

奈良女子大学大学院修士論文

地震短期予測のための
統合的データベースシステムの構築

奈良女子大学大学院人間文化研究科情報科学専攻
(学籍番号：I06-003)

石川 千里

指導教官： 城 和貴

平成19年1月

概要

地震の被害を最小限に食い止めるためには数週間から数日単位の短期予測が必要であり，地震短期予測システムの早期実現が望まれている．現在，様々な電磁気学的な地震短期予測が研究されている．高精度な予測を行うためには，短期予測の結果を組み合わせる方法が有効であるが，効果的な組み合わせは解明されていない．この組み合わせ方を解明する手法として，データマイニングが考えられる．データマイニングを行うためには，ある観測値データに対してある解析を適用した結果からシグナルがどのような形式で検出されたかという情報を格納するデータベースが必要である．また，シグナルを効率よくデータベースに記録していくためには，観測値データを統合的に扱い，様々な解析が行えるようにし，解析結果から検出されるシグナルを格納するシステムが必要である．本研究では地震短期予測のための統合的データベースシステムの構築に着手する．本稿では，データベースシステム構築の一環として，解析プログラムにICA(Independent Component Analysis)を適用する．

電磁気学的な地震短期予測の方法の1つに，VAN法という，地面に埋めた電極で地電流の電圧を測定し地電流に稀に含まれる矩形波の電流を検出することで地震の予知を図る手法がある．VAN法を用いる場合，“強い電車ノイズが地電流へ混入するため目視での矩形波の検出が困難になる”という問題が生じる．この問題を解決するために，ICAを適用することで電車ノイズを分離させ矩形波を検出する試みがされている．すでに，ICAが地電流に対して有効であることが示されている．

実際の地電流データへのICA適用実験を行うために，新潟県中越地震震源地点から約120km離れた長野県・松代観測点で観測されたデータを用いる．また，ICA適用結果から矩形波を検出するための専用ツールを開発する．専用ツールを使用してICA適用実験の結果から矩形波を検出した結果，多くの矩形波が検出された．中でも，中越地震発生翌日の地電流データの分離結果からは，中越地震の後に発生した多数の余震の前駆的シグナルと推定されるものが検出された．この結果より，矩形波検出の手段としてICAを用いることの有効性が高いことがわかる．ゆえに，データベースシステムの適用手法の1つとして実装を行う．

データマイニングに使用できるようなデータベースシステムにするためには，システム内に，観測値データを保存する観測値データベース，観測値データベースの任意のデータに対して解析プログラムを適用する機能，解析結果を適切な形式で表示させる機能，解析結果表示によって見つけられたシグナルを登録する機能，シグナルの情報を記録するデータベース，シグナル情報を適切な形で検索する機能を組み込む必要がある．そこでまず，データベースシステムに組み込むデータベースの設計を行う．観測値データベース設計に関しては，パイ

ナリ形式を用いた膨大な数値データの効率的な格納と，XMLを用いた柔軟なメタデータの設計について述べ，シグナルデータベースの設計では，数値実験において重要な実験履歴に触れ，それを取り入れたデータベースの設計について述べる．また，データマイニングに使用するためのシグナル問い合わせ機能について設計する．

データベースシステム開発の一環として，地電流データに対してICAを適用し適用結果を表示する機能OnlineICAを開発し，データベースシステム内に実装する．OnlineICAは，地電流データに対してICA適用などの処理をサーバに自動で行わせ，その結果をグラフで表示させることができるものとする．OnlineICAのプロトタイプをデータベースシステムとは別個に作成した後，データベースシステムへ実装する．データベースシステム内に実装されたOnlineICAでは，同研究室豊島によって開発中であるデータベースシステムインターフェースで地電流データや適用処理を選択することができる．インターフェース内の実行ボタンを押すと，OnlineICAが選択地電流データにICAを適用させ，適用結果のグラフ表示を行うという流れになっている．

キーワード：地震短期予測，データベース，ICA

目次

概要	ii
目次	iv
図目次	v
第 1 章 はじめに	1
第 2 章 地震予測研究	2
2.1 地震予測の分類	2
2.2 電磁気学的な短期予測	2
2.2.1 VAN 法と地電流データに対する工学的手法適用	3
第 3 章 ICA を用いた新潟中越地震前駆的シグナルの抽出	5
3.1 ICA とは	5
3.2 地電流データ	5
3.3 ICA の適用方法と適用例	6
3.4 ICA の出力結果 Viewer	7
3.5 矩形波の自動検出	9
3.6 矩形波の検出と考察	10
第 4 章 データベースシステム	12
4.1 データベースの設計	12
4.1.1 観測値データベース	12
4.1.2 シグナルデータベース	14
4.1.3 シグナル問い合わせ機能の設計	16
4.2 OnlineICA の構築	18
4.3 データベースシステムの実装状況	19
4.3.1 実装中のデータベースインターフェイスと観測値データベース	19
4.3.2 データベースシステムへの OnlineICA 組み込み	20
第 5 章 まとめ	22
謝辞	23
参考文献	24
研究業績	26

目 次

2.1	1999.1.17, 1:30-2:30 の松代観測所の地電流	4
2.2	1999.1.17, 一日間の松代観測所の地電流	4
3.1	松代観測所の測線配置図	6
3.2	ICA 適用結果のうち矩形波を含むもの	7
3.3	グラフ Viewer	8
3.4	ヒストグラム作成の流れ	9
3.5	矩形波自動検出結果	9
3.6	ICA 適用結果内での矩形波発生回数	10
4.1	RootXML ファイル内階層図	15
4.2	DetailXML ファイル内階層図	15
4.3	観測所 XML ファイル内階層図	15
4.4	解析ログ XML ファイル内階層図	17
4.5	SignalXML ファイル内階層図	17
4.6	解析手法 XML ファイル内階層図	17
4.7	OnlineICA 実行例	19
4.8	データベースシステム内 OnlineICA データフロー図	21
4.9	データベースシステム内 OnlineICA フローチャート	21

第1章 はじめに

地震国であるといわれている日本は、過去に幾度となく、地震によって甚大な損害を受けている。この地震による損害を軽減させるためには、地震の発生時期・規模・場所を予測することが重要である。従来の地震学において、地震予測として認められている手法は、統計的手段を用いた長期予測である。しかし、地震の被害を最小限に食い止めるためには数週間から数日単位の短期予測が必要であり、地震短期予知システムの早期実現が望まれている。

現在、地震を短期的に予測するために、様々な電磁気学的な地震短期予測が研究されている。電磁気学的短期予測では、地震が発生する前に生じる岩石圧縮・岩石破壊などによる電磁異常を地震前駆的シグナルとして検出し、地震の短期予測を図る [2]。地震前駆的シグナルが含まれると考えられるデータは ULF 帯や FM 波、地電流など多様であり、シグナルの形式は観測の種類によって異なる。また、シグナルを検出するための解析手法もそれぞれの観測データに対して多々考えられている。

高精度な予測を行うためには、短期予測の結果を組み合わせる方法が有効であるが、効果的な組み合わせは解明されていない。この組み合わせ方を解明する手法として、データマイニングが考えられる。データマイニングを行うためには、ある観測値データに対してある解析を適用した結果からシグナルがどのような形式で検出されたかという情報を格納するデータベースが必要である。また、シグナルを効率よくデータベースに記録していくためには、様々な解析が行えるように観測値データを統合的に扱い、観測値データの解析結果から検出されるシグナルを格納するシステムが必要である。

そこで、本研究では地震短期予測のための統合的データベースシステムの構築に着手する。データベースシステム内には、観測値データを保存するデータベース、観測値データベースの任意のデータに対して解析プログラムを適用する機能、解析結果を適切な形式で表示させる機能、その表示によって見つけられたシグナルを登録する機能と、シグナルの情報を記録するデータベースを組み込む。

データベースシステム構築の一環として、本稿では、解析プログラムに ICA (Independent Component Analysis) を適用したシステムを開発する。ICA を用いて工学的に地電流データを解析する試みは、すでに行われており、VAN 法 [1] による地震前駆的シグナルの抽出を行っている。

本稿の構成は以下のとおりである。

第2章では、既存の地震予測の研究について述べ、電磁気学的な短期予測の例について説明し、我々が着目している VAN 法 について述べる。第3章において、地震前駆的シグナルを含んでいると思われる地電流データへの ICA の適用方法や適用結果から矩形波を検出するためのツールの開発、そして検出結果について説明する。第4章にて、データベースシステムを設計し、更に地電流データに対して ICA を適用するための機能として OnlineICA を構築し、データベースシステムへの実装を行う。

第2章 地震予測研究

地震発生前に発生の時期・場所・規模を予測することを表す単語として，“地震予知”と，“地震予測”がある．上田 [2] は、「日本語の予知という言葉ではなにかしら『預言者の所行』といったニュアンスがあるらしい」という理由で，“予知”ではなく“予測”という言葉を用いている．本稿でも“地震予知”ではなく，“地震予測”の単語を用いる．なお，地震予測は，地震予知に較べて，一段確度の下がった推定の意味で用いられることもある．

本章では，地震予測について，予測時間および時刻精度という観点から予測手法を分類することで説明し，本稿で述べるデータベースシステムと深く関わる電磁気学的な手法を用いた短期予測について述べる．なお，本稿では予測時間および時刻精度を「予測から地震発生までの時間」，「地震は発生時期をどこまで詳細に予測できるか」という意味で定義する．

2.1 地震予測の分類

地震予測の方法は，統計的手法を用いて予測する手法や，地震の前兆と思われる現象を直接捕らえて予測する方法など様々なものが考えられている．これらの地震予測は，以下のようにより，予測時間や時刻精度によって分類することができる．

長期予測：地殻変動の経年的観測を用いて確率的・統計的な予測が行われる．予測時間は長い，時刻精度が低い．政府の地震調査研究推進本部によって発行される長期評価では，10年以内・30年以内・50年以内についてそれぞれ地震の発生確率が予測される．

短期予測：地震発生のごく初期の段階に，なんらかの前駆的な現象をとらえることで，予測を行う．前駆的現象の発生が地震発生の数時間から数週間であるため，予測時間は短い，時間精度が高い．

直前予測：大地震の初期微動（P波）を検知することで，遅れて伝わる主要動（S波）を知らせる．予測時間が主要動の数秒～数十秒前と，きわめて短い．

現在，有効な予測として広く認められているものは長期予測・直前予測である．長期予測の結果は公的に発表され，直前予測は緊急地震速報として既に実用化されている．しかし，時刻精度が数十年単位と低い長期予測や，予測時間が主要動の数十秒まえと短い直前予測だけでは，十分に地震に備えることが難しい．そのため，短期予測の早期実現が望まれている．

2.2 電磁気学的な短期予測

地震が発生する前には，岩石破壊や岩石圧縮などによって様々な電磁気学的な現象が生じていると考えられており，その現象を観測することで地震の短期予測を行うという研究が世

界各国で行われている。以下で述べる ULF 法，見通し外 FM 電波観測，VAN 法は電磁気学現象を用いた短期予測の一種である。

ULF 法では超低周波である ULF 帯の電界，磁界を測定対象とする。この周波数帯では地中における表皮効果による減衰が低いため，地震に伴って発生する電磁放射を観測しやすいという特長を持つ。南北方向，東西方向，地面に垂直な方向に反応するセンサーを地面に埋め，観測機器として使用する。

見通し外 FM 電波観測は，地震の発生に伴って瞬間的に電離柱が生じるという現象を利用している。電離柱により，通常は電離層を突き抜けてしまう VHF 帯が反射される。通常，遠方からの FM 波は受信できないが，VHF 帯に含まれる FM 波は，電離柱による反射で通常より遠方へと伝わる。見通し外 FM 電波観測では，その性質を利用して，遠方からの FM 波を受信することで，地震の発生に伴う電離柱の発生を検出を行う。観測機器として，FM を受信するためのアンテナやチューナーが用いられる。

地電流の観測による VAN 法は，既にギリシャで被害の軽減に多大な功績をあげている。そのため，日本での地震予測にも効果が期待される。そこで本研究グループでは，この VAN 法に着目する。VAN 法の詳細については 2.2.1 で述べる。

2.2.1 VAN 法と地電流データに対する工学的手法適用

地震の数週間から数時間まえには，岩石圧縮・破壊によって電流が発生していると考えられる。この電流の変化を検出し，短期的に地震を予測する方法が考えられた。地中に流れる電流は，直接測定することが困難である。そのため通常は，2 地点の地面に電極を埋め，その間の電位差を測定することで地電流の観測を行う。岩石圧縮・破壊による電流の変化は，測定された電位差の時系列データ内に矩形波として現れる。この矩形波は，地震前兆現象または地震前兆シグナルとよばれることがあるが，上田は，“前兆”という単語に神がかり的なニュアンスがあるとして，“前駆”という単語を積極的に用いている。本稿でも，この矩形波のことを地震前駆的シグナル (SES: Seismic Electric Signal) とよぶ。

SES を捉えることで行う地震予測には，ノイズ問題が伴う。地面には，雷などの自然現象や，工場の漏電・電車などの人工活動などが原因の電流が流れている。これらの電流が SES を検出する際のノイズとなり，1 つの時系列データではノイズから SES を検出することはできないのである。

VAN 法では，東西南北両方向に測線と呼ばれる電極の対をいろいろな間隔で埋め込み，電位差を測定する。各観測点には，100 メートル程度の測線と，数キロ程度の測線が設置される。1 つの観測点に多数の測線を張って同時に観測することで，ノイズと SES を区別しようとするのが VAN の特徴である。VAN 法を用いた地震予測は，1981 年にギリシャで始められた。ギリシャでは，1988 年に初めて地震予測に成功してから，1995 年までに 70% 以上の確率で予測に成功している。ここでの成功とは，誤差がマグニチュードで ± 0.5 以内，震央位置で 100km 以内，期日で数日～2ヶ月程度以内 (信号の種類で異なる) であることを指す。

日本においても，ギリシャの VAN 法と同様の観測手法によって，1999 年 1 月 17 日未明に矩形波の電位変化が観測されている。図 2.1 は，矩形波が発生した時間帯の電位変化を示すグラフである。1:35 から 2:05 にかけて，矩形状に電位が変化しているのが確認できる。しかし，日本はギリシャと異なり，電車の線路から十分に離れた地点に測線を設置することが出来ないため，日本で観測される地電流には，強力な電車のノイズが含まれてしまう。図 2.2

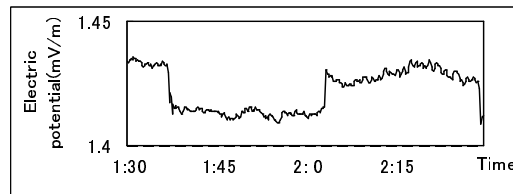


図 2.1: 1999.1.17, 1:30-2:30 の松代観測所の地電流

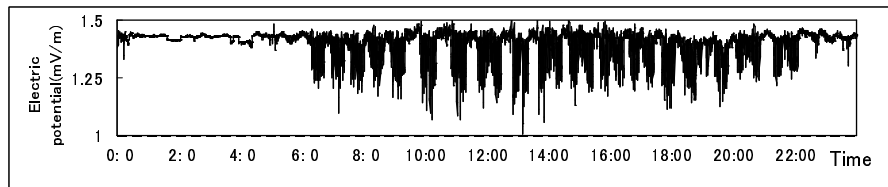


図 2.2: 1999.1.17, 一日間の松代観測所の地電流

は 1999 年 1 月 17 日の一日間の電位変化を示すグラフである．午前 6 時頃から午後 11 時頃までにかけて電車によるノイズが強く含まれている．矩形波は電車が走っていない時間に現れているため確認しやすいが，電車の走っている時間帯から SES をを見つけることは専門家でも困難であるといわれている．そのため，電車ノイズの工学的な除去が期待されている．

電車ノイズを除去するための方法として，機械学習の適用が考えられる．機械学習の 1 つであるニューラルネットワークを用いた電車ノイズの分離が既に研究されている [7]．ニューラルネットでは，特定地域の電車ノイズのパターンを学習し，地電流データから除去することに成功している．しかし，学習に膨大な計算時間を必要とするという問題があった．一方で，ICA では電車の影響を受けるとされる 2 観測点の地電流データに ICA の適用を試みた結果では，電車ノイズに関しては高い確率で正確に分離できることが確かめられており [9]，また，電車ノイズと矩形波を足し合わせて作った合成データに適用した結果，矩形波を取り出すことに成功している [8]．ICA には計算時間がニューラルネットの計算時間と比較して格段に短くてすむという利点がある．地震予測を効果的に行うためには，日々得られる観測データをすばやく解析する必要がある．ゆえに，計算時間の短い ICA の方がニューラルネットワークより，地電流データからの電車ノイズの分離に適していると考えられる．

第3章 ICAを用いた新潟中越地震前駆的シグナルの抽出

ICAで電車ノイズを分離できる可能性は高いと考えられる。しかし、実際に矩形波を含むと思われるデータに対しての適用実験は行われていない。そこで、本稿では、中越地震の地震前駆的シグナルが多数含まれていると予測されるデータに対するICAの適用実験を行う。実験に用いるデータは、新潟県中越地震震源地点から約120km離れた長野県・松代観測点で観測されたものである。このデータには中越地震の地震前駆的シグナルが多数含まれていると予測される。本章では、まず、使用するICA技術およびICAへの入力データについて述べ、ICAの適用方法について説明する。そして、矩形波を抽出する際に用いるツールを開発し、ツールを用いて矩形波の抽出を試みる。

3.1 ICAとは

ICAとは、独立した信号が、未知の割合で混合されて出来た混合成分より、もとの信号を推定する技術である。例えば、未知の場所に、独立した未知の信号 α , β を発信する信号源A, 信号源Bがあるとする。そして、信号源からの信号を観測するため、観測点X, 観測点Yを配置したとする。ここで得たいのは、信号 α の情報と、信号 β の情報である。しかし、観測点X, Yで観測されるデータ x, y は信号 α, β が線形に混合された混合信号である。ICAを観測データ x, y に適用することで、元の信号 α, β を推測することが可能となる。

現在、ICAとして様々なアルゴリズムが考えられている。ICAは音声分離問題にも使用されるが、音声分離においては、音の速度が約340m/sと、信号源と観測点の距離および観測点間の距離に対して十分に速くないため、信号源と観測点の距離に応じた波形の遅れを考慮したアルゴリズムのICAが適している。しかし、地電流データにおいては、電気の伝達速度は光の速度と同等であり、信号源と観測点の距離に対して十分に速いため、遅延の問題を考慮する必要がないと考えられる。また、対象となるデータの性質によって適した評価関数が異なるため、ICAで用いる評価関数も選ぶ必要がある。小金山(2004)は、松代の地電流データは非対称であるため、奇数次のキュムラントで表現されるデータに対して有効な評価関数が適していると述べ、評価関数として、3次のキュムラントを使用するCubICA[13]を採用している[10]。CubICAは、波の遅延を考慮していないアルゴリズムであるが、地電流データに適用する際には問題は無い。よって、本稿でもCubICAを用いる。

3.2 地電流データ

地電流は直接測定することは出来ない。そのため、長さ40cm・太さ3cm程度の鉛-塩化鉛平衡電極を地面に埋め、2地点間の電位差を測定し、地電流データとしている。この測定

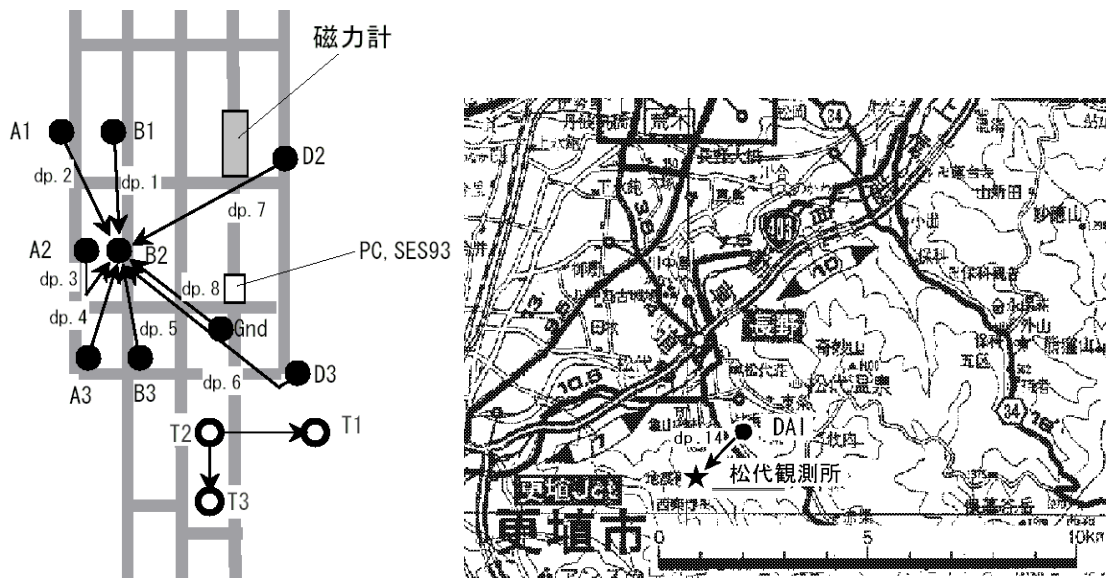


図 3.1: 松代観測所の測線配置図
 右:観測所内設置図, 左:観測所周辺図兼観測所外測線設置図

に使用される電極の対は測線と呼ばれる。

地電流の観測点は、1998年までに東海・北陸地方を中心に、42地点に設置されている。各観測点では8または16本の測線がそれぞれ異なる方向に埋設されている。これらの測線を、dp.1, dp.2, , dp.16と呼ぶ。地電流は、測線毎に決められた時間間隔で測定され、地電流データとして保存される。

次に、本稿で用いるデータとそのデータを観測した松代観測所について述べる。松代観測所には、16本の測線が埋められており、2004年は、その内の dp.1 ~ dp.14 で地電流の測定が行われた。図 3.1 は、松代の測線の配置図である。dp.1,4,5,6,7,9,10,11 では、1秒間隔で地電流の測定が行われ、それ以外の測線では10秒間隔で地電流の測定が行われている。1秒間隔で測定が行われる測線に対しては、10秒間隔での測定データも存在する。本稿での実験では、2004年1月から11月にかけて dp.2,5,6,7,14 で観測された、10秒間隔の地電流データを用いる。地電流データは測線となる対の電極間の電位差で記録されるため、測線の長さによって大きさが変わってしまう。そこで、ICAを適用する前に、観測値を測線の長さで割り、1メートル当たりの電位差という単位に変換する処理を加える。

3.3 ICA の適用方法と適用例

ICA は複数のチャンネル入力によって地電流に含まれる要素を1組もとめることができるが、入力したデータのいずれかに異常値や観測機によるノイズが含まれていると、出力結果が正常でないおそれがある。つまり、ICA の出力結果に含まれる矩形波には、ノイズの影響によって出力されたものと、シグナルの発生を示すものがあるということである。出力結果の信頼性を高めるためには、ノイズの影響による矩形波と、シグナルの発生を示す矩形波を区別できるようにしなければならない。ここで、ノイズと地震前駆的シグナルの特徴について考える。異常値は測線に依存するため、同じ異常値が複数の測線で観測されるこ

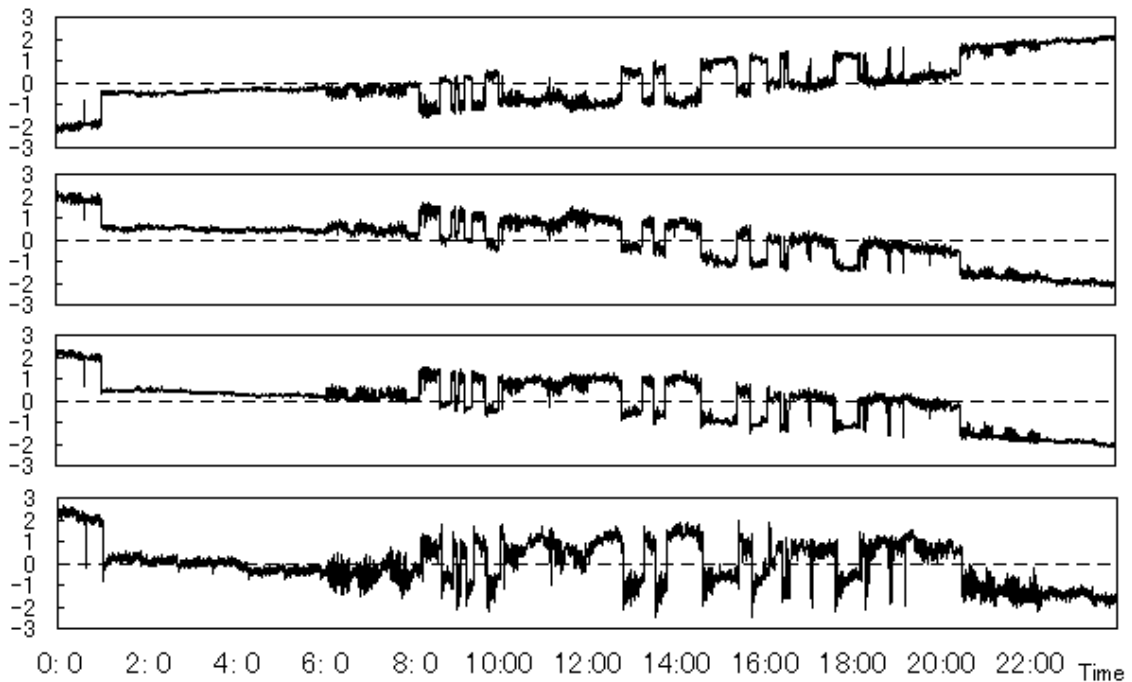


図 3.2: ICA 適用結果のうち矩形波を含むもの
 入力測線の組み合わせは上から (2,5,6),(2,5,7),(2,6,7),(5,6,7)

とが無いと考えられるのに対し、地電流に含まれる地震前駆的シグナルは複数の測線で観測される可能性が高いと考えられる。複数の測線で観測されていれば、入力測線の中にある任意の測線が存在しなくても出力されるので、そういった矩形波を探せば、それは異常値による出力ではないと考えられ、地震前駆的シグナルである可能性が高いといえる。そこで、dp.2,dp.5,dp.6,dp.7,dp.14の中から3測線選び、それを入力データセットとしてICAの適用を試みるということを行う。そうすることで、矩形波を検出する際に、ある任意の測線を含まない組み合わせで出力されている矩形波を探すことができる。使用データは2004年に10秒間隔で観測された1月から11月のデータであり、1日ごとのデータに対してICAを適用する。よって、入力データ長は8640点となる。

図 3.2 は、以上の方法で2004年10月24日のデータに対してICAの適用させ、その結果出力された独立成分の内、矩形波がみられたもののグラフである。dp.14を含んでいる組み合わせを除くと、すべての組み合わせでいずれかの成分に矩形波が確認できる。これらの矩形波は、複数の測線で観測されたものであると推定される。

3.4 ICA の出力結果 Viewer

目視での矩形波検出を効率よく行うために、ICAの結果をグラフ形式で表示するViewerを作成する(図 3.3)。ノイズの影響を考慮するには、同時系列のデータを比べながら確認することが必要であるため、このViewerでは1日分の出力結果を一覧で表示させる。ユーザは、セレクトボックスで月を選択し、任意の月の初めのデータに飛ぶことができる。Next・Backのボタンにより、次のデータや前のデータに移動することが可能である。

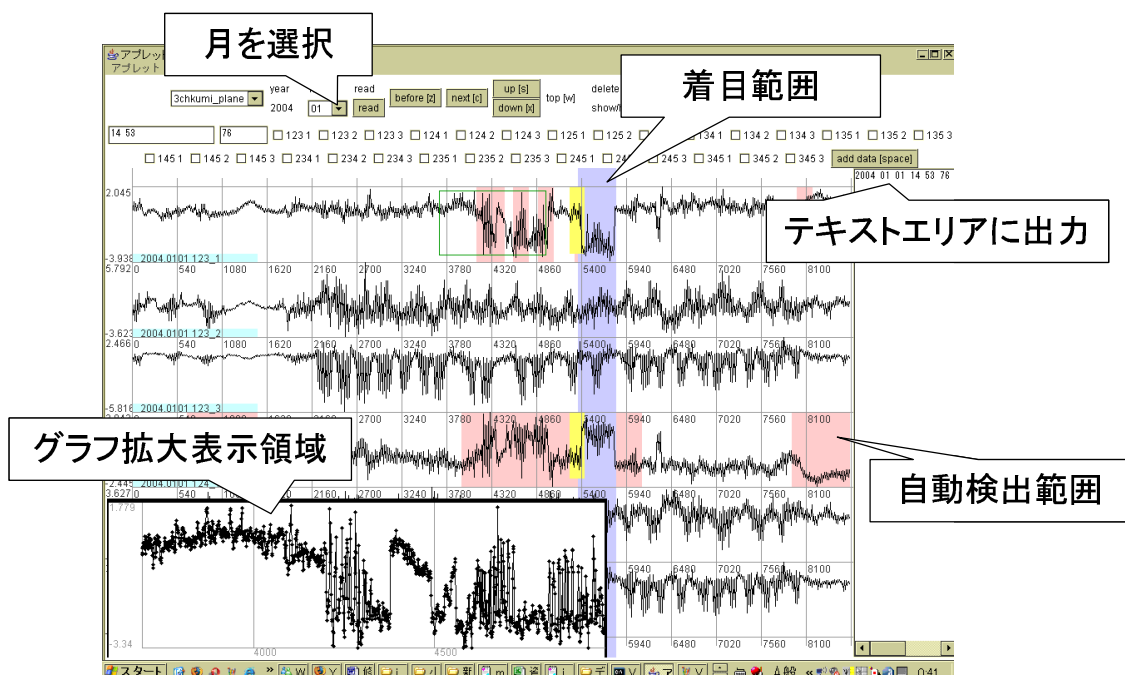


図 3.3: グラフ Viewer

本ツールは、1024*768 程度の解像度を持つ PC を想定してウィンドウを作成している。一日分の ICA 出力結果のデータ長は 8640 点であるので、一日のグラフを一画面で表示させるために、10 点から 1 点選択するというサンプリングを行い、864 点のサンプリングデータをグラフ化することで、グラフ全体表示を行う。このとき、サンプリングは等間隔で行うのではなく、前の点からの変位が最も大きいデータをサンプルとして取る。しかし、矩形波を抽出する際には、1 点単位でサンプリング前のデータの電位変化を検証する必要があるため、10 点を 1 点で表すグラフ全体表示だけでは不十分である。そこで、任意の範囲がドラッグされると、その部分を拡大表示するという機能を持たせる。拡大表示では、データを表すマーカーの表示非表示を切り替えることができる。このマーカーは拡大範囲に存在するデータ全てに対して表示されるため、発見した矩形波が電位の変位を何点で完了しているかを確認することができる。

1 つのシグナルは、複数のグラフに同時に出現する可能性がある。もし、各グラフを個別にチェックすると、1 つのシグナルを複数のシグナルとして記録してしまう可能性がある。これを回避するために、補助機能を実装する。次に、このシグナル記録補助について、使用例を用いて説明する。まず、ユーザは、1 日長のグラフと拡大機能を用いて、矩形波を探し、見つけたシグナルの開始点から終了点までドラッグする。その結果、ドラッグ開始点のグラフに対応するチェックボックスにチェックが入る。それと同時に、すべての表示グラフ内で、ドラッグした時間範囲が青い背景へと変化する。この背景が青くなった範囲を着目時間範囲と定義する。ユーザは着目範囲内で、別のグラフに矩形波が出ていないかを探し、見つけた場合はその部分をクリックする。最後に、add-data ボタンをクリックすると、シグナルが発生した年月日時分およびシグナルの波長、チェックされたグラフのラベルが、テキスト入力フィールドに出力される。

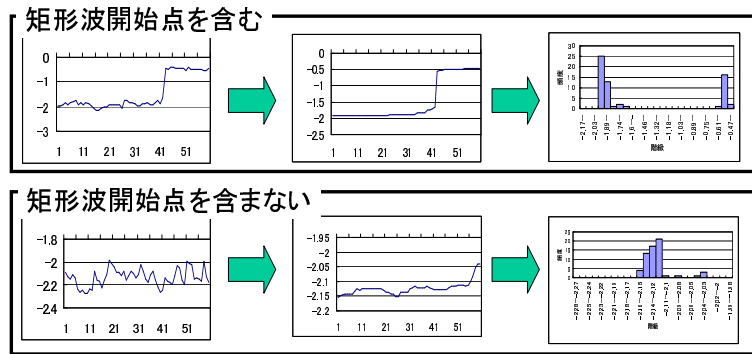


図 3.4: ヒストグラム作成の流れ

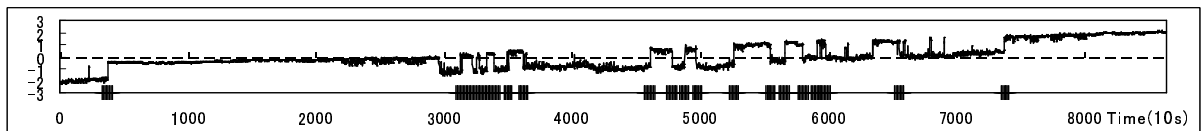


図 3.5: 矩形波自動検出結果

3.5 矩形波の自動検出

地震の短期予測をすばやく行うためには、3.3 節で述べた電車ノイズの分離だけでなく、矩形波を自動的に検出することも必要である。そこで、矩形波自動検出について考案する。

本稿の自動検出では、60(*10s) 内の範囲を着目範囲とし、範囲内で急激に平均電位が変化しているかどうかを判別することで、矩形波発生を検出を行う。判別は次のように行う。まず、着目範囲内において、データの値を最小値から最大値の間で 25 の階級に分け、値の出現分布をあらわすヒストグラムを作成する。ノイズの影響を少なくするため、メディアンフィルタを用いて着目範囲の平滑化させたものを用いて各階級の頻度を数え、ヒストグラムとする。そして次に、ヒストグラムの分布が多峰型であるかどうかを判定する。もし、平均電位が急激に変化していれば、値の出現分布は大きく分裂すると考えられるので、変化が電位の分布が多峰型であるときに、その範囲で矩形波が発生していると判別する。図 3.4 は、矩形波開始点が含まれる範囲と含まれない範囲のヒストグラム作成の流れをそれぞれ示している。

また、ここでは、下記の a,b の条件を両方満たすものを多峰型であるとする。

- a) ヒストグラムの各階級で、その階級が周囲 10 階級の中で頻度が最大値を取るかどうか判定し、その結果最大値を取る階級が 2 つ以上存在する。
 - b) そのうち頻度が 10 よりも大きい階級が 2 つ以上存在する。
- a の条件により、値の出現分布が分裂しているかどうか判断する。しかし、a のみで判断を行うと、頻度の小さい階級も検出されてしまう。そのため、大きく分裂しているかどうかを b の条件を用いて判断する。

ICA 適用結果に自動検出プログラムを適用する。自動検出プログラムは 60(*10s) の範囲に適用するものであるため、ICA の適用結果の各ファイル内で、 $i*10$ から $i*10+59$ ($i=0,1,,858$)

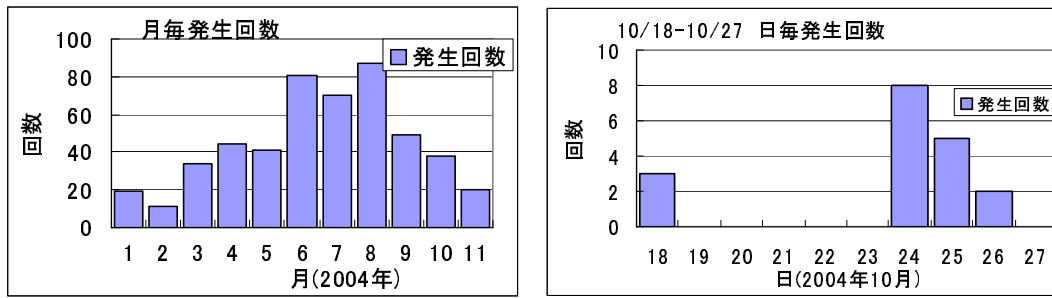


図 3.6: ICA 適用結果内での矩形波発生回数
(右: 1-11 月の月毎の合計回数, 左: 10/18-10/27 の日毎の合計回数)

の範囲に対して自動検出の適用を行う。自動検出結果の例を図 3.5 に示す。横軸が太くなっている部分が、矩形波発生範囲として検出された時間帯である。このシグナル自動検出で検出された範囲は、前節で提案した Viewer に赤く表示させる。

3.6 矩形波の検出と考察

3.4 節の Viewer を用いて、3.3 節で述べた ICA 適用の出力結果から矩形波の検出を目視で行う。この目視検出の結果は、自動検出の検証にも用いるため、ここでは Viewer の自動検出範囲表示を使用しない。検出する矩形波は、立ち上がりまたは立下りがほぼ直角、波長が 10 分以上で、複数の測線で観測されたと推定できるものとする。立ち上がり立下りがほぼ直角であるとは、電位の変位が 30 秒以内で完了することと同意とする。そして、複数の測線で観測されたと推定できる矩形波とは、ある任意の測線を除いても出力されている矩形波のことである。

矩形波検出の結果、およそ 494 もの矩形波を検出した。図 3.6 は、発生回数の統計グラフである。右が、月毎の矩形波発生回数を示すもので、左は、10/18 から 10/27 の間の日毎に発生した矩形波数を示すグラフである。月毎発生回数のグラフでは、夏の発生が多く、冬の発生が少ないという分布になっている。これは、季候依存する矩形形状のノイズの存在を示唆している。このノイズを取り除くには、複数の測線で観測された矩形波を探すというノイズ区別法だけでなく、他のシグナル判断基準も取り入れる必要があると考える。図 3.6 の右側に示す日毎の発生回数のグラフからは、地震が発生した翌日から 3 日間の間に、特に多く矩形波が発生していることがわかる。我々は、その 3 日間で発生した矩形波は、中越地震本震の後多数発生した余震の前駆的シグナルであると推定している。このように、ICA を用いることで、元の地電流データを目視するだけでは検出できなかった矩形波を多数検出できた。

以上の結果から、矩形波を選別する方法を更に考える必要はあるが、ICA が地電流から矩形波を抽出するための手法として有効である可能性はより高くなったと考える。ICA の効果をより確かめるためには、本節で用いた地電流データ以外のものに対して、ICA の適用実験を行っていく必要がある。また、それと同時に、ノイズ除去の適用実験も行う必要がある。第 4 章で述べるデータベースシステムを使用すれば、効率よく適用実験が行えるであろう。

目視検出を行った ICA 結果と同じものに対して、3.5 節の矩形波自動検出を適用し、自動検出された時間帯と目視検出の結果とを比較する。自動検出では 60(*10s) の範囲で矩形波

が発生しているかどうか判別しており，目視での検出では矩形波の開始時間および矩形波の長さを記録している．本実験では，自動検出されたある範囲内で目視によって矩形波の開始または終了が検出されていれば，その範囲の自動検出が正しいとする．そのために，正しい自動検出の時間の総和を全ての自動検出の総和で割ったものを精度とする．また，目視で矩形波の開始が検出されているある点が，自動で検出された範囲に含まれていれば，その目視検出された矩形波開始点は再現されているものとする．再現された矩形波開始点の総数を全ての矩形波開始点の総数で割ったものを再現率とする．以上の比較方法に関する結果は，再現率は84%，精度は10%となった．また，自動検出にかけた全ての時間のうち，自動検出された時間は3%であった．矩形波検出を自動検出のみで行うためには，矩形波であるかどうかの判断条件を増やし，精度を上げる必要があるが，現段階でも再現率は高いため，目視検出の補助として使うことが出来る．

第4章 データベースシステム

第3章で述べた地電流データに対するICA適用結果からの矩形波抽出の他にも、様々な電磁気学的な短期予測が存在する。2.2節で挙げているULF法と見通し外FM電波観測もそのうちの例である。精度の高い地震予測を行うためには、これらの短期予測の結果を組み合わせることが有効である。結果の効果的な組み合わせ方を見つけるには、ある観測値データにある解析を適用した結果からシグナルがどのような形式で検出されたかという情報をデータベースに格納し、そのデータベースを用いてデータマイニングを行えばよいと考える。それを行うためには、様々な解析が行えるように観測値データを統合的に扱い、その結果を格納するためのデータベースシステムが必要である。

そのためのデータベースシステム内には、観測値データを保存するデータベース、観測値データベースの任意のデータに対して解析プログラムを適用する機能、解析結果を適切な形式で表示させ、その表示によって見つけられたシグナルを登録する機能と、シグナルの情報を記録するデータベースを組み込む必要がある。

本章では、4.1節において、データベースシステム内に組み込む2種類のデータベースについて設計を行い、4.2節で、地電流データに対してICAを適用する機能を実現するためのモジュールとして、OnlineICAの実装を行う。そして4.3節で、4.1節で述べた設計に基づき同研究室の豊島が開発中であるデータベースとデータベースシステムインターフェイスについて説明し、それらの機能を組み込んだデータベースシステムへのOnlineICA機能の実装について述べる。

4.1 データベースの設計

現在扱っているデータは地電流のみであるが、将来的にはULF波やFM波などのデータを扱っていく。そのため、様々なデータの属性情報に対応できるようなデータベースにする必要がある。本節では、地電流やULF波などの観測値を保存するための観測値データベースや、それらの観測値から検出される地震前駆的シグナルを記録するためのシグナルデータベースについて設計を行う。

4.1.1 観測値データベース

システムで扱う観測値データは膨大な量となる。膨大な観測値データを効率良く扱うためには、観測値やその属性情報を記録するデータベースが必要である。そこで、まず、観測値データベースの設計を行う。データの種類によって、記憶すべき属性情報や属性間の依存が異なるため、階層構造を柔軟に記述できるXMLを用いて属性情報を記録する。一方で、地電流データなどの観測値データは、時系列の数値データであり、そのデータ量は膨大である。ゆえに、観測値データはバイナリファイル形式で保存する。

以上の考えから、観測値データベースは、データの属性情報を保存する XML ファイル、属性情報を検索する XML 検索プログラム、観測値データを保存するバイナリファイル、バイナリファイルから任意の地電流データを取り出すプログラムから構成されるものとする。本稿では、データの属性情報を保存する XML ファイルを観測情報保存ファイル、観測値データを保存するバイナリファイルを観測値データファイルと呼ぶ。

4.1.1.1 観測値データファイル

観測値データファイルの設計については、地電流データをどのように分け、どのような順序で記録するかを考えなければならない。そのためには、主にどのようなデータアクセスがあるのかを知る必要がある。

ICA を適用する際、使用する測線ごとに、その測線で観測された地電流の値が複数記録されているファイルを用意し、入力データとしている。各ファイル内の値は観測された順に並んでいる。そのため、最も効率のよいデータアクセスを実現するには、観測値を保存するバイナリファイルを測線ごとに作成し、それぞれのファイルで各測線での観測順にデータを保持させるようにすればよい。バイナリファイルには、観測されたデータのみを保存する。任意の日時の観測値データ取得は、XML ファイルに保存する記録開始日時と観測間隔の情報からファイルポインタを計算すれば可能である。それぞれの観測機の不具合などにより、データが存在しない日時が存在するため、欠乏値として NaN を用いる。

4.1.1.2 観測値属性保存ファイル

それぞれの観測値データの属性は、観測値データの種類によって決まる。それらをすべて 1 つのファイルに格納すると、ファイル内の階層構造が複雑になり、ファイルの管理やファイルに対する操作も複雑になる恐れがある。この問題を回避するために、データタイプと観測所の組み合わせごとに XML ファイルを作成し、それを管理する XML ファイルを 1 つ作成することにする。また、データの重複を避けるため、観測所の場所に関するデータを保存する XML ファイルを 1 つ作成し、その中に位置情報をまとめる。

次に、作成する XML ファイルについてそれぞれ階層図を用いて説明する。ここで用いる階層図では、長方形は XML のタグを、タグの横の五角形はその要素に付加する属性を、楕円は真上のタグ要素内に入れる値の意味を表す。アスタリスク付のタグは、同じタグが複数存在する可能性があることを示している。

- RootXML ファイル

図 4.1 は、観測値データベースの Root XML ファイルの階層図である。

このファイルでは、データタイプ 1 つにたいして、category という要素を 1 つ作成する。属性としてデータタイプ番号を登録する。category に登録される子要素として、次のものが挙げられる。データタイプ名、データタイプの単位、適応可能処理、そして、観測所 ID と Detail XML ファイルへのパス情報をもつ place という要素である。

- Detail XML ファイル

Detail XML ファイルはデータタイプと観測所の組み合わせ全てに対して 1 つずつ作成する。例として、データタイプが地電流である場合の Detail XML ファイルの階層図を図 4.2 に示す。

ファイルに 1 つ親要素 “detail” を作成し、それにデータタイプ ID と観測所 ID を属性として付加し、“dp_num”、“interval” と、複数の “dp” 要素を子要素として登録する。

“dp_num”要素は、測線の数を値としてもち、“interval”要素は、観測間隔を値として持つ。そして、“dp”要素は、測線の情報を表す次の要素を子要素としてもつ：記録開始日、測線の長さや方向、観測値データファイルへのパス。“dp”の属性として、測線IDを付加する。

- 観測所 XML ファイル

複数のデータを観測している観測地が存在するため、Detail XML ファイル内に観測所の位置情報を記録すると、同じ情報が重複して記録されることになる。データの重複を避けるため、位置データは観測所 XML ファイルに保存する。図 4.3 は、観測所 XML ファイルの階層図である。

観測所 1 箇所に対して“place”という要素を 1 つ作成し、その“place”には属性として観測所 ID を付加し、子要素として観測所名・緯度・経度を登録する。

4.1.2 シグナルデータベース

観測値データに解析処理を加えた結果から検出したシグナルを登録し、登録したシグナルの情報を用いてデータマイニングを行い、なんらかの相関関係を発見することがデータベースシステム構築の主な目的である。その目的を果たすためには、どの時刻にどのようなシグナルが発生したかといった情報を記録するシグナルのデータベースをデータベースシステムに組み込む必要がある。本節では、シグナルデータベースの設計について述べる。

相関関係を発見するためには、解析によって検出されたシグナルの有無が重要である。この 2 種類の情報を効率よく保存するためには、これらの情報を「シグナルデータ」と「解析ログ」に分ける必要がある。そして、シグナルデータとしては見つかったシグナルの情報だけを記録する。

シグナルデータに属性として解析ログ ID をつけることで、シグナルデータと解析ログの関連付けをおこなう。関連付けを用いることによって、シグナルの有無や、シグナルの検出方法が確認できる。

この解析ログを記録するための XML ファイルおよび、シグナルデータを記録するための XML ファイルの設計を行う。4.1.1.2 と同様に、作成する XML ファイルについてそれぞれ階層図を用いて説明する。

- 解析ログ XML ファイル

図 4.4 に解析ログを保存する XML ファイルの階層構造を示す。このファイル内では一連操作に対して、要素“history”を作成する。

“history”には、属性として履歴 ID を付加する。この履歴 ID は、検出されたシグナルと解析ログの関連を付けるために使用する。“history”の子要素としては、行った解析の情報をもつ“analysis”要素、解析を行ったユーザーを特定するための“userID”、使用データの情報をもつ“data”要素を登録する。

“data”は、子要素として、データタイプ ID、どこで観測されたデータであるかを特定するための観測所 ID と使用観測機の番号、いつ観測されたものかを特定するための適用範囲開始点および終了点、測定間隔を登録する。

“analysis”には、その処理が一連操作の中で何番目に適用された処理であることを示す

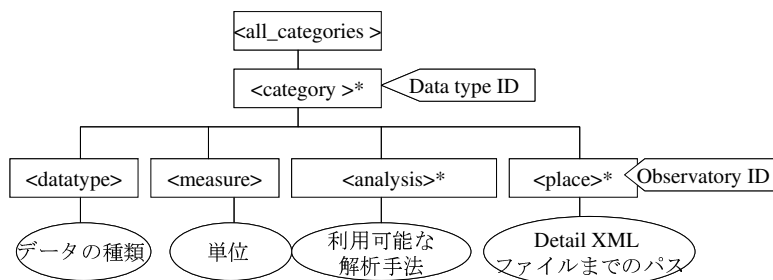


図 4.1: RootXML ファイル内階層図

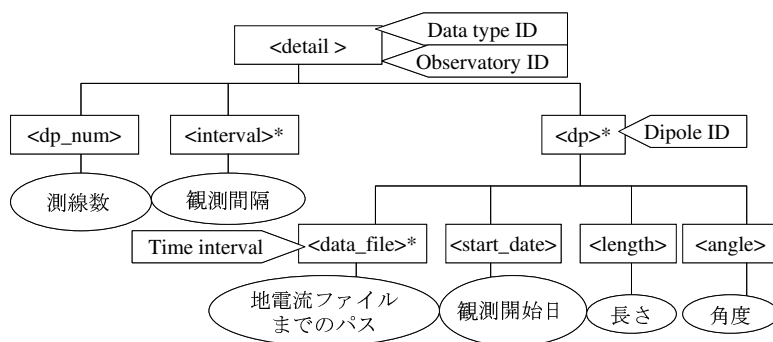


図 4.2: DetailXML ファイル内階層図

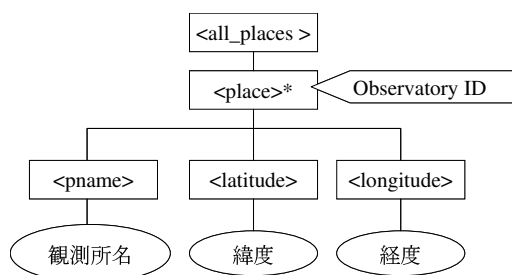


図 4.3: 観測所 XML ファイル内階層図

順番という属性を付加する．そして子要素として，適用処理 ID・必要パラメータを登録する．パラメータの例としては，入力ファイル名・出力ファイル名・閾値などの値があげられる．

- シグナルデータ XML ファイル

図 4.5 は，シグナルデータを保存する XML ファイルの階層図である．

地震前駆的シグナルの属性は，もとの観測値のデータタイプと加えた処理の組み合わせによって異なると考えられるため，データタイプと加える主処理の組み合わせごとに XML の要素型を定義する必要がある．そこで，データタイプと加える主処理の組み合わせごとに 1 つの要素を作成し，それぞれに子要素 “data” を定義する．“data” は記録するシグナルごとに作成する．例えば，地電流と ICA という組み合わせには “tcd_ica_signal” という要素を作成する．そして地電流に ICA を適用した結果から検出されるシグナルを，“tcd_ica_signal” の子要素 “data” として格納する．

ここで，“tcd_ica_signal” の子要素 “data” について階層の定義を行う．“data” には属性としてシグナルを特定するためのシグナル ID を付加し，子要素として，シグナルがどのような処理の結果から検出されたかに関するデータとして履歴 ID や，矩形波の開始時間・矩形波の長さ・矩形波の振幅を登録する．

- 解析手法 XML ファイル

解析手法 XML は処理 ID と，処理名・バージョンの対応付けを行うために作られる．図 4.6 にそのファイルの階層構造を示す．

処理 1 つに対し，“calculation” という要素を作成する．“calculation” には属性として処理 ID を付加し，子要素として処理名と処理のバージョンについて登録する．このファイルによって，解析プログラムに関する情報を一元管理することができる．

4.1.3 シグナル問い合わせ機能の設計

シグナルの属性情報は用いられた観測地データ種類と解析手法の組み合わせごとに異なる．データマイニングを行う際に，その属性の違いを意識せずに，シグナルデータ全体に対して共通の Query を出し，回答が得られると便利である．

1 つの問い合わせで，シグナルの必要な情報を取得するためには，問い合わせを各手法のシグナル保存情報に対して適切に変換することが必要となる．そこで，シグナルデータ全体に出す Query を設計し，データベースシステム内部でどのように Query を扱うかを考案する．

4.1.3.1 問い合わせ文

検索に用いる属性として，次のようなものが考えられる：

subitem 観測所の位置 (緯度・経度)，観測所名，観測値データ種類名，シグナル波長，シグナル発生時間 (開始時間および発生中の判定)，検出者 ID

扱う観測値データ種類や解析手法により，新たなシグナルの属性が出てくる可能性がある．新しい属性は，その属性の特性によって，ユーザが指定可能な属性とするか，ユーザの問い合わせを直接データベースへアクセスする問い合わせに変換する際に内部で用いるのみにするかを決め，実装を行えるようにすればよい．

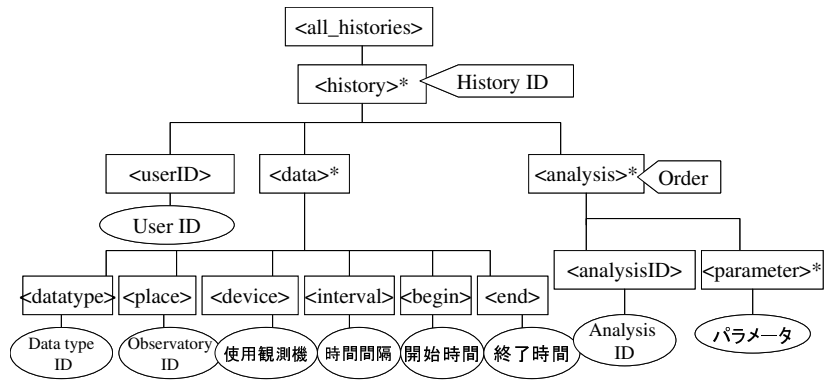


図 4.4: 解析ログ XML ファイル内階層図

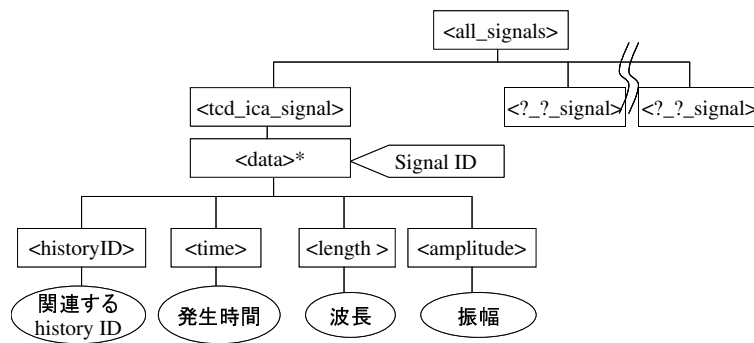


図 4.5: SignalXML ファイル内階層図

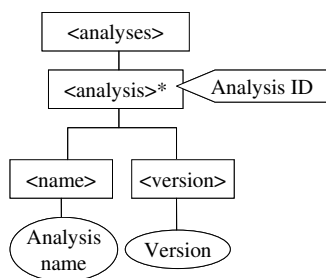


図 4.6: 解析手法 XML ファイル内階層図

属性と検索情報の比較文としては、EQUAL, NOT, GRATER, LOWER が考えられる。比較文と、和集合演算, 差集合演算, 共通集合演算, 直積演算, 選択演算, 射影演算, 結合演算, 商演算, 外結合演算, 外和集合演算などを組み合わせて、1つの Query として実行できるようにする予定である。そして、Query に対して、指定属性情報表示・該当シグナル数表示・解析履歴情報表示などのオプションをつけることで、適切な形式で回答させるようにする。

4.1.3.2 機能の実装方法に関する考案

シグナルデータベースには、検出に用いた観測値データ種類と解析手法ごとにシグナルを登録する階層を作成し、それぞれ属性を定義している。Query を解析し、検索対象の観測値データ種類に該当する階層に対して、階層の定義と対応する検索モジュールを用いて検索を行う。たとえば、地電流データに対して ICA を適用した結果として検出されたシグナル情報に対しては、tcd-ica 以下の階層を検索するモジュールを用いて検索を行う。このようにしてモジュールを分割すれば、それぞれのシグナルの情報について問い合わせ手法を独立して適宜変えることが出来る。

シグナルデータベースには、同一のシグナルと考えられるデータが異なる属性で重複して登録される場合があると考えられる。第3章で述べた、複数測線で観測されるというシグナルの特性も、重複登録の原因の一つである。シグナルの重複登録をどのように判断し、データマイニングのためにどのようにまとめるかは観測値データ種類と解析手法の組み合わせによって異なる。モジュールの分割によって、シグナルデータ重複に関する処理の違いも記述することが出来る。この処理の記述は、データマイニングを行うことと独立して行うことが出来るため、データマイニング手法の考案が効率よく行える。

4.2 OnlineICA の構築

地電流データに対して ICA を適用する機能を実現するためのモジュールとして、OnlineICA の実装を行う。本稿で構築する OnlineICA は、次のようなものを想定している。ユーザがデータを加える処理を選択し、選択情報をサーバに送信することで、データに対する処理を全てサーバに自動で行わせ、その結果をグラフで表示させる。

OnlineICA を構築するためには、Linux のローカルマシン上で使用している ICA 関連のプログラムをサーバ上で適切に動かせばよいと考えられる。

現在、ICA を適用させるために、Octave という数値演算言語を用いている。また、前処理を加えるのに C 言語、前処理から ICA 適用までを自動で実行させるために bash を採用している。ICA 適用に用いる地電流データは、全測線分が1日毎に1つのテキストファイルにまとめられている。まずはこれらのプログラムやデータを利用して、データベースシステムとは別に OnlineICA プロトタイプの構築を行う。

ユーザは Web 上に設置した HTML のフォームを使って、ICA に入力するデータの時間範囲および使用測線を指定する。そしてその入力データに加える処理を選択し、実行ボタンをクリックすると、ICA 適用前の状態と ICA 結果が共にグラフで表示される。結果表示には Java アプレットを用いている。この結果表示アプレットは、3.4 節の Viewer と同様に、グラフの拡大機能とシグナル自動検出範囲表示機能を持つ。その他に、選択範囲の開始日時と縦横幅をテキストエリアに出力する機能も持つ。この機能により、ユーザが矩形波を囲うよ

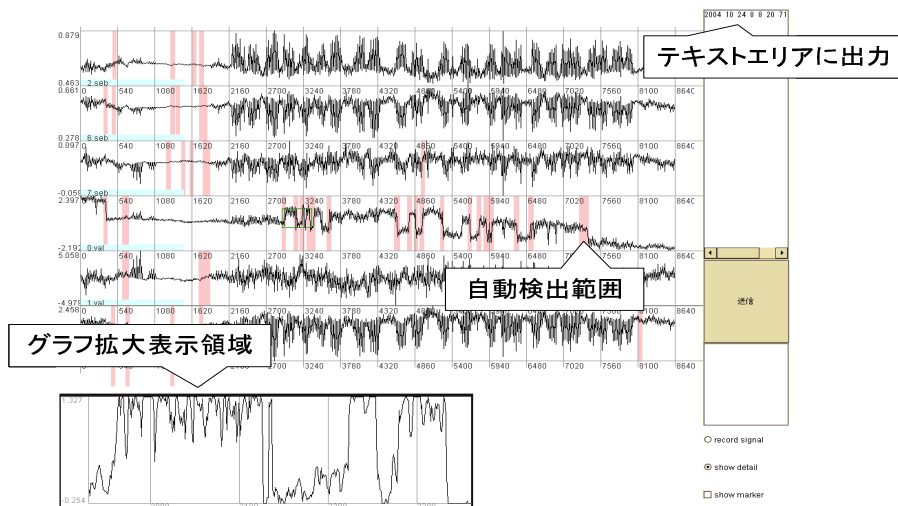


図 4.7: OnlineICA 実行例

うに範囲選択をすることで、矩形波の波長と振幅を記録することができる。

ここで構築を行った OnlineICA プロトタイプで扱うデータとして、2004年に松代で観測された地電流データを用いる。図 4.7 は、このシステムで、10月24日に測線 2,6,7 で観測されたデータを選択し、実行させた結果である。

4.3 データベースシステムの実装状況

本節では、データベースシステムの実装状況について述べる。4.3.1 では、4.1 節で述べた設計に基づき同研究室の豊島が開発中であるデータベースとデータベースシステムのインターフェイスについて説明する。そして、4.3.2 において、筆者はそれらの機能を組み込んだデータベースシステムに OnlineICA を統合させる。

4.3.1 実装中のデータベースインターフェイスと観測値データベース

同研究室の豊島は、データベースシステムの一部である観測値データベースの XML ファイルおよび XML の検索プログラムが作成している。豊島はまた、データを決定するための属性情報選択を適切に行えるようなインターフェイスの実装も行っている。そのインターフェイスでは、属性情報を指定していくことで、観測値が保存されるバイナリデータへのパスや測線の長さなど、観測値を得、観測値を用いて解析を行うために必要な情報を取得することが出来る。

インターフェイスは JSP を用いて実装されているため、アプリケーションサーバである Tomcat を利用することで動作する。現在、Windows2000,XP 上の Tomcat5.5 を用いて、ローカルサーバ上での動作確認がおこなわれている。

豊島による上記の実装にあわせ、筆者は 2004 年の松代観測所の地電流データを 4.1.1.1 で述べた形式へ変換した。また、観測値データファイルから任意の範囲を一時ファイルへコピーするモジュールも作成した。

以上のものをあわせて、データベースシステムインターフェイスと観測値データベースとする。

4.3.2 データベースシステムへの OnlineICA 組み込み

4.3.1 で述べた豊島によって開発されているデータベースシステムインターフェイスと観測値データベースと、4.2 節で述べた筆者が開発している OnlineICA を結合させる。結合したものはデータベースシステムの一部となる。この結合作業は筆者が行う。まず、現在 Windows 上で動作しているインターフェイスとの結合作業を行うために、OnlineICA のモジュールを、Linux および Windows で動作可能なものへと書き換えを行う。4.2 節の OnlineICA 内の、Windows サーバで動作不可能なモジュールは、Octave で記述された ICA プログラムと bash である。Octave は多くの数値計算ライブラリを備えており、行列計算を多用する ICA のプログラムの記述に向いている。しかし、Windows サーバ上では動作しないため、インターフェイスとの結合作業を行うことが出来ない。また、インタプリタ形式の言語であるが故に、コンパイルで最適化される言語よりも動作が遅いという欠点もある。そこで、C++ でコンパイルできるように、移植を行う。Octave で使用されている関数ライブラリは、C++ から呼び出して使用することができる。同じライブラリを用いて ICA で使用する行列計算を行うことで、Octave で書かれた ICA プログラムの結果と同じ結果を C++ の ICA プログラムから得られる。bash は Perl に書き換え、上位の Perl モジュールの一部とする。使用データはバイナリで保存されているため、OnlineICA で作成する一時ファイルがテキスト形式であると情報落ちが生じる。そこで、一時ファイルもバイナリ形式で作成し、各プログラムのデータ読込・書込み部分をバイナリデータに対応するように書き換える。4.3.1 で述べたモジュールと、書き換えを行った OnlineICA を結合させる。結合後の OnlineICA 機能動作部に関するデータフロー図、フローチャート図をそれぞれ図 4.8、図 4.9 に示す。フローチャートは、処理適用のモジュールをすべて管理する Perl プログラムのフローチャートである。フローチャート内の太線の四角は、Perl 以外の言語で作られたモジュールを表している。インターフェイスでは、実行ボタンがクリックされると、XML 検索プログラム、XML ファイルを用いて、使用データの情報や適用する処理などからなる処理ステータスファイルを生成する。それと同時に、OnlineICA 機能のフローが開始する。

OnlineICA 機能が開始すると、まず観測値切り出しプログラムによる、地電流ファイルからの地電流データの切り出しが実行される。このプログラムでは、観測値ファイルまでのパスと観測開始日と使用データ期間から、使用データに対する観測値データファイル内のファイルポイントを求め、指定範囲を使用地電流データとして一時ファイルへコピーする。

次に、選択適用処理に“測線で割る”という前処理が含まれていれば、使用地電流データをそれぞれ測線の長さで割るという処理を行い、処理適用後のデータを ICA 適用前のデータとして登録する。“測線で割る”前処理が含まれていない場合は、使用地電流データを ICA 適用前のデータとして登録する。

そして、選択適用処理に“ICA”が含まれていれば、ICA 適用前のデータとして登録されているデータに対して、ICA の適用を行い、その出力結果を ICA 適用後のデータとして登録する。含まれていない場合は、ICA 適用後の結果として、NULL を登録する。

最後に、登録されている ICA 適用前のデータおよび ICA 適用後のデータをグラフ表示する。ICA の適用が行なわれていない場合、ICA 適用後のデータとして NULL が登録されているため、表示は ICA 適用前のデータのみとなる。

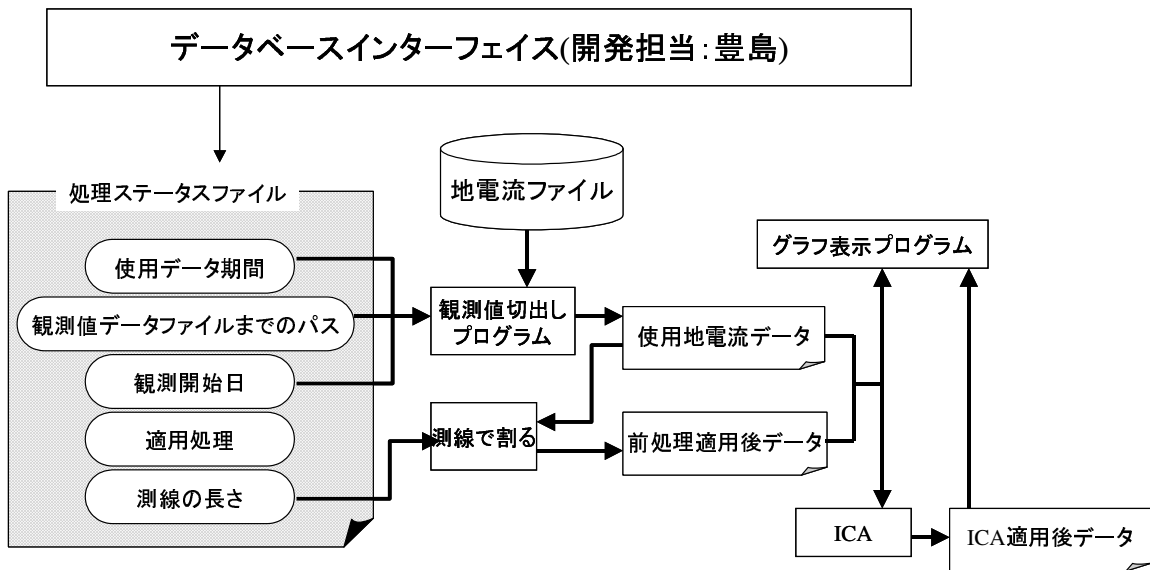


図 4.8: データベースシステム内 OnlineICA データフロー図

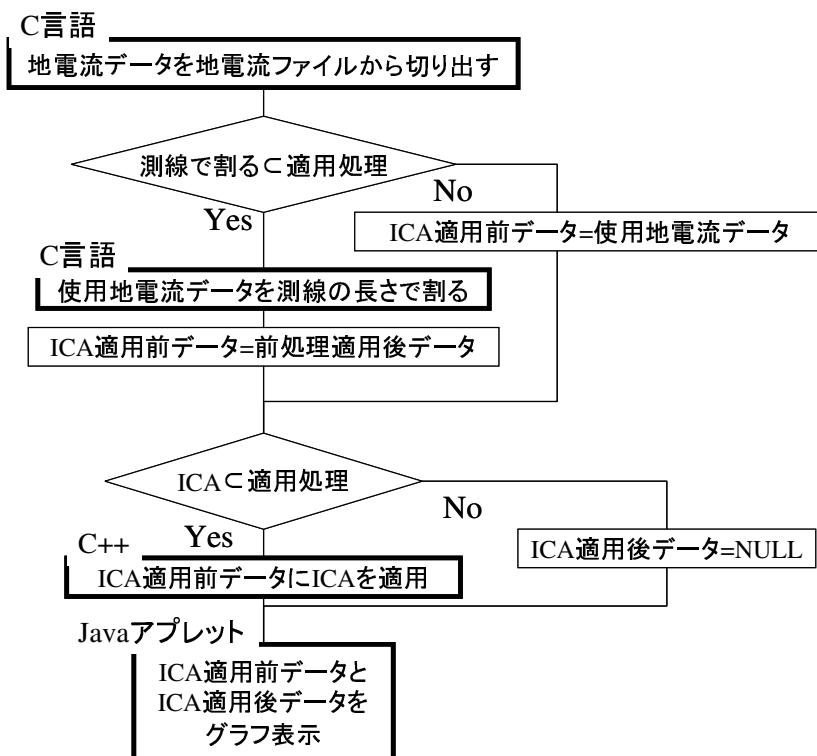


図 4.9: データベースシステム内 OnlineICA フローチャート

第5章 まとめ

本稿では、地震短期予測のための統合的データベースシステム設計および、地電流データに ICA を適用する機能として OnlineICA の実装について述べた。また ICA を新潟県松代の地電流データに適用する実験を行い、出力結果から多くの矩形波が検出されたことを述べた。

ICA の適用実験は、松代で 2004 年に観測されたデータを用いて行った。ノイズによって出力された矩形波と、シグナルに該当する矩形波を区別するために、測線の組み合わせを複数セット用いた。ある任意の測線を含まない組み合わせに対する出力結果に現れている矩形波の検出を行ったところ、494 の矩形波を検出することができた。検出された矩形波の中には、中越地震本震の後多数発生した余震の前駆的シグナルであると推定されるものもあった。これらの矩形波は、地電流データを目視で観測するだけでは検出できなかったものである。つまり、ICA が地電流から矩形波を抽出するための手法として有効である可能性がより高くなったと言える。しかし、矩形波の月毎の発生回数について統計をとった結果からは、検出された矩形波には季節依存するノイズが含まれている恐れがあることがわかった。地電流データに対する ICA の有効性をより確実なものとして示すためには、他のシグナル判断基準を取り入れて矩形波検出を行う必要があると考える。

データベースシステムの設計については、観測地データベースの設計およびシグナルデータベースの設計について述べた。観測値データベース設計に関しては、バイナリ形式を用いた膨大な数値データの効率的な格納と、XML を用いた柔軟なメタデータの設計について述べ、シグナルデータベースの設計については、数値実験において重要な実験履歴に触れ、それを取り入れたデータベースの設計について述べた。

また、データベースシステム実装の一環として、OnlineICA を開発し、それをデータベースシステムに適用することで、ICA 適用機能の実装を行った。この実装では、まずデータベースシステムから独立した OnlineICA を既存のプログラムとデータ形式を用いて実装し、観測値データベースとデータベースシステムのインターフェイスの完成にあわせて OnlineICA をデータベースシステムへ実装した。なお、このインターフェイスは豊島が開発したものである。OnlineICA 実装後のデータベースシステムでは、インターフェイスで使用する地電流データを選択し実行ボタンを押すことで、選択地電流データに ICA が適用され、その結果をグラフ形式で確認することができるようになった。

今後、データベースシステムをデータマイニングに用いることができるようにするために、観測値データベースに地電流以外のデータを格納していく。また、4.1.2 で行った設計に従い、シグナルデータベースを作成し、シグナルを登録できるようにする。そのために、OnlineICA には、結果表示で検出されたシグナルをシグナルデータベースに送信する機能を加え、地電流データに ICA を適用した結果から検出されるシグナルをシグナルデータベースに登録していく。さらに、4.1.3 で述べた、シグナルデータベースに対するシグナルの問い合わせに応じる機能を実装する予定である。

謝辞

本研究を行うにあたり，指導教官である城和貴教授には，本研究だけでなく，学生生活においても大変お世話になりました．厚くお礼申し上げます．

また，地電流データを提供して頂きさらに前駆的シグナルに関してご指導頂いた東海大学地震予知研究センターの長尾年恭教授，データベースに関してご指導頂いたお茶の水女子大学の渡辺知恵美先生，そしてICAに関してご指導頂いた日本IBM東京基礎研究所の小金山美賀氏と山口大学の庄野逸助教授に深く感謝いたします．

さらに，副指導教官の高田雅美先生には多大な助言をいただきました．ありがとうございました．

最後に，城研究室の皆様のおかげで，とても楽しい研究生生活を送ることができました．そして，大学生活を送る上で，終始経済的支援だけでなく温かく見守ってくださった両親に感謝します．ありがとうございました．

参考文献

- [1] Uyeda, S.: Introduction to the VAN method of earthquake prediction, Critical Review of Van, Sir Lighthill, J.(Ed.), pp.3-28, World Scientific, London, Singapore,(1996).
- [2] 上田誠也.: 地震予知はできる 岩波書店,(2001).
- [3] 社団法人日本地震学会: <http://wwwsoc.nii.ac.jp/ssj/index.html>
- [4] 地震調査研究推進本部: <http://www.jishin.go.jp/main/index.html>
- [5] 活断層における地震予知技術開発のための地電流等観測報告書, 気象庁 地磁気観測所, (2002).
- [6] Fukuda, K., Koganeyama, M., Shoumo, H., Nagao, T. and Joe, K.: Detecting Seismic Electric Signals by LVQ based Clustering, PDPTA2001, pp.1305-1311,(2001).
- [7] 小金山美賀, 長尾年恭, 城和貴.: ニューラルネット用いた地電流データからの電車ノイズ除去, 情報処理学会論文誌:数理モデル化と応用, Vol.42, No.SIG14, pp.124-133,(2001).
- [8] 小金山美賀, 庄野逸, 長尾年恭, 城和貴.: ICA を用いた地電流データからの電車ノイズおよび地震前兆シグナルの分離, 情報処理学会論文誌:数理モデル化と応用, Vol.43, No.SIG7, pp.92-104,(2002).
- [9] 沢小百合, 小金山美賀, 庄野逸, 長尾年恭, 城和貴.: ICA を用いた 2 観測点の地電流データに影響を及ぼす電車ノイズの抽出とその統計的評価, 情報処理学会研究報告:数理モデル化と問題解決, MPS42-24, pp.95-98,(2002).
- [10] 小金山美賀.: 地震前駆的シグナル検出を目的とした地電流データに対する工学的手法の適用に関する研究, 奈良女子大学大学院人間文化研究科複合領域科学専攻 学位論文, (2004).
- [11] 向井 良, 澤田 宏, 荒木 章子, 牧野 昭二.: 狭間隔・広間隔の複数マイクロホン対を用いた周波数領域ブラインド音源分離, 日本音響学会, 3-P-16, pp.627-628,(2004).
- [12] Sawada, H., Mukai, R., and Makino, S.: Direction of arrival estimation for multiple source signals using independent component analysis, in Proc. ISSPA2003 (International Symposium on Signal Processing and its Applications), pp. 411-414, (2003).
- [13] Blaschke, T. and Wiskott, L.: CuBICA: Independent Component Analysis by Simultaneous Third- and Fourth-Order Cumulant Diagonalization. (2003).

- [14] Blaschke, T. and Wiskott, L.: An Improved Cumulant Based Method for Independent Component Analysis, Proc.Int.Conf. on Artificial Neural Networks, (2002).
- [15] 村田昇. 入門独立成分分析 電機大出版局 (2004).
- [16] 増永良文. リレーショナルデータベース入門 サイエンス社 (2003).

研究業績

査読付

1. Ishikawa, C., Watanabe, C., Nagao, T. and Joe, K.: Extracting Seismic Electronic Signals from the Telluric Current Data for the Nigata Chuetsu Earthquake by ICA, Vol.II, PDPTA, pp.488-494, (2005).
2. Ishikawa, C., Kamo, H., Nide, N., and Joe, K.: Design of an Integrated Database System for Short-term Earthquake Prediction, Vol.II, PDPTA, pp.853-859, (2006).

査読無し

1. 石川千里, 長尾年恭, 城和貴.: ICA を用いた新潟中越地震前駆的シグナルの抽出, MPS04-53, pp. 57-61, (2005).
2. 豊島良美, 石川千里, 高田雅美, 城和貴.: 短期地震研究のための統合的なマルチデータベースの設計, MPS06-62, pp.65-68, (2006).