1M3-OS-24b-1in

系列ラベリング手法による海洋観測データの良否識別

Error Detection of Oceanic Observation Data Using Sequential Labeling

松山 開 *1 Haruki Matsuyama　田中舜也 *1 Syunya Tanaka　西元千恵 *1 Chie Nishimoto　小野 智司 *1 Satoshi Ono　福井 健一 *2 Ken-ichi Fukui　細田 滋毅 *3 Shigeki Hosoda

*1鹿児島大学大学院 理工学研究科 情報生体システム工学専攻
Department of Information Science and Biomedical Engineering, Graduate School of Science and Engineering, Kagoshima University

*2大阪大学 産業科学研究所
The Institute of Scientific and Industrial Research, Osaka University

*3独立行政法人 海洋研究開発機構
Japan Agency for Marine-Earth Science and Technology

This paper proposes a method for error detection on observation data of globally-covered ocean monitoring system Argo. The proposed method utilizes Conditional Random Field (CRF) to assign quality labels for observed temperature and salinity values in each depth. This paper also proposes a method for feature design support in CRF using Support Vector Machine (SVM).

1. はじめに

全球海洋監視システム「アルゴ」では，海洋の水温と塩分の自動計測が可能なアルゴフロートを投入することにより，全球海洋データのリアルタイムな観測を実現している．アルゴフロートにより自動観測されたデータには観測エラーを含むことがあるため，観測値ごとに観測の信頼度を表す品質管理フラグを割り当てている．品質管理フラグは初めに各国のデータ管理機関が定めた品質管理手法に則り，自動的に割り当てられる [1]．しかし，水温や塩分などは自然変動の影響により変化するため，観測エラーとの切り分けが困難であり，既存の自動品質管理では観測エラーの見落としや誤検出が発生している．そのため，最終的に専門技術者が目視で確認を行い，手動で補正を行わなければならない状況にある．また，専門技術者ごとの補正基準の相違や心的要因による判断の揺らぎ，補正スキルを持ち合わせていない国の存在などから，全球データの品質の均一性が担保できないという懸念も生じる [2]．

本研究では，海洋データの観測エラーの検出および指標決定の問題を系列ラベリング問題として捉え，条件付き確率場 (Condtional Random Field:CRF) により自動識別する方式を提案する．一方，CRF に与える素性の設計には，問題に対する経験的知識が必要であり，特に閾値を考慮する場合，経験的知識がない限り困難である．そこで，素性の条件式に機械学習を組み込むことで，閾値設定を不要にした素性の設計を可能とする．北太平洋海域の観測データを用いた実験により，従来の自動品質管理方式や，ラベルの組み合わせを考慮せずに各層毎に推定を行う方式と比較し，提案方式の精度および閾値調整の自動化について基本的な有効性を示す．

2. 全球海洋監視システム「アルゴ」

2.1 概要

全球海洋監視システム「アルゴ」は,「アルゴ計画」のもとで運営される国際プロジェクトであり，2000 年より全球海洋データのリアルタイムな取得を目的として開始された [3]．このプロジェクトでは，全球アルゴ観測網を実現するために，アルゴフロートと呼ばれる自動昇降可能な海洋観測ロボットを海へ投入し，海洋内部の水温や塩分の自動観測を行っている．観測データが衛星を介して地上局に送られると，品質管理を経てインターネットを通じて公開される．現在，このプロジェクトには全世界で 30 カ国以上が参加し，アルゴフロートは 3,500 台以上が常に稼動している．現在までに 100 万点を超えるプロファイルの蓄積に成功しており，従来知り得なかった地球規模の変動が捉えられ，気候変動のメカニズム解明に向けて研究が進められている．

連絡先: 松山開，鹿児島大学大学院 理工学研究科 情報生体システム工学専攻，〒 890-0065 鹿児島市郡元 1-21-40，k4391399@kadai.jp

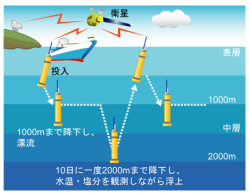


図 1: 観測サイクル

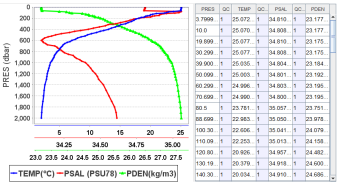


図 2: プロファイルの例

2.2 海洋観測

アルゴフロートによる観測サイクルを図 1 に示す．アルゴフロートは海に投入されると，海流の影響が弱い水深 1,000[m] で漂流する．観測時になると，水深 2,000[m] 付近まで降下し，水温と塩分を観測しながら海面まで浮上する．1 回の浮上によって生成される観測データはプロファイルと呼ばれ，プロファイルには各観測層における水温値および塩分値が記録される．アルゴフロートは，この観測サイクルを 10 日間隔で自動的に行う．プロファイルの例を図 2 に示す．縦軸は圧力 [dbar] であり，これは水深 [m] とおよそ同義である．青色のグラフは水温 [℃]，赤色のグラフは塩分 [PSS-78]，緑色のグラフは水温と塩分から算出される密度 [kg/m^3] を表す．

2.3 品質管理

アルゴにおける品質管理では，アルゴフロートによって観測されたプロファイルに対して観測値の信頼度を決定する．品質管理フラグの種類として，1(正)，2(おそらく正)，3(おそらく誤)，4(誤) の 4 段階の信頼度が使用される．

品質管理には，即時品質管理 (Real-time quality control: RQC) と遅延品質管理 (Delayed-mode QC: DQC) の 2 種類

がある．RQC はプロファイルが観測されてから 24 時間以内に公開することを目的として行われる簡易的な品質管理である．リアルタイムなデータの公開を優先するため，技術者による目視確認は基本的に行われない．一方，DQC は研究や解析用のデータとして提供することを目的として行われる高精度な品質管理である．専門技術者により観測されたプロファイルや，近傍で観測されたプロファイルとの比較などを含めて目視確認が行われ，必要に応じて補正が行われる．

2.4 エラーの種類

2.4.1 密度逆転

密度逆転は，ハードウェア・ソフトウェアの問題，汚濁物質や生物の付着などの外的要因などにより発生する観測エラーである．密度は水温，塩分によって決定される値で，海域によらず深度とともに単調増加するという傾向がある．そこで，ある閾値より大きな鉛直密度の重軽の関係の逆転を検知することで観測エラーを検出し，その深度を特定する．密度逆転は，水温または塩分のどちらか一方の観測エラーによって引き起こされることが多く，上下層の値の関係で決まる．

2.4.2 同値エラー

アルゴフロートが観測する際に，電圧低下により観測が行われず，直前に観測した値がそのままコピーされることがある．すなわち，連続した観測層で同じ観測値が格納されていることになる．本稿ではこの観測エラーを同値エラーと呼称する．

2.4.3 オフセット

オフセットは過去のプロファイルや近傍で観測されたプロファイルと比較したとき，観測量のグラフ全体が平行移動したような観測値が得られる観測エラーである．観測層全体に現れる場合や，深層のみに現れる場合もあり，自動品質管理では検出対象外とされている．専門技術者が過去や近傍のプロファイルとの比較を行うことでオフセットの発生を把握することができるが，即座に解決することは困難である．

3. 提案手法

3.1 概要

本研究では，プロファイルに対する品質管理フラグの良否識別を系列ラベリング問題と捉え，条件付き確率場（Conditional Random Field: CRF）[4] を用いて解く方式を提案する．隠れマルコフモデル（Hidden Markov Model: HMM）[5] と比較して，CRF は様々な素性を同時に考慮できる点に特徴がある．一方，アルゴフロートの観測データは実数であるため，素性関数で適切な閾値が必要となる．このため，本研究では素性関数の閾値をサポートベクタマシン (Support Vector Machine: SVM) [6] により自動的な決定を試みる．

3.2 条件付確率場（CRF）

CRF は系列ラベリング問題に用いられる識別モデルであり [4]，系列データやラベルの前後の依存関係の特徴を素性として記述できるため，多様な属性を参照しつつ，前後のラベルの組み合わせを考慮した設計を行うことができる．CRF ではラベル間の依存関係に応じて様々なモデルが存在するが，本稿では入力データと 1 つ前のラベルに依存する Linear-chain モデルを前提に説明する．

系列の長さが T である入力データ $\boldsymbol{x} = (x_1, x_2, \cdots, x_T)$ が与えられたとき，出力ラベル $\boldsymbol{y} = (y_1, y_2, \cdots, y_T)$ となる条件付き確率 P を式 (1) のようにモデル化する．

$$P(y_t, y_{t-1}|\boldsymbol{x}) = \frac{1}{Z_{\boldsymbol{x}}} \exp\left(\sum_{t=1}^{T}\sum_{k} \lambda_k f_k(\boldsymbol{x}, y_t, y_{t-1})\right) \quad (1)$$

ここで，f_k は素性関数，λ_k は素性関数 f_k に対する重み，$Z_{\boldsymbol{x}}$ は $\sum_{\boldsymbol{y}} P(y_t, y_{t-1}|\boldsymbol{x}) = 1$ を保証する正規化係数である．素性関数 f_k は識別器を学習させる際に，識別の情報として与える特徴量であり，式 (2) のように定義される．

$$f_k(\boldsymbol{x}, y_t, y_{t-1}) = \begin{cases} \phi & \text{if condition} = \text{true} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

ここで，ϕ は素性値と呼ばれる実数値であり，素性に応じて任意に設定できる．素性関数は入力 $\boldsymbol{x}$ および t 番目と $(t-1)$ 番目の出力ラベルに依存しており，問題に応じて設計する．

学習では入力データ $\boldsymbol{x}$ と出力ラベル $\boldsymbol{y}$ を 1 組とする学習データ $\boldsymbol{D} = \{(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)})\}_{i=1}^{N}$ を与え，最尤推定により重み λ_k を求める．条件付き確率 P の最尤推定をとった目的関数 $L(\boldsymbol{D})$ を最適化する．

$$L(\boldsymbol{D}) = \sum_{\boldsymbol{D}} \log P(y_t, y_{t-1}|\boldsymbol{x}) - \sum_{k} \frac{\lambda_k^2}{2\sigma^2} \quad (3)$$

ここで，式 (7) の第 2 項は過学習を防ぐための正規化項である．目的関数 $L(\boldsymbol{D})$ を最適化するために，最急勾配法により重み λ_k を決定する．

学習した識別器を使って，入力データ $\boldsymbol{x}$ に対する出力ラベル $\boldsymbol{y}_{\mathbf{out}}$ を予測する際は，式 (4) に示す最大化問題を解くことにより決定できる．

$$\boldsymbol{y}_{\mathbf{out}} = \underset{\boldsymbol{y}}{\operatorname{argmax}}\, P(\boldsymbol{y}|\boldsymbol{x}) \quad (4)$$

3.3 系列ラベリング問題としてのモデル化

提案手法では，アルゴにおいて自動観測されたプロファイルに対して，機械学習により品質管理フラグを自動で識別する． 品質管理フラグには 1(正), 2(おそらく正), 3(おそらく誤), 4(誤) が用いられているが，フラグ 2,3 は曖昧さを含むフラグであるため，本実験では 2 を 1，3 を 4 とみなして，2 値の品質管理フラグに識別する．塩分の品質管理フラグを識別対象とし，密度逆転（正側，負側），同値エラー，その他観測不良，オフセットエラーの 5 種類の観測エラーを検出する．

学習に用いる素性は表 1 のように設計した．$PRES$ はプロファイルの圧力，$TEMP$ は水温，$PSAL$ は塩分，$PDEN$ は密度であり，添え字は t 層目における観測値を表す．また，$PDEN_{\max}$ は $(t-1)$ 層目までの最大密度値，$PDEN_{\min}$ は $(t+1)$ 層目以降の最小密度値を表す．f_1 は塩分値のとりうる許容範囲を表す素性であり，出現回数が多いことを考慮して素性値 $1/T$ とした．f_2 は塩分値の許容範囲外を表す．f_3 は圧力 800[dbar] 以深において，連続した層で観測値が同値になっていることを表す素性，f_4 は水深値が同じ値，または逆転が発生したことを表す．f_5, f_6 はそれぞれ負側密度逆転，正側密度逆転による観測エラーを表す素性で，SVM を導入した素性である（3.5 節）．

3.4 オフセットエラーの検出

オフセットエラーの検出には，World Ocean Atlas 05(WOA) のデータセットや過去に観測されたプロファイル利用する．WOA の 1 度格子データは欠損箇所が見られるため，5 度格子データを用いる．表 1 における WOA は入力プロファイルと比較可能な WOA データ値であり，観測位置および圧力値を揃える必要があるため，入力プロファイルの観測位置を囲む 4 つの WOA データから重み付き内挿を行い，観測層間で線形補間をして算出した．

表 1: 使用した素性

No.	内容	素性の条件	素性値
f_1	正常	$2 \leq PSAL_t \leq 41$	$1/T$
f_2	不良	$PSAL_t < 2, 41 < PSAL_t$	1.0
f_3	同値	$PSAL_t = PSAL_{t+1}, 800 < PRES_t$	1.0
f_4	水誤	$PRES_t <= PRES_{t-1}$	1.0
f_5	負密	$SVM(PDEN_t - PDEN_{MAX}, PRES_t) = 4$	1.0
f_6	正密	$SVM(PDEN_{MIN} - PDEN_t, PRES_t) = 4$	1.0
f_7	オフ	7 割の層が WOA と標準偏差の 3 倍以上の差がある	1.0
f_8	オフ	7 割の層が過去データと 0.05 以上の差がある	1.0
f_9	オフ	過去アルゴデータの塩分ラベルがすべてエラー	0.1

3.5 閾値の自動決定

CRF の素性設計の一選択に機械学習を加えることで，素性設計の支援を試みる．CRF が学習を行う際，識別情報として素性を与える必要があるが，本問題の素性を設計するためには，品質管理フラグ識別の経験的知識を要する．特に閾値を考慮する場合，経験的知識がない限り素性の設計が困難である．このような素性に対して，素性の条件に機械学習の導入を可能とし，閾値が不明なデータを学習させることで，自動的な閾値境界の決定を試みる．

SVM には圧力値および密度逆転幅を入力として与えるが，学習に不必要なデータを取り除くため，圧力値 0〜2200[dbar] のうち，密度逆転が発生している観測層のデータ，すなわち，密度逆転幅が 0 以上のものを入力として採用した．また，外れ値による学習の影響を考慮し，密度逆転幅が 0.03 以上のものは 0.03 とみなすことにした．以上のデータを SVM の入力として与え，密度逆転エラーの識別境界を学習させて分類器を構築する．識別時にフラグ 4 を出力すると素性が成立する．

4. 評価実験

4.1 実験設定

提案方式の有効性を検討するため，自動品質管理（RQC）[7]，SVM[6]，提案方式 1（f_5，f_6 において手動で設計した素性を使用 [8]，CRF と表記），提案手法 2（f_5，f_6 において Advanced automatic QC[9] の閾値を使用，CRF-AQC と表記），提案手法 3（f_5，f_6 において SVM により閾値を決定，CRF-TA と表記）を比較する．本実験では，塩分の観測値に対する品質管理フラグの識別を試みる．評価の指標として，観測層単位の誤検出数および見逃し数，プロファイル単位の適合率（$Precision$），再現率（$Recall$），正解率（$Accuracy$）に着目する．オフセットエラーについてはプロファイル単位で評価する．適合率・再現率・正解率は，以下の式で表される．

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$Acurracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (7)$$

ここで，観測エラーを含むプロファイルにおいて，観測エラーを検出できたプロファイル数を TP，見逃した数を FN とし，正常プロファイルにおいて，観測エラーを誤検出しなかった数を TF，誤検出した数を FP とする．なお，上記プロファイル単位での評価では，フラグ 4 を付与する層が正しくない場合も，観測エラーの存在を知らせているとみなし正解とする．

実験に使用するプロファイルは，日本で遅延品質管理が施されたもののうち，北太平洋海域で観測され，観測エラーを含んだ 1,904 プロファイルを用いる．詳細には，負側密度逆転 1,384 点，正側密度逆転 213 点，同値エラー 307 点である．観測不良は密度逆転のプロファイルに含めている．実験は 10-fold cross validation を行うこととし，評価を行う際は観測エラーを含まない正常なプロファイルを同数加えた．すなわち，学習データ 1,712 点，予測データ 382 点（観測エラーを含むプロファイル：190 または 191，正常なプロファイル：190 または 191）として，予測データを入れ替えながら 10 回の実験を行った．また，CRF-TA については上記の 1,904 プロファイルに加えて，オフセット 660 点を含めた 2,564 プロファイルで実験を行い，オフセット検出を試みた．

比較対象手法として用いる SVM は，入力として注目する層の圧力，水温圧力，水温，塩分，密度と，上層の圧力，密度との差分，下層と圧力，水温，塩分，密度との差分の 10 項目を与え，カーネル関数には RBF カーネルを用いた．学習に偏りが生じてしまうことを考慮し，品質管理フラグ 4 のサンプルに対して重みを付加した．上記の重みは 10 から 30 まで 5 間隔ずつ（5 通り）変更を行いながら実験を行った．

表 2: 1 試行あたりの平均誤識別数 (単位:層数)

手法	誤検出数	見逃し数
RQC	68.5	130.6
SVM(10)	250.6	110.5
SVM(15)	334.3	106.8
SVM(20)	450.8	105.0
SVM(25)	567.0	101.8
SVM(30)	671.6	100.2
CRF	24.3	35.6
CRF-AQC	8.0	111.3
CRF-TA	**31.0**	**26.0**

表 3: プロファイル単位の適合率・再現率

手法	TN	FP	TP	FN	Precision	Recall	Accuracy
RQC	**189.5**	**1.0**	109.2	81.2	**0.991**	0.570	0.784
SVM(10)	93.6	96.8	30.5	159.9	0.623	0.840	0.666
SVM(15)	77.8	112.6	23.4	167.0	0.598	0.877	0.643
SVM(20)	58.9	131.5	17.9	172.5	0.568	0.906	0.608
SVM(25)	43.6	146.8	13.6	176.8	0.546	0.929	0.579
SVM(30)	30.6	159.8	10.8	179.6	0.529	0.943	0.552
CRF	182.6	7.8	**184.0**	**6.4**	0.959	0.966	0.963
CRF-AQC	187.2	3.0	116.6	73.5	0.975	0.613	0.799
CRF-TA	185.9	4.5	182.1	8.3	0.957	**0.976**	**0.966**

4.2 実験結果

表 2 に観測層単位での評価結果を，表 3 にプロファイル単位での評価結果をそれぞれ示す．表 2 について，誤検出数，エラーの見逃し数はともに CRF に基づく手法が少なく，品質管理ラベルの識別精度が高いことが伺える．1 試行あたりに存在するプロファイルのエラー数は 618.0 個であり，RQC はおよそ 1/5 で見逃しが発生しているが，CRF-TA は約 1/24 に抑えることができている．誤検出に関しても，CRF に基づく手法が少なく抑えることができているものの，表 3 から正常なプロファイルに対する誤検出が多くみられるため，適合率は RQC を下回る結果となった．

SVM を分類器として用いた場合は，フラグ 4 であるサンプルに対する重みが高くなるにつれて，再現率は上昇しているが，適合率は下がっている．重みを増やしても，誤検出が多くなるに対してエラーの見逃し数はあまり改善が見られず，ラベリングの品質は良好であるとは言えない．カーネル関数のパラメータを調整することで改善の余地はあると考えられる．

より詳細な結果として，RQC，CRF および CRF-TA のエラー別の誤識別の層数を表 4 に示す．自動品質管理は，正側密度逆転や同値エラーに対する誤識別が多いことが確認できる．これに対し，提案手法は正常データに対する誤検出数は多いものの，全体的に誤識別が少ないことがわかる．

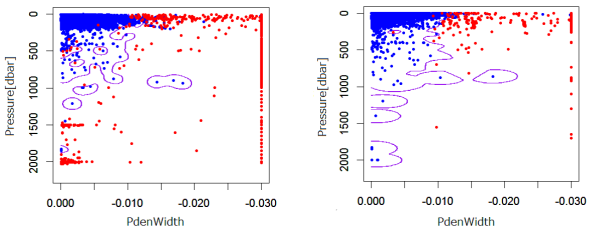
オフセットの検出は，CRF-TA を用いた場合のみ試みた．表 5 に結果を示す．オフセットはすべて検出することができていたが，一方で誤検出が発生した．

表 4: 1 試行あたりの平均誤識別回数（単位:層数）

エラーの種類	総エラー数	RQC		CRF		CRF-TA	
		誤検出	見逃し	誤検出	見逃し	誤検出	見逃し
正常	–	1.1	–	9.2	–	13.7	–
負側密度逆転	413.8	39.2	73.8	9.9	16.9	9.1	15.9
正側密度逆転	55.0	18.8	21.8	3.0	3.3	5.2	3.0
同値エラー	149.2	9.4	35.0	2.2	15.4	3.0	7.1

表 5: 1 試行あたりの誤識別回数 (単位:プロファイル数)

エラーの種類	総エラー数	CRF-TA	
		誤検出	見逃し
オフセットエラー	66.0	11.6	0.0



(a) 負側密度逆転　　(b) 正側密度逆転

図 3: SVM 素性の学習結果

また，機械学習を導入して設計した素性である，負側密度逆転と正側密度逆転の学習結果の一例を図 3(a), および 3(b) に示す．縦軸は深度 [dbar]，横軸は密度逆転幅を表す．分類結果から，圧力値 1,000[dbar] までは，0.01 程度の密度逆転は許容されていることがわかった．誤検出や見逃しも比較的少なく，上手く識別境界を自動で決定できている．表 3 および表 2 の結果からも CRF と CRF-TA の精度に差はなく，人間が設定した閾値と同程度の精度を実現できたことがわかる．

4.3 考察

評価実験により，自動品質管理，SVM，CRF を比較し，提案手法である CRF による品質管理フラグの識別が再現率の面で良好であることを確認できた．特にエラー別の識別結果では，自動品質管理が苦手とする正側密度逆転についても提案手法は誤識別が少ない結果となっていた．系列データの前後の関係を素性として記述することで，負側と正側のどちらの密度逆転エラーであるかを上手く判断できていると考えられる．

しかし，本実験では識別対象の観測エラーを絞っており，複数層にわたる密度逆転などの識別に誤りが生じていた．これは，複数層にわたる密度逆転が生じている学習データが少ない，いわゆる不均衡な学習データであることが理由として考えられ，単層の密度逆転の学習に打ち消されていると推測される．

素性の設計に機械学習 (SVM) を取り入れることにより，閾値の設定を要することなく，識別境界を決定できることが確認でき，今後，素性設計の支援に貢献できるものと考える．しかし，例外サンプルなどの影響により，ノイズのような識別境界が発生しており，誤検出を誘発している可能性がある．

5. おわりに

本研究では，全球海洋監視システム「アルゴ」において，自動観測されたアルゴデータに対し，機械学習により品質管理フラグを自動識別する方式を提案した．アルゴデータの特徴から本問題を系列ラベリング問題として捕捉し，CRF を適用することにより識別を試みた．評価実験により，観測エラーの誤検出・見落とし数を削減し，再現率の改善が見られ，提案手法の有効性を確認することができた．また，CRF における素性の設計に機械学習を導入することで，素性に用いる閾値設定を行わずに自動で決定することを可能とした．これにより，経験的知識を必要とせず，入力データと属性の選択のみで素性の定義が可能となり，素性設計を支援できる可能性を示した．

本研究では識別対象を限定しており，その他の観測エラーや水温エラーにも対応させる必要がある．識別対象としていた塩分の観測エラーの中でも，複数層にわたる密度逆転は識別できず，課題が見られた．これらのエラーに対応するために，さらなる素性の拡充やモデル選択，不均衡学習データの問題を払拭する対策が必要である．

謝辞

本研究を進めるにあたり，独立行政法人海洋研究開発機構・地球環境変動領域・アルゴデータ班に協力頂いた．また，本研究の一部は，倉田記念日立科学技術財団 倉田奨励金によるものである．ここに記して感謝の意を表する．

参考文献

[1] Argo Data Management Team, Report of the Argo Data Management Meeting. Proc. Argo Data Management Third Meeting, Marine Environmental Data, 2002.

[2] 細田 滋毅, 全球海洋監視システム「アルゴ」, 人工知能学会第 27 回全国大会, 3K1-OS-08a-1, 2013.

[3] Argo sicence team, Argo: The global array of profiling floats, in Observing the Oceans in the 21st Century, edited by C. J. Koblinsky and N. R. Smith, pp. 248–258, GODAE Project Office, Bureau of Meteorology, 2001.

[4] Lafferty, John, Andrew McCallum, and Fernando CN Pereira. ”Conditional random fields: Probabilistic models for segmenting and labeling sequence data.” (2001).

[5] Seymore, Kristie, Andrew McCallum, and Roni Rosenfeld. ”Learning hidden Markov model structure for information extraction.” AAAI-99 Workshop on Machine Learning for Information Extraction. 1999.

[6] C. Cortes and V. Vapnik “Support-vector networks”, Machine Learning, vol.20, no.3, pp.273-297, Sep. 1995.

[7] Wong, Annie, Robert Keeley, and Thierry Carval. ”Argo quality control manual.” 2014.

[8] Ono, Satoshi, et al. ”A Preliminary Study on Quality Control of Oceanic Observation Data by Machine Learning Methods.” Proceedings of the 18th Asia Pacific Symposium on Intelligent and Evolutionary Systems, Volume 1. Springer International Publishing, 2015.

[9] JAMSTEC. ”データセット「Advanced automatic QC (AQC) Argo Data」について” 2014.