域で低品質データとして削除されているデータは, 「ある船舶の同一航海」で得られたデータであり, その航海間の全プロファイルは他のプロファイルと比べて約0.1psu程度低い塩分値を示す。統計処理を用いた場合, 深層部分の低品質データを取り除くことは可能であろうが, 表・中層では他のデータとの区別が難しくなり, 統計処理で取り除くことは極めて困難である(実際, North Pacific HydroBaseにおいてもMSN 1017海域の低品質データは取り除けていない)。

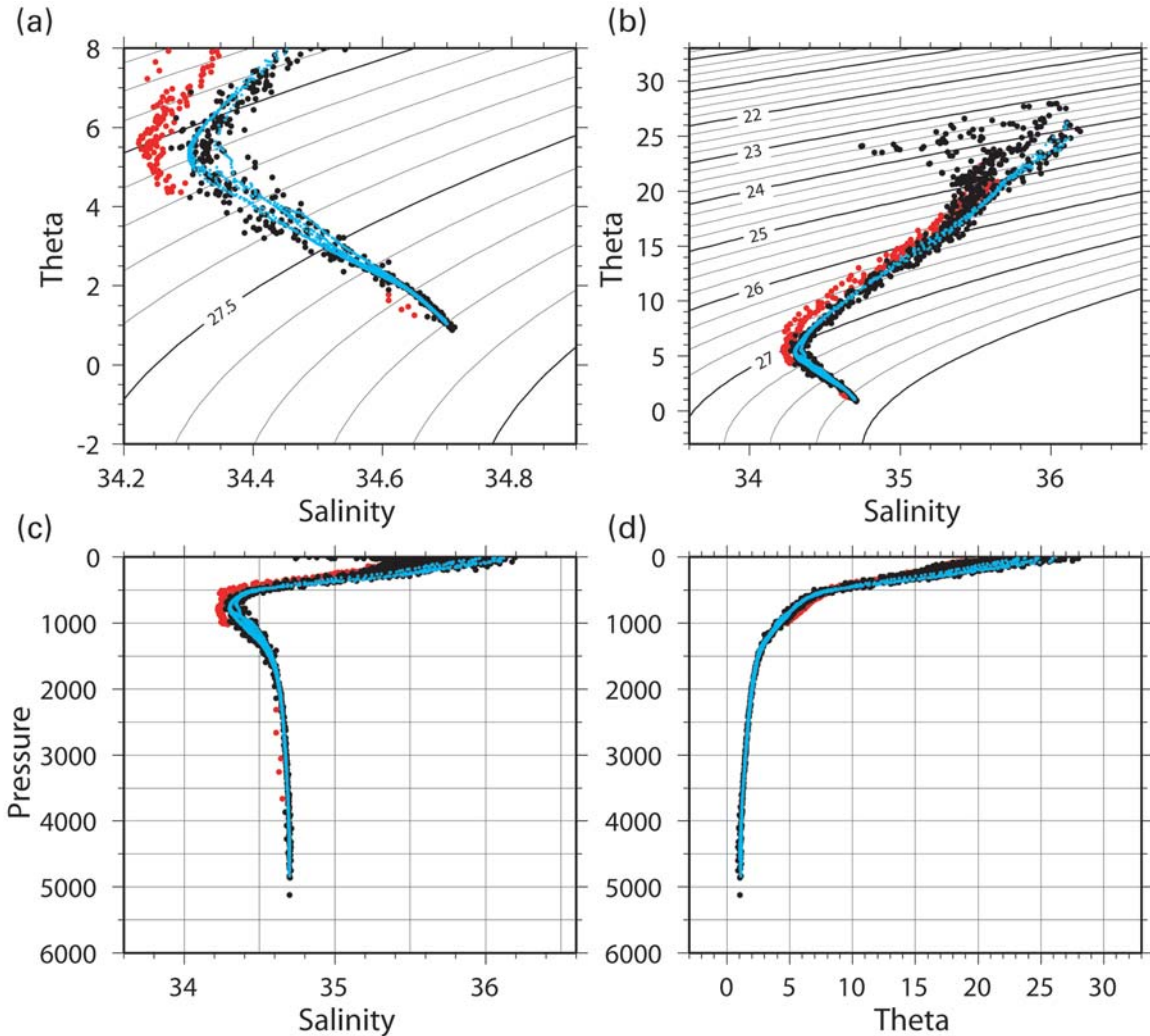


図8 MSN 5214海域 (20-30° S, 140-150° W)にて目視による品質管理を施した結果。詳細は図7に同じ。
 Fig. 8 Quality control by visual inspection for the profiles in the region of MSN 5214 (20-30° S, 140-150° W). Otherwise, as described for Fig. 7.

2.4.2. 統計的な品質管理

統計的な品質管理は、水塊構造の空間分布を考慮して行うことが望ましい (HydroBase型の統計チェック)。深層では水塊構造の空間変化が比較的小さいため、かなり広い空間にわたって共通の統計値を用いることが可能である。しかし、表層では水塊構造の空間変化が激しく、より細分化された空間ビンが必要になる。しかし、選択されたプロファイル (2.3節参照) だけでは、統計処理を行うにデータ数が十分とは言えない。

そこで、SeHyD version 1.0では、深層1000m以深に対してのみ $10^\circ \times 10^\circ$ 格子毎にHydroBase型の統計チェックを施している。表層の統計チェックは、将来行う予定のPacific HydroBaseの作成 (更新) に併せて実施する計画である。

3. 統計情報

結果を示す前に、SeHyDの統計情報として歴史的データの分布密度 (図9) およびノイズ/シグナル比 (N/S比: 図10) を

示す。また参考として、Wong et al. (2003) で作成されたデータセットのN/S比を示す (図11)。N/S比は、データセットから直接計算している。なお付録Bにてシグナルおよびノイズの計算方法を簡単に示した (詳細はKobayashi and Minato (2003) を参照)。

3.1. データ密度分布

データの分布密度は、日本近海および北米大陸沿岸で非常に高い。これらの海域における単位面積あたりのデータ数は、水温面によって若干の差があるが、100個を越えることもある。またハワイ近海や 150° Eよりも西側の北太平洋亜熱帯/熱帯域も比較的多くのデータが分布する。これに対して、南太平洋の中部や東部ではデータ分布密度が極めて低く、1990年代に行われたWOCE観測を除けば、歴史的データはほとんどないと言える。

14° Cの等温面では、南北太平洋ともに 40° よりも高緯度側ではデータ数が激減する。この水温面は 35° 付近にある

亜熱帯前線で海面付近に達している。それよりも高緯度側では季節温度躍層内部にしか存在できず、データ数が減少するものと考えられる。また、低温(2.6°C)の等温面上のデータ数は、亜熱帯に比べて亜寒帯で多くなっている。これは、この等温面の分布深度は亜寒帯域ほど浅く、同規模の観測を行った場合に観測される頻度は亜寒帯のほうが高いためである。たとえば2.6°Cの等温面は、北太平洋の亜寒帯域では800-1200dbarに存在するのに、熱帯や北太平洋亜熱帯域ではより深い1400-1800dbarに分布する。そのため、1000-1500dbarまで観測を行った場合、亜寒帯では最深部が2.6°C等温面に達するが、亜熱帯では難しい。

3.2. N/S比

N/S比の分布(図10)は非常に興味深い分布をしている。14°Cでは日本南方にN/S比の非常に大きな領域が存在している。これは、ちょうど北太平洋亜熱帯モード水の分布域に対応する(水温・塩分の一様な水塊が分布する場合、シグナルは小さくなるのに対し、観測誤差に起因するノイズは小さくならない)。オーストラリアとニュージーランドの間の30°S付近にも、高い値を示す領域が存在するが、これもモード水の分布に(何らかの形で)対応していると考えられる。

太平洋赤道域は、高温層ではN/S比の極めて小さな領域として特徴づけられる。水温の低下に伴い、東部太平洋の陸岸付近からN/S比の高い値が出現し始め、6°C以下では、かなり低温部(少なくとも約2°C)まで、赤道上にほぼ帯状の極めて大きな値の領域が広がっている。

また、N/S比分布はデータセットに残っている低品質データの発見にも資することができる。たとえば図10の2.6°Cや2.2°CにおけるN/S比分布には、15°N、130°W付近に周囲よりも値の大きな、非現実的な楕円形の領域が見られる。このほかにも6°Cの25°S、125°W付近、2.2°Cの20°S、160°W付近にも楕円形の分布が認められる。これは、この中心付近に存在する低品質データのために、現実海洋で現れる以上に大きな観測データの分散やノイズの値が計算され、結果的に非現実的なN/S比の値となって現れる。なお、2.2°Cおよび2.6°Cは、約45°Sよりも南側で水温極大の水温にほぼ相当しており(図4を参照)、同時に海面に露出する可能性がある。そのため、N/S比分布にも大きなバラつきが見られる。

N/S比の空間分布の特徴に関して、SeHyDはWong et al. (2003)のデータセットと定性的な差はない。しかし、品質管理を行っていることを反映して、値は全体的に小さくなっており、特に高N/S比領域で「負のシグナル分散」が計算される部分が非常に少なくなっている。さらに、非現実的な「楕円」領域が少なくなっていること、南米エクアドル・ペルー沖の高N/S比領域が小さくなっていることが挙げられる。

4. 塩分修正のパフォーマンス

太平洋領域で得られたフロートデータに対して、SeHyDを用いて塩分修正を行った結果を図12-17示す。また、比較のため、Wong et al. (2003)による歴史的塩分データセットを用いた塩分修正結果も併せて示す。計算に使用した

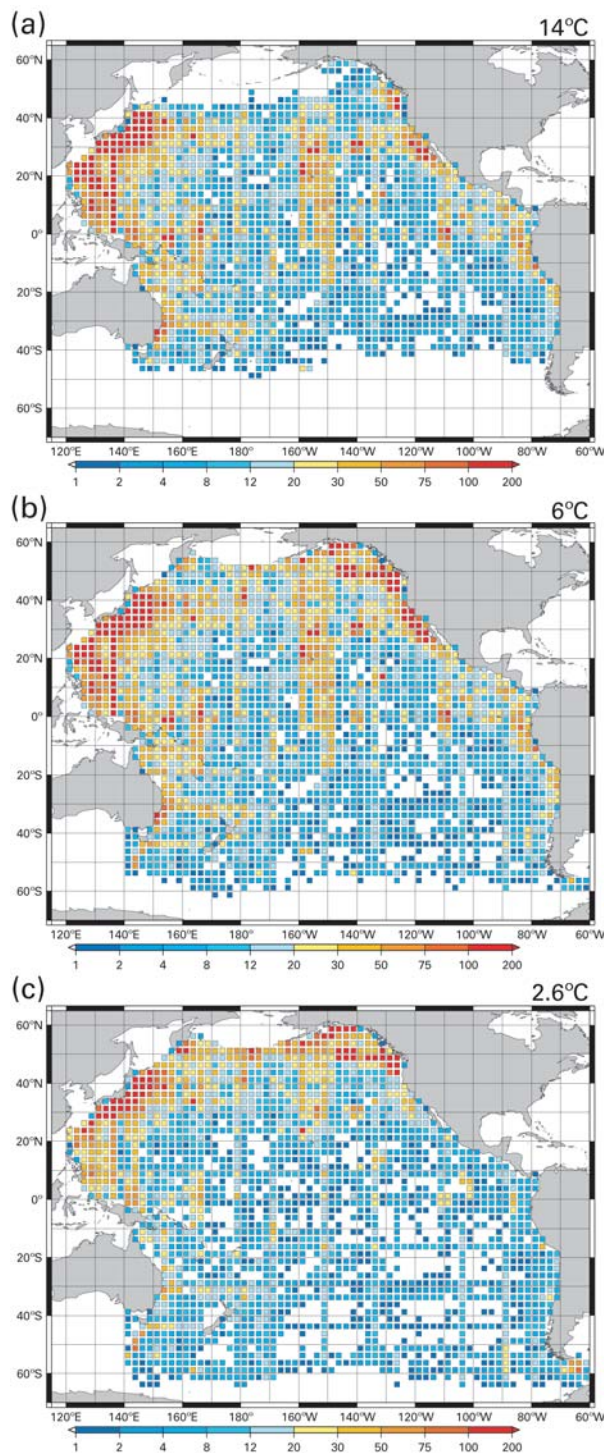


図9 SeHyDにおける、ポテンシャル水温(a) 14°C、(b) 6°C、(c) 2.6°Cの等温面上の2.5°×2.5°格子内の歴史的塩分データの数。
Fig. 9 The number of historical salinity data per a 2.5°×2.5° square on (a) 14°C, (b) 6°C, and (c) 2.6°C in SeHyD.

パラメータは同一である。

日本南岸域に展開されているフロート(WMO 2900186: 図12)の場合、両者のケースともにフロートデータに塩分バイアスがあることは表現されている。ただし、SeHyDの場合には、フロートの投入・回収時のCTD観測結果にほぼ一致す

るのに対して、Wong et al. (2003) のデータセットの場合は修正量がやや小さく、CTD観測とは0.01psu程度のズレが存在する。また西部亜寒帯域 (WMO 2900056: 図13) の場合、修正結果それ自体は両者ともにほとんど差はない。しかし、推定された気候学塩分値は、Wong et al. (2003) のデータセットを使った場合には内挿誤差が非常に大きくなる場合が頻繁にあるが、SeHyDの場合はその頻度が大幅に小さくなる。そのため、修正誤差が小さくなることを期待できる。

それ以外の修正結果についても、両者に若干の違いは認められる。たとえば、WMO 5900110 (図15) では、Wong et al. (2003) に比べてSeHyDの方がその後半部分のフロートデータについてやや高塩分な推定値を示している。またWMO 39034 (図17) では、プロフィール番号60-80にかけて、SeHyDの場合は推定塩分値がほぼ一定であるのに対し、Wong et al. (2003) では一時的な高塩化が見られる。しかしながら、どちらの修正結果が優れているか、という点については不明とせざるを得ない。実際、SeHyDを用いた修正結果の方が良いことがはっきりしているWMO 2900186の場合でも、フロート展開・回収時のCTDデータとの比較によってのみ判断が可能であり、修正結果それ自体での評価は極めて難しい。言い換えれば、SeHyDを用いた塩分修正結果、つまりSeHyDの性能を評価するためには、フロート観測の時空間的近傍において行われたCTD観測データが大量に必要であるといえる。逆説的ながら、このことは、塩分修正のperformanceが非常に高いことを示しており、また修正塩分誤差もほとんどの計算結果で0.01psu以内に収まっている。今後詳細な検討が必要になるが、おそらく修正塩分値と真値とのズレは ± 0.01 psu以内に収まるものと推測される。

また、塩分修正結果は、歴史的塩分データセットとしてWong et al. (2003) とSeHyDのどちらを用いても明瞭な差が現れないことは、大幅な改善を期待していただけに、筆者にとってはかなり残念な結果である。この理由は次の2点にあると考えられる。

1. 2003年4月に大幅に改訂されたWong et al. (2003) のデータセットには、筆者らが2002年9月に第2回アルゴデータ管理チーム会合にて行った研究発表 (Argo Data Management Team 2002; Kobayashi and Minato (2003) の先行結果) が反映されている (Wong 2003, 私信)。
2. Wong et al. (2003) のシステムでは、最適内挿法による気候学的塩分プロフィールの計算には、歴史的塩分データセットから直接計算されたノイズとシグナルの値を用いている。以下では、それぞれの理由について考察する。

4.1. 先行結果の反映

第2回アルゴデータ管理チーム会合において、筆者らは、Wong et al. (2003) のシステムにより効果的に塩分修正を行うためには、優れた歴史的塩分データセットを準備することが重要であるとの発表を行った (Argo Data Management Team 2002)。その際、水温逆転層の処理やデータの補間の観点から、2.3節で挙げた基準を用いてデータ選抜を行って北太平洋の歴史的データセットを試験的に作成し、

非常に良好な結果を得られることを示した。Wong et al. (2003) は新データセット作成に際し、WOD01からのデータ選抜に我々が示した選抜基準を適用している。つまり、Wong et al. (2003) の新データセットとSeHyDは、ほとんど同じデータを格納している。そのため、大きく異なる結果とはならないだろう。また、このことは良好な塩分修正結果が期待できる歴史的塩分データセットの作成は、必ずしも困難な作業ではない可能性の高いことを示唆する。

4.2. 歴史的データセットから直接推定されたN/S比の使用

最適内挿法では、計算に用いるN/S比の大きさによって、推定点の極めて近傍のデータの取り扱いが大きく変わる。N/S比が小さい場合、個々のデータはその海域の状況を良く反映していると考えられるので、推定点の内挿値はその極めて近傍のデータを非常に重視して推定される。逆にN/S比が大きい場合、計算対象海域は、想定する時空間スケールで見るとデータのバラつき具合 (誤差成分) が大きいと判断される。そのため、遠方のデータも相対的に重視し、多数のデータを使って推定点の値を内挿しようとする (このあたりのことは水野 (1995) に詳しい)。

Wong et al. (2003) では、データセットから計算されたN/S比を使って最適内挿を行い、気候学的基準塩分プロフィールを推定している (ただし、シグナルの値として観測値の標準偏差を与えており、本研究よりも簡単な式を用いている。しかし、シグナルを必ず推定できる計算上の利点もある)。その結果、低品質のデータが気候学的塩分プロフィールの計算に用いられた場合、計算海域のN/S比は大きくなり、特定のデータが推定結果に大きな影響を及ぼすことはなくなる。従って、低品質データによって結果が大きく歪められる可能性は、それが少数のデータにとどまる限り、小さい。このことは、アルゴ計画における遅延品質管理システムが、塩分修正の結果に対して「頑丈」であることを意味しており、歓迎すべきことである。

4.3. アルゴ遅延品質管理におけるSeHyD作成の意義

前述したように、塩分修正 (遅延品質管理) を施したアルゴデータには、歴史的塩分データセットとしてWong et al. (2003) とSeHyDのどちらを用いた場合も、明瞭な差は現れなかった。このことは、格納するプロフィールデータに入念な品質管理を施すなど、作成するのにかなりの労力を要するSeHyDの作成意義を再検討する必要に迫られる。

SeHyDは品質管理された歴史的データのみを格納している。その結果、SeHyDには、低品質データに起因した異常なN/S比分布は少なくなっている (最終的には全て取り除くことを目標とする)。このことは、歴史的データセットに新しいデータ、たとえばフロート投入時のCTD観測データなどを加える時に大きな意味を持つ。N/S比の分布が現実的なものであったとき、新しいデータはその海域の特徴に見合った形で参考にされる。多くの場合、追加されるデータはフロート観測点の時間的近傍にあるから、その情報は相対的に重視され、修正結果はこの追加データに対して敏感

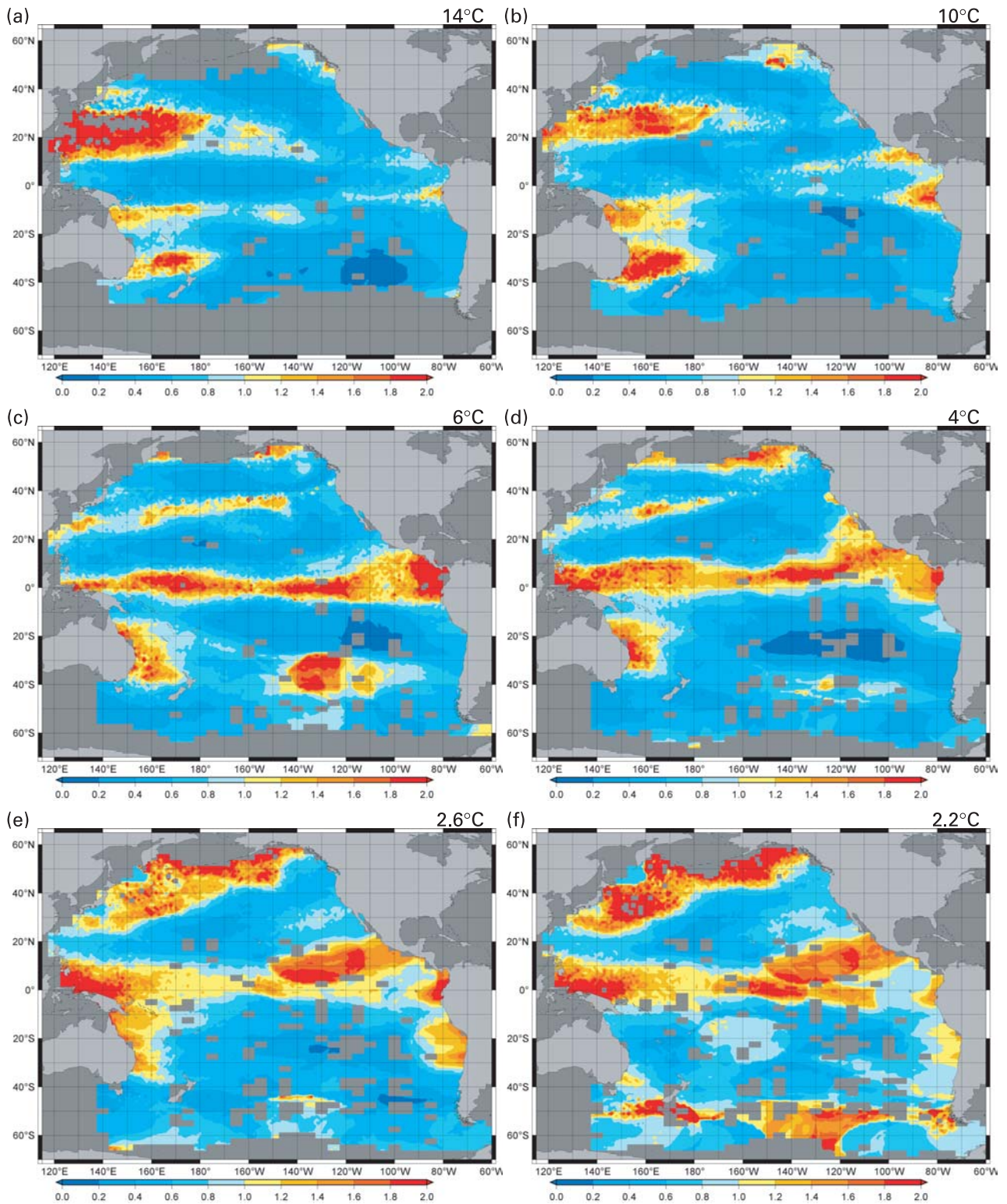


図10 SeHyDにおける、ポテンシャル水温 (a) 14°C, (b) 10°C, (c) 6°C, (d) 4°C, (e) 2.6°C, (f) 2.2°Cの等温面上の塩分データのN/S比分布。灰色のパッチ領域は、次のいずれかに当てはまる。1) 計算上「負のシグナル分散」が得られ、N/S比が推定できない。2) 定義点の周辺2°以内にデータ点がない。

Fig. 10 Distributions of noise/signal ratios in the dataset of SeHyD on the potential temperature of (a) 14°C, (b) 10°C, (c) 6°C, (d) 4°C, (e) 2.6°C, and (f) 2.2°C, respectively. The gray square patches show the regions where noise/signal ratio can not be estimated due to "negative variance of signal" or where estimating location can find no data within 2°.

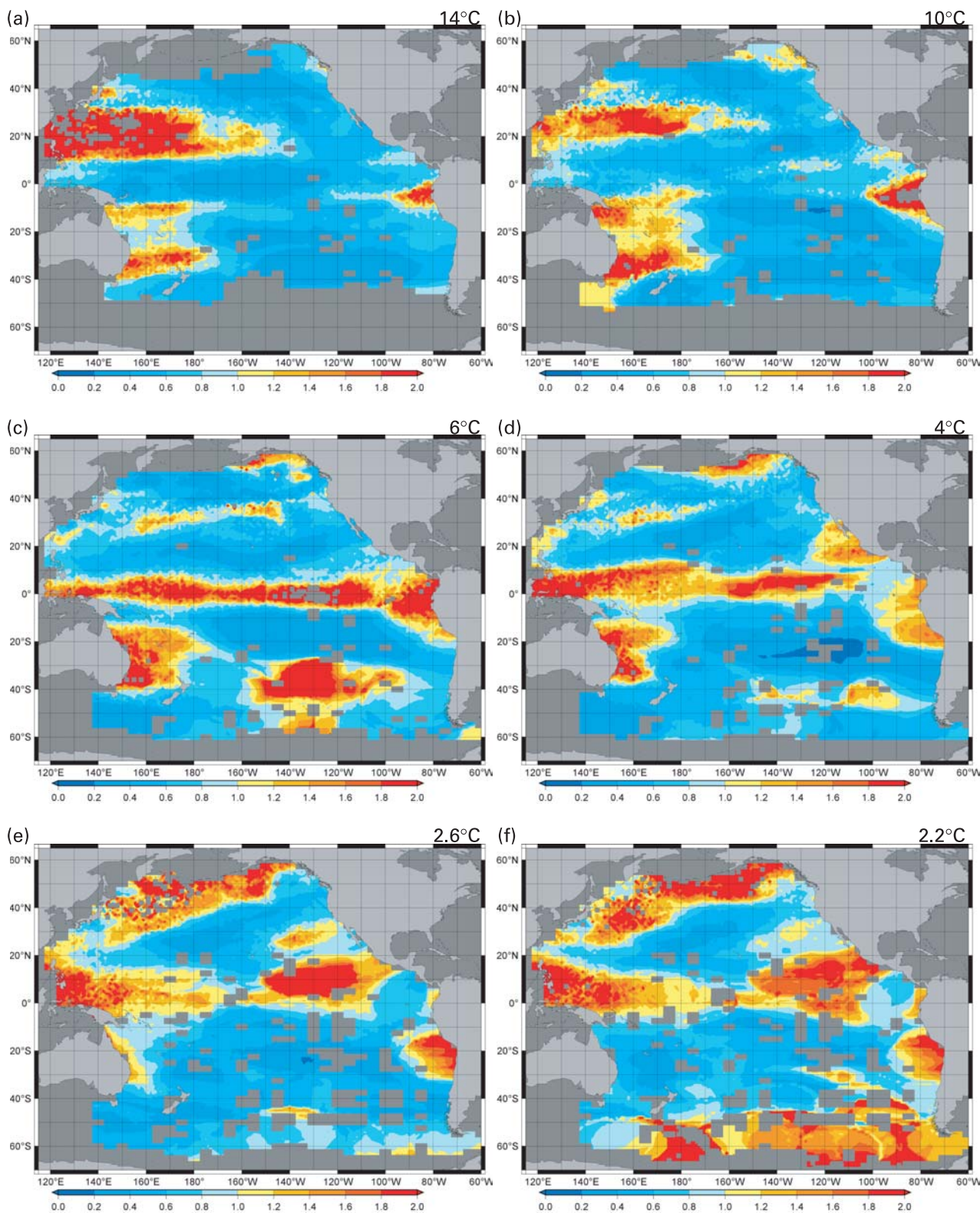


図11 Wong et al. (2003)が作成した歴史的データセット(2003年4月更新版)における塩分データのN/S比分布。詳細は図10に同じ。
 Fig. 11 Distributions of noise/signal ratios in the dataset updated by Wong et al. (2003) in April 2003. Otherwise, as described for Fig. 10.

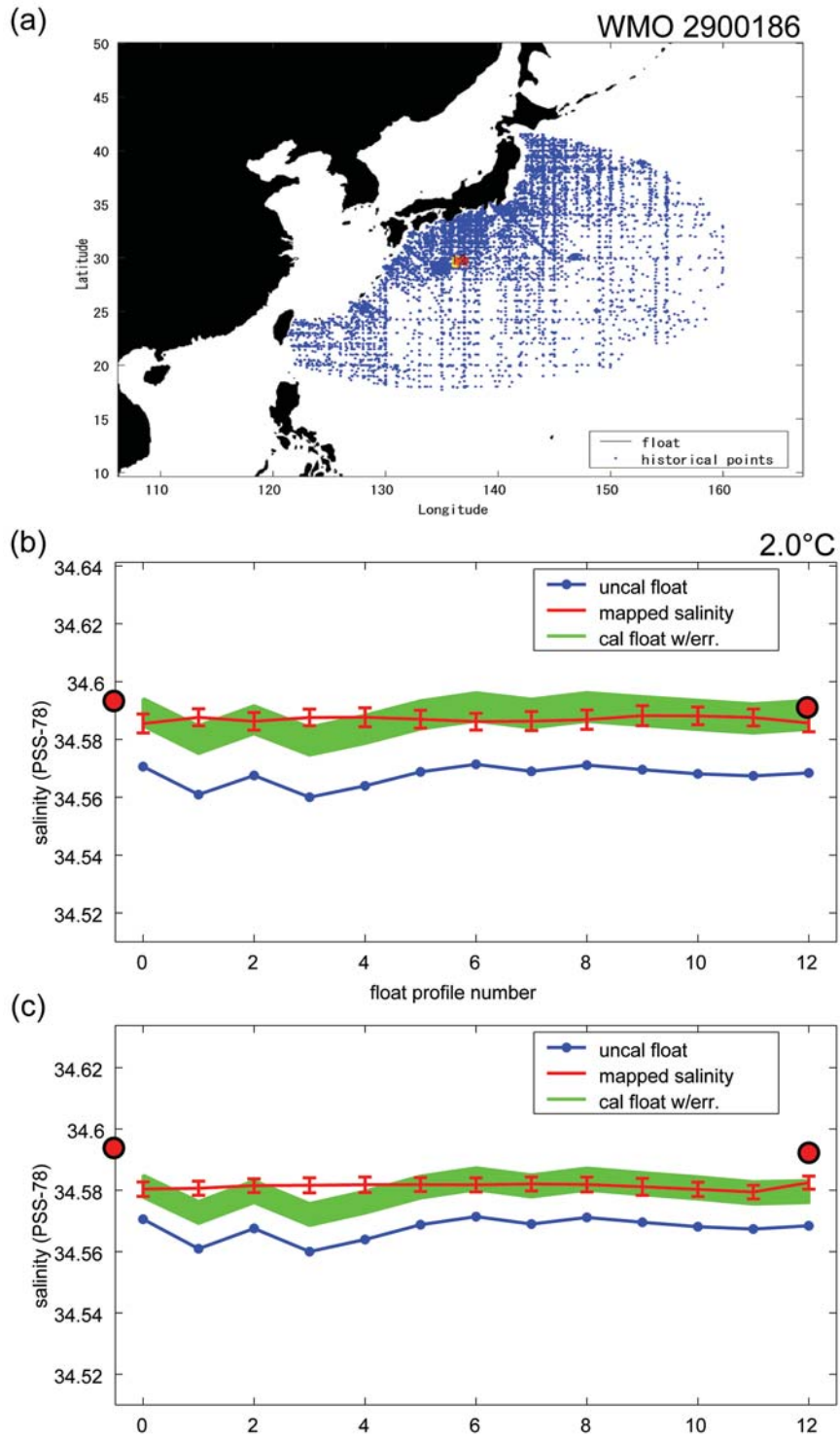


図12 WMO 2900186フロートデータの塩分修正結果。(a)フロート観測点の位置およびSeHyDに格納されている歴史的データの分布。(b)および(c)は歴史的データセットとしてSeHyDもしくはWong et al. (2003)のデータセットを用いた場合の結果(2.0°C等温面)。青はフロート観測値, 赤は推定されたフロート観測点での気候値および内挿誤差を示す。緑帯は修正結果で, 帯の中央が最適修正値, 帯の幅は修正誤差(の2倍)を示す。赤丸は投入・回収時のCTD観測結果。回収後のpost-calibrationもほぼ同じ結果を示す。

Fig. 12 Salinity correction of the float WMO ID 2900186. (a) The locations of float measurements and the historical data in SeHyD. The time series of salinity correction on 2.0°C isotherm with the dataset of (b) SeHyD and (c) Wong et al. (2003). The blue line with circles shows the measured salinity, and green belt represents the corrected salinity (its center represents the optimal values and its width represents correction errors). The red line with vertical bars represents the estimated climatological salinity and its mapping errors, respectively. The red circles represent the "true" salinity obtained by independent CTD measurements at float deployment and recovery. Almost the same value is obtained by the post-calibration in laboratory.

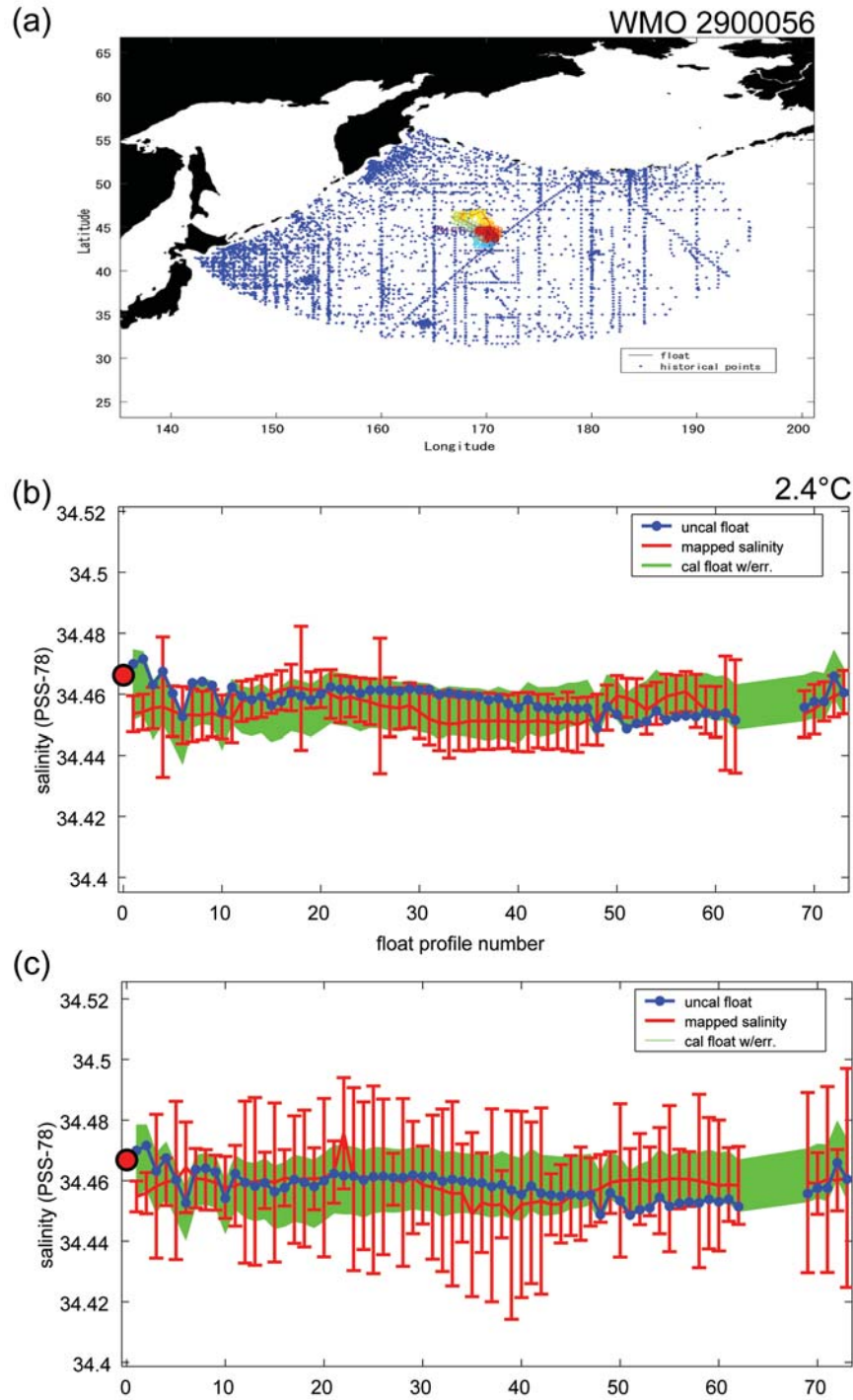


図13 WMO 2900056フロートデータの塩分修正結果(2.4°C等温面)。詳細は図12と同じ。
 Fig. 13 Salinity correction of the float WMO ID 2900056 on 2.4°C. Otherwise, as described for Fig. 12.

に反応する。しかし、低品質データが含まれているなどしてN/S比が極めて大きな値を示している場合、多数のデータを使って気候値の推定を行うので、個々のデータの重要度は相対的に低下する。そのため、新しいデータを導入しても、そのデータによって修正結果は大きく変わることはない。つまり、今後のデータセット拡充に対しては、歴史的塩分データセットとしてSeHyDを用いていることが有利に働くと予想される。

この点については、別の観点からの指摘もある。観測データが非常に少ない海域では「歴史的データによるアルゴデータの品質評価」を行うこと自体が困難であり、フロート投入時などに行われるであろうCTD観測データを用いる(歴史的データセットに加える)必要があることが指摘されている(Brian 2003, 私信)。また、WOD01などのデータセットに格納されていない、最近行われた観測データをデータセットに組み込むことによって、塩分修正の結果が良くなる(可能性

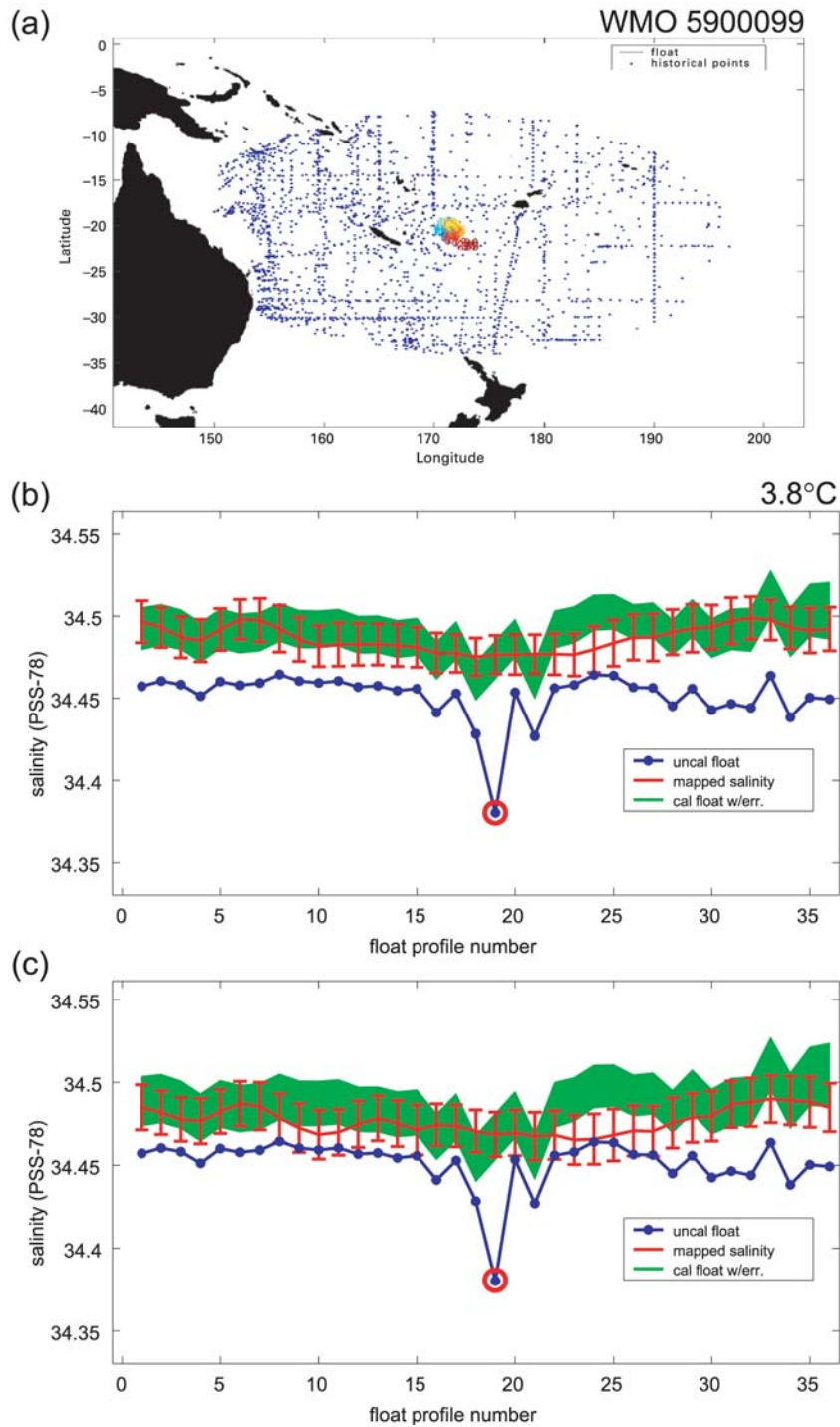


図14 WMO 5900099フロートデータの塩分修正結果(3.8°C等温面)。詳細は図12と同じ。赤丸を付けた塩分データ(プロファイル番号19)は、1プロファイルのみの塩分ジャンプのためにデータ修正ができない。

Fig. 14 Salinity correction of the float WMO ID 5900099 on 3.8°C. The salinity data observed in profile No. 19 (marked by red circles) can not be corrected because Wong's system can not deal with "one profile salinity jump". Otherwise, as described for Fig. 12.

が高いことも指摘されている(Gould 2002, 私信; Desaubies 2003, 私信)。さらに、最近の研究によれば、南大洋では10年程度の時間スケールの水塊変動が生じていることが明らかになりつつある(Brian 2003, 私信)。しかし、南大洋は歴

史的データが少ない上、水塊変動の時間スケールに比べて観測の行われた時間間隔が広く、歴史的データとの比較によって結果が歪んでしまう可能性も指摘されている(Brian 2003, 私信)。

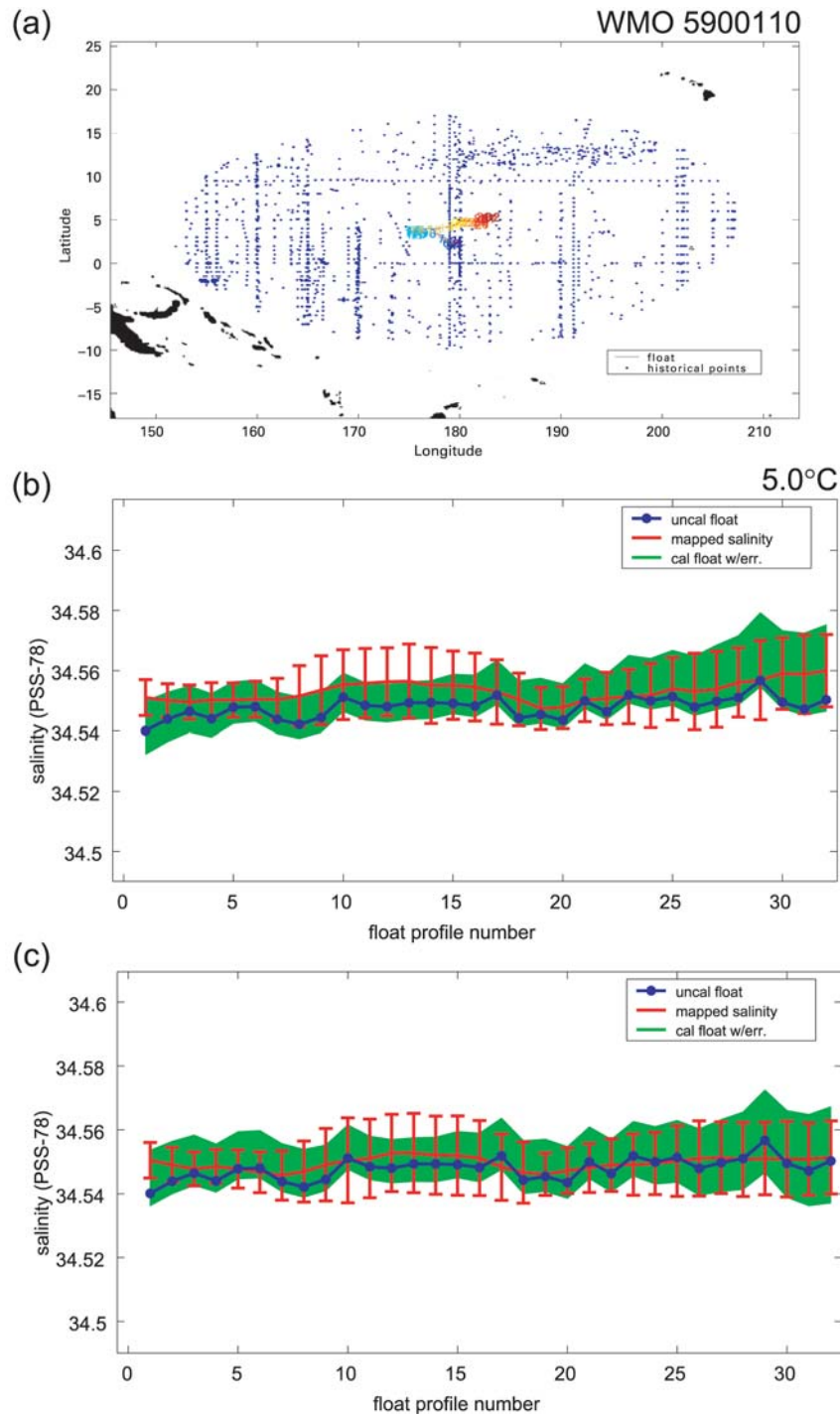


図15 WMO 5900110フロートデータの塩分修正結果 (5.0°C等温面)。詳細は図12と同じ。
 Fig. 15 Salinity correction of the float WMO ID 5900110 on 5.0°C. Otherwise, as described for Fig. 12.

つまり、SeHyDは遅延品質管理の結果を、新たに歴史的塩分データセットに加えられた最新の観測データに敏感に対応させることができるので、アルゴ遅延品質管理の性能(performance)の向上に大きく寄与すると言える。

5. 今後の計画

現在、SeHyDは縁辺海を含まない太平洋版のみが完成

している。しかしながら、縁辺海(日本海・ベーリング海・オホーツク海・南シナ海)においてもアルゴ計画およびArgo equivalentのフロートが展開されている(いた)。これらのフロートデータについても品質管理は必要であり、早急な歴史的塩分データセットの作成が必要である。

また、SeHyDでは日本近海および北米大陸沿岸ではデータセットの品質の面からmodHBを使っている。しかし、

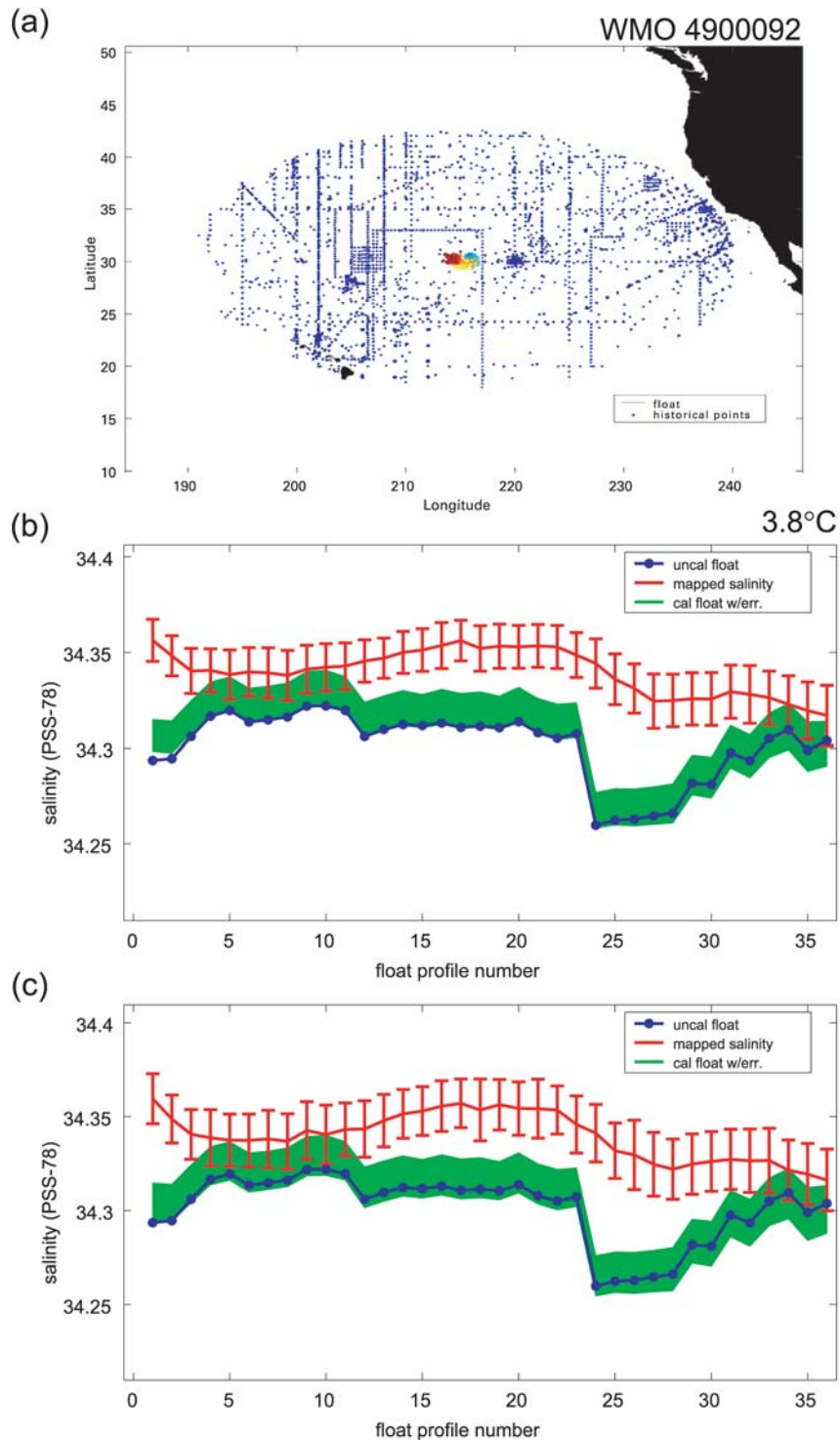


図16 WMO 4900092フロートデータの塩分修正結果 (3.8°C等温面)。詳細は図12と同じ。
 Fig. 16 Salinity correction of the float WMO ID 4900092 on 3.8°C. Otherwise, as described for Fig. 12.

データセットの拡充と言う点では、十分な品質管理を施した上でデータ数の多いWOD01を使うことが望ましい。また2.3節で述べたように、太平洋の熱帯・亜熱帯領域では、プロフィールデータに対して十分な品質管理がなされていれば、プロフィールの選択基準を緩和することによってデータの格納数を増やすことができる。これら2点による遅延品質管

理のperformanceの向上を目的として、WOD01全体にHydroBase方式の品質管理を施してPacific HydroBaseを作成・完成させた上で、SeHyDの更新・品質向上を図る計画である。

なお、作成されたSeHyDは、現在JAMSTEC-Argo Web サイト上にて公開されている。また、同サイト上にて今後も

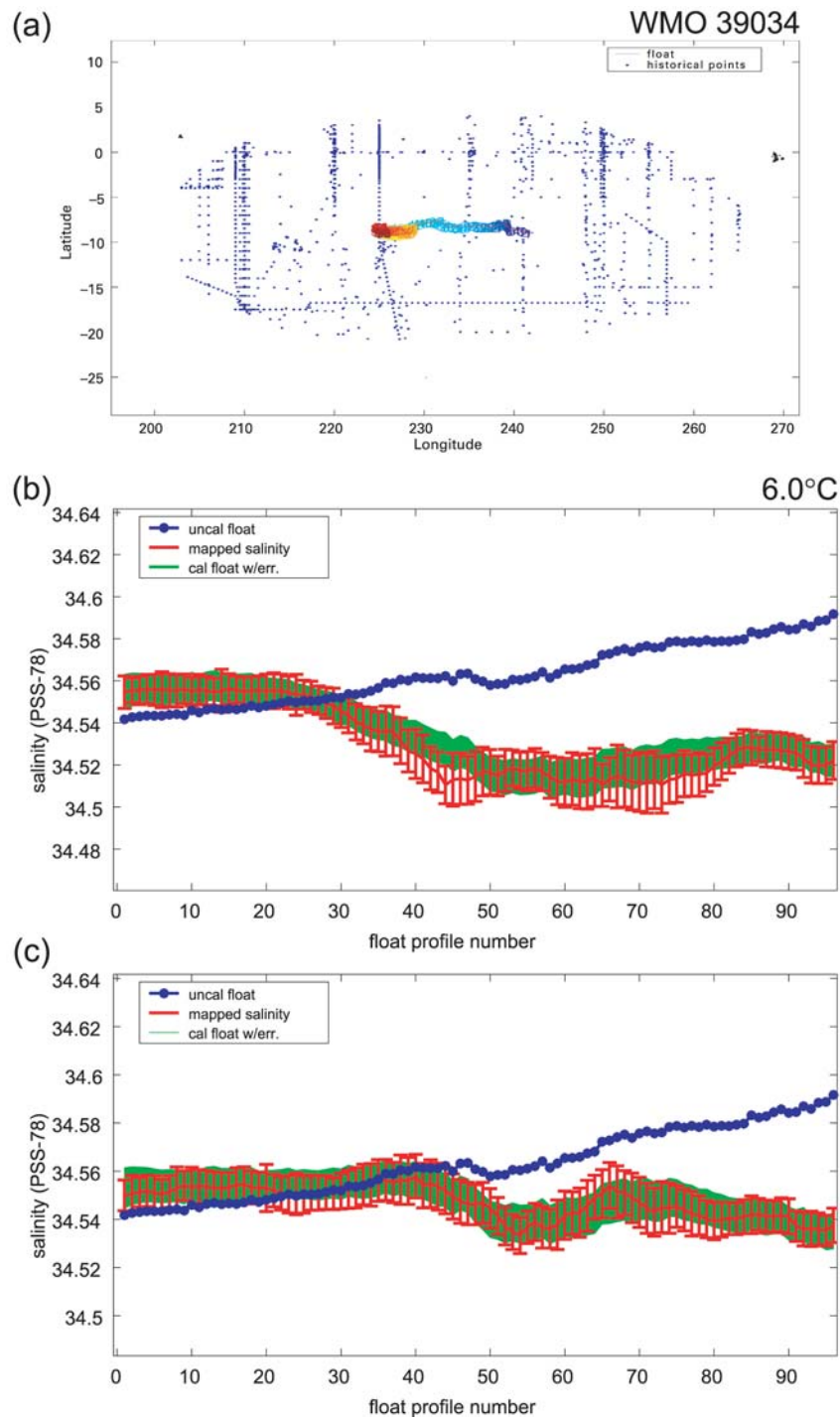


図17 WMO 39034フロートデータの塩分修正結果(6.0°C等温面)。詳細は図12と同じ。
 Fig. 17 Salinity correction of the float WMO ID 39034 on 6.0°C. Otherwise, as described for Fig. 12.

最新版を公開する予定である。

ダウンロードサイト(2003年11月末現在; SeHyD version 1.0)
 : http://www.jamstec.go.jp/ARGO/product/SeHyD_1.zip

付録A: 水温逆転層により引き起こされる問題

Wong et al. (2003)による塩分修正には、歴史的塩分データセットとして、所定水温における塩分値データセットを準備する必要がある。データセットに格納された各データは、

各歴史的プロファイルのスプライン補間することにより計算される。プロファイル中に同一水温が複数存在する場合、最深層における塩分値を用いる(Wong et al. 2003)。通常、海洋は温度成層をしているので、これは水温逆転層が含まれている場合の処理に相当する(なお、筆者が独自に調べた結果、水温逆転層では塩分データを選択しないようなアルゴリズムが用いられていることが判明した)。この処理は、水温逆転層が存在する海域で、深層における基準塩分プ

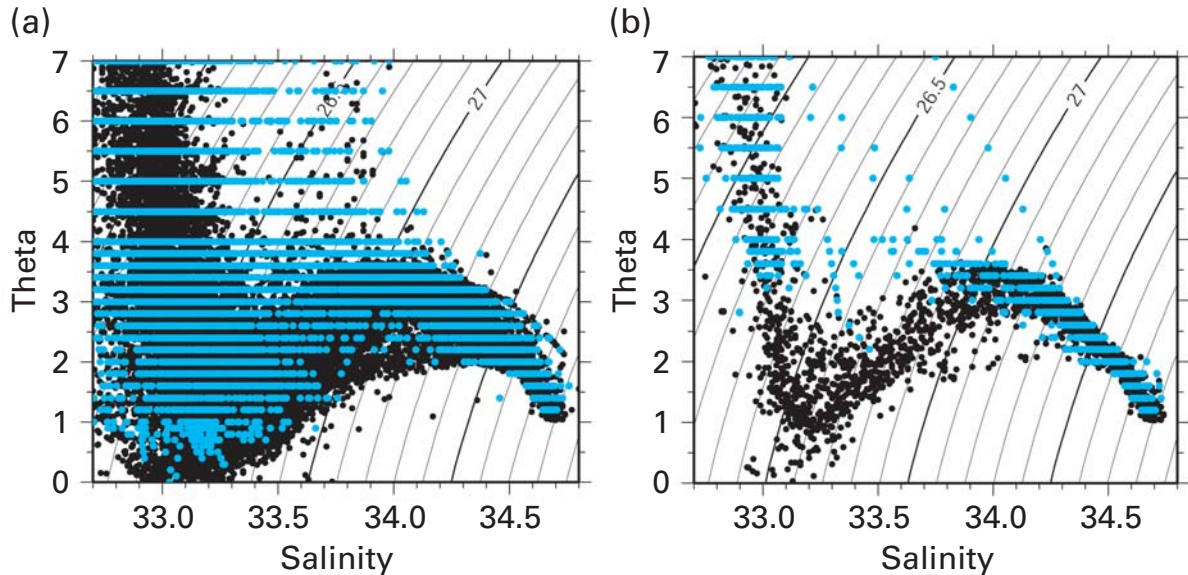


図18 オホーツク海を除く44-54°N, 150-160°Eの領域において求められた歴史的塩分データの水温-塩分分布(青)。ただし、元になる観測プロファイルの最深層が(a) 100dbar以深の場合と(b) 2000dbar以深の場合。黒は計算に用いたプロファイルデータ。
 Fig. 18 Potential temperature - salinity diagram of historical salinity data in the region of 44-54°N, 150-160°E except for the Sea of Okhotsk, which are calculated from the original profile data (the black dots), the deepest measurements of which exceed (a) 100dbar or (b) 2000dbar, respectively.

ロファイルを作成するに当たり、(水温のほぼ同じ)表層近くの水塊の影響を受けないようにすると言う意味で、極めて妥当なものである。

しかし、水温逆転層を含むプロファイルから深層部分を徐々に取り除いていったときに、このアルゴリズムによってどのような塩分データが選択されるかという問題を考えると、このアルゴリズムの持つ問題点がよく分かる。それを端的に示したのが図18である。十分深層まで観測が行われた歴史的プロファイルだけをを用いた場合、水温2.5°C以下の深層部分の基準データとして北太平洋深層の水塊を示す塩分データのみが選択されている。これに対し、短いプロファイルの使用を許可した場合は、34psuよりも低塩の表層塩分データが低温層のデータとして多数選択されている。これは、プロファイルが水温逆転層をすべて含んでいるかどうかを、アルゴリズム自体が判別できないためである。つまり、塩分データセットの作成に用いるべき歴史的プロファイルは、十分に深層まで観測されていること、少なくとも水温逆転層をすべて含んでいることが必要である。しかし、十分に深層まで達する観測データの数は少ないので、条件を厳しくしすぎると最適内挿法による推定精度は低下する。そのため、水温逆転層の底、つまり水温極大層の分布の様子(海域・深度)を十分に調べた上で、なるべく多くの歴史的プロファイルデータを使用できるように、選択基準を定める必要がある。

付録B: ノイズおよびシグナルの計算について

ノイズおよびシグナルの計算に用いられるデータは、推定点から東西20°, 南北10°の楕円形の領域の中から、次の基

準に従って600点を選択する。なお領域内のデータが600に満たない場合、すべてのデータを選択する。

- ・まず、領域内からランダムに200点を選択
- ・残りのデータから、推定点から(楕円の形状を考慮して)距離的に近いもの400点を選択

なお、ランダムにデータを選択しているのは、データが特定のもの(たとえば定点観測や特定の観測航海)に偏るのを防ぐためである。

観測データの分散は、シグナルの分散とノイズの分散の和である(たとえば水野(1995))から、シグナルの分散は次のように求められる。

$$\text{Signal variance} = \frac{1}{n} \sum_{i=1}^n (d_i - \langle d \rangle)^2 - \text{noise variance}.$$

ここで、 n はデータ数であり、 $\langle d \rangle$ はデータ d_i の平均値である。

$$\langle d \rangle = \frac{1}{n} \sum_{i=1}^n d_i$$

ノイズの分散は、データの構造関数を仮定し、そのデータ間の距離を0にすることで近似的に求めることができる。実際の計算には以下の式を用いる(Wong et al. 2003)。

$$\text{Noise variance} = \frac{1}{2n} \sum_{i=1}^n (d_i - d_j)^2$$

ここで d_j は、データ d_i から最も近い位置にあるデータである。これらによりノイズおよびシグナルが求められるので、N/S比をデータセットから直接計算することができる。

なお、データの分布が小さくなると、ノイズにはgeophysical noiseと呼ばれる、場の空間変化の影響が混入するので、実際の値よりも大きくなる。そのため、観測データのまばらな海域では、実際よりも大きなN/S比が推定される傾向が強く、場合によっては「負」のシグナル分散が得られることもある(詳細はKobayashi and Minato (2003)を参照)。また、Wong et al. (2003)ではシグナルとして観測データの標準偏差の値を用いている。

引用文献

- 1) Argo Data Management Team, "Report of Argo Data Management 2nd Meeting", Argo Data Management 2nd Meeting, 18-20 September 2002, Marine Environmental Data Service, Ottawa, Canada (2002).
- 2) Kobayashi, T. and S. Minato, "Performance of Argo automatic delayed-mode salinity calibrations and suggestions for historical database improvements", J. Atmosph. and Oceanic Tech. (submitted) (2003).
- 3) 小林大洋, 市川泰子, 高槻靖, 須賀利雄, 岩坂直人, 安藤健太郎, 水野恵介, 四竈信行, 竹内謙介, "高品質気

候学データセット(HydroBase)を用いたアルゴデータの品質管理 I", 海洋科学技術センター試験研究報告, 44, 101-114 (2001).

- 4) Macdonald, A. M., T. Suga, and R. G. Curry, "An isopycnally averaged North Pacific climatology", J. Atmosph. and Oceanic Tech., 18, 394-420 (2001).
- 5) 水野恵介, "大規模な海洋変動の研究に必要な水深/水温データの品質管理について", 遠洋水産研究所研究報告, 32, 147-171 (1995).
- 6) Ueno, H. and I. Yasuda, "Distribution and formation of the mesothermal structure (temperature maximum) in the North Pacific subarctic region", J. Geophys. Res., 105, 16 885-16 898 (2000).
- 7) Wong, A. P. S., G. C. Johnson, and W. B. Owens, "Delayed-mode calibration of autonomous CTD profiling float salinity data by θ -S climatology", J. Atmosph. and Oceanic Tech., 20, 308-318 (2003).

(原稿受理:2003年11月25日)