

# 最尤法による潮流の調和解析と最適な分潮の組み合わせの探索方法

## The Harmonic Analysis of Tidal Currents by Maximum Likelihood Method and a Way of Selecting the Dominant Tidal Constituents

田中陽二<sup>1</sup>

Yoji TANAKA

Tidal currents are one of the major attributes of flows in the sea. This paper describes the maximum likelihood method for the harmonic analysis of tidal currents. Assuming that measurements of tidal currents contain errors that follow the multivariate normal distribution, it can be explored the way to select the dominant components of tidal currents. Tidal currents were synthesized with harmonic constants by numerical simulation, and AIC and BIC were employed to select the dominant components. As a result, BIC was a more reliable criterion than AIC when using many sample data. In addition, this method was applied to surface currents in Tokyo bay.

### 1. はじめに

沿岸海洋域の流速は、潮汐現象によって引き起こされる成分(潮流)、風に起因する成分(吹送流)、密度差によって駆動される成分(密度流)などが合成されたものである。そのうち潮流成分は天体運動によって駆動され、時刻と位置のみの関数として評価され得る成分である。したがって、観測された流速から潮流成分を精度よく抽出することは、他の流速成分を理解しやすくするとともに、その海域の特性を議論する上で重要な情報を提供することができる。

現在、潮位の調和解析には最小二乗法による計算法が一般的に用いられている(気象庁, 1999)。分潮の組み合わせ(以下、分潮群)の判定として、AIC (Akaike Information Criterion)を用いた方法も池谷(1993)により提案されている。潮流の場合は、例えば東西・南北成分に分けてそれぞれを調和解析する方法が用いられている(村上, 1981)。しかしながら、最適な分潮群を選択する判断基準はこれまでなかった。これは、潮流の場合は潮位と異なり、多次元であるため、一般にお互いの流速成分が非直交になってしまうためである。また、多数の分潮群から最適な組み合わせを自動で、かつ実用的な計算コストで行える方法も確立されていない。

そこで本研究では、誤差の共分散性を考慮した最尤法による潮流の調和解析方法を示し、情報量基準を利用した最適な分潮群を自動で選択できる方法を提案する。加えて、実際の適用例として東京湾の海洋短波レーダーの観測値に本研究手法を用いて調和解析を行う。

### 2. 調和解析モデルと線形化

実際の潮流観測では鉛直方向流速は水平方向流速に比べて非常に小さいので無視され、2次元の流速ベクトルを観測対象とすることが多い。しかし、ここでは一般性

を考慮して  $p$  次元の流速ベクトルの調和解析方法を示す。 $p=1$  の場合は潮位の調和解析と一致する。

#### (1) 調和解析モデル

いま、 $N$ 個の観測された  $p$  次元の流速ベクトル  $U_n$  ( $n=1, \dots, N$ ) から  $K$ 個の分潮の調和解析を求めようとする(  $U_n$  は縦ベクトル)。流速は潮流成分とそれ以外に分離されるので、次式で表わされる。

$$U_n = u_0 + \sum_{k=1}^K f_k H_k \cos(v_k + \mu_k - \kappa_k) + \epsilon_n \quad (1)$$

ここで、 $u_0$  は平均流速ベクトル、 $f_k$  は分潮  $k$  の振幅に対する補正量、 $H_k$  は流速の振幅、 $v_k$  は時間とともに変化する項と定数項を含んだもの、 $\mu_k$  は位相に対する補正量、 $\kappa_k$  は観測地点における位相遅れ、 $\epsilon_n$  は観測値と、計算値(式(1)右辺第1項と第2項の和)の誤差である。式(1)の変数のうち、 $f_k$ 、 $v_k$ 、 $\mu_k$  に関しては分潮の種類、観測時刻、観測地点によって計算される変数である(計算方法は、気象庁(1999)を参照)。よって、未知数は平均流速  $u_0$  と調和解析定数  $H_k$ 、 $\kappa_k$  である。

#### (2) 線形重回帰モデルへの変換

調和解析を行いやすくするため、式(1)を線形重回帰モデルに変換する。

$$U_n = \sum_{m=1}^M Z_{n,m} a_m + \epsilon_n \quad (2)$$

ここで、 $M=2K+1$ 、および、

$$\begin{cases} a_1 = u_0 \\ a_{2k} = H_k \cos \kappa_k & (k=1, \dots, K) \\ a_{2k+1} = H_k \sin \kappa_k & (k=1, \dots, K) \end{cases} \quad (3)$$

$$\begin{cases} Z_{n,1} = 1 \\ Z_{n,2k} = f_k \cos(v_k + \mu_k) & (k=1, \dots, K) \\ Z_{n,2k+1} = f_k \sin(v_k + \mu_k) & (k=1, \dots, K) \end{cases} \quad (4)$$

である。調和解析定数は  $a_m$  ( $m=1, \dots, M$ ) が求まれば、次式で計算することができる。

$$H_k = \sqrt{a_{2k}^2 + a_{2k+1}^2} \quad (k=1, \dots, K) \quad (5)$$

$$\kappa_k = \tan^{-1}(\mathbf{a}_{2k+1}/\mathbf{a}_{2k}) \quad (k = 1, \dots, K) \quad (6)$$

### 3. 最尤法による潮流の調和解析と情報量基準

最尤法は尤度(対数尤度)を最大にするようにパラメータを選ぶ手法である。また、対数尤度の最大値は情報量基準を求める際に必要な値である。はじめに最尤法によるパラメータの推定方法を示し、次にモデル式の情報量基準を求める計算式を示す。

#### (1) 最尤法による調和解析

誤差成分  $\varepsilon_{in}$  ( $i=1, \dots, p; n=1, \dots, N$ ) について以下の仮定を置く。a)  $\varepsilon_{in}$  と  $\varepsilon_{i'n}$  は互いに独立である ( $n \neq n'$ ;  $i, i'$  は任意)。b)  $\varepsilon_{in}$  の期待値はゼロである。c)  $\varepsilon_{in}$  の分散は  $i$  ごとに等しく、 $\varepsilon_{in}$  と  $\varepsilon_{i'n}$  との共分散は  $i$  と  $i'$  の組み合わせごとに等しい。d)  $\varepsilon_{in}$  は多変量正規分布に従う。

上記の仮定より、 $p$ 次元流速ベクトル  $\mathbf{U}_n$  ( $n=1, \dots, N$ ) の尤度  $L$  は次式で表わされる(添え字の  $T$  は転置を表す)。

$$L = (2\pi)^{-\frac{pN}{2}} |\Sigma|^{-\frac{N}{2}} \exp\left\{-\frac{1}{2} \sum_{n=1}^N (\boldsymbol{\varepsilon}_n)^T \Sigma^{-1} \boldsymbol{\varepsilon}_n\right\} \quad (7)$$

ここで、 $\Sigma$  は分散共分散行列である。尤度  $L$  を最大にするパラメータの値は最尤推定値と呼ばれる。その場合、対数尤度  $l = \log L$  は尤度  $L$  に対して単調増加関数であるから、対数尤度を最大化するパラメータ値を求めても同じである。対数尤度は次式で表わされる。

$$l = -\frac{pN}{2} \log(2\pi) - \frac{N}{2} \log|\Sigma| - \frac{1}{2} \sum_{n=1}^N \left( \mathbf{U}_n - \sum_{m=1}^M Z_{n,m} \mathbf{a}_m \right)^T \Sigma^{-1} \left( \mathbf{U}_n - \sum_{m=1}^M Z_{n,m} \mathbf{a}_m \right) \quad (8)$$

対数尤度  $l$  を最大にするには対数尤度をパラメータ変数で偏微分した値がゼロとなればよい。

$$\frac{\partial l}{\partial \mathbf{a}_j} = \sum_{n=1}^N \Sigma^{-1} Z_{n,j} \left( \mathbf{U}_n - \sum_{m=1}^M Z_{n,m} \mathbf{a}_m \right) = \mathbf{0} \quad (j = 1, \dots, M) \quad (9)$$

上式を整理すると、 $M$ 本の連立一次方程式が  $p$ 組得られる。

$$\sum_{n=1}^N \sum_{m=1}^M Z_{n,j} Z_{n,m} \mathbf{a}_m = \sum_{n=1}^N Z_{n,j} \mathbf{U}_n \quad (i = 1, \dots, p; j = 1, \dots, M) \quad (10)$$

式(10)を解くことで  $\mathbf{a}_{ij}$  の最尤推定値  $\hat{\mathbf{a}}_{ij}$  が得られる。なお、式(10)は結果的に最小二乗法による正規方程式と一致するため、最尤推定値  $\hat{\mathbf{a}}_{ij}$  は最小二乗推定値と同じである。式(10)を解く際に分散共分散行列の情報が必要ないが、 $N \geq 2K+1$  を満たす必要がある。

同様に、対数尤度を  $\Sigma$  で偏微分した値をゼロとすることによって、分散共分散行列の最尤推定値  $\hat{\Sigma}$  が得られる(導出の詳細はAnderson(1958)を参照)。

$$\hat{\Sigma} = \frac{1}{N} \sum_{n=1}^N \left( \mathbf{U}_n - \sum_{m=1}^M Z_{n,m} \hat{\mathbf{a}}_m \right) \left( \mathbf{U}_n - \sum_{m=1}^M Z_{n,m} \hat{\mathbf{a}}_m \right)^T \quad (11)$$

#### (2) モデル式の情報量基準

現象を説明するモデルとして複数のモデルが考えられる場合、それらのモデルの比較として情報量基準が役立つ。情報量基準としてはK-L情報量を利用したAICとベイズ的アプローチによるBIC (Bayesian Information Criterion) がよく用いられる。以下に本研究モデルのAICとBICの計算式を示す。

##### a) AIC (Akaike Information Criterion)

AICは最大対数尤度  $\hat{l}$  と自由パラメータ数  $m$  より次式で定義される(Akaike, 1974)。

$$AIC \equiv -2\hat{l} + 2m \quad (12)$$

最大対数尤度  $\hat{l}$  は、式(11)を式(8)に代入することによって次式が得られる。

$$\hat{l} = -\frac{pN}{2} \log(2\pi) - \frac{N}{2} \log|\hat{\Sigma}| - \frac{pN}{2} \quad (13)$$

自由パラメータ数は  $\mathbf{a}_1, \dots, \mathbf{a}_m$  の  $p \times M$ 個に加えて、 $\Sigma$  の  $(p+p)/2$ 個である。したがって、AICは次式で計算される。

$$AIC = N \log|\hat{\Sigma}| + 2p(2K+1) + \alpha \quad (14)$$

ここで、 $\alpha$  は定数項であり、モデル比較には関係ない項である。AICは選択する分潮群によって変化し、AICを最小にするモデルが最適な分潮群と判定される。

式(14)で実質的に比較対象となるパラメータは、右辺第1項の分散共分散行列の行列式  $|\hat{\Sigma}|$ 、および第2項の分潮数  $K$  である。 $|\hat{\Sigma}|$  は誤差の広がりという意味、説明する分潮数が増えれば減少していく。しかし、第2項によって説明変数が増えることによるペナルティが加わるので、モデルの過適合(over-fitting)を防いでいる。ただし、データ数が非常に大きい場合は  $|\hat{\Sigma}|$  に対する重みが大きくなるので、第1項のみの影響で決まってしまう恐れがある(Burnhamら, 2002)。

##### b) BIC (Bayesian Information Criterion)

BICは次式で定義される(Schwarz, 1978)。

$$BIC \equiv -2\hat{l} + m \cdot \log(N) \quad (15)$$

本研究モデルのBICは次式で表わされる。

$$BIC = N \log|\hat{\Sigma}| + p(2K+1) \log(N) + \beta \quad (16)$$

ここで、 $\beta$  はモデル比較に関係ない定数項である。BICの場合は第2項に  $\log(N)$  が入っているのでデータ数の増加に伴って説明変数のペナルティが大きくなる。

なお、2次元の場合、 $|\hat{\Sigma}|$  は次式のようになる。

$$|\hat{\Sigma}| = \sigma_u^2 \sigma_v^2 (1 - \rho^2) \quad (17)$$

ここで、 $\sigma_u^2$ 、 $\sigma_v^2$ 、 $\rho$  はそれぞれ流速成分  $u$ 、 $v$  の分散および相関係数である。もし、 $u$ 、 $v$  が互いに独立であると仮

定すると  $\rho = 0$  であるから、 $|\hat{\Sigma}|$  が大きく見積もられてしまう可能性がある。よって、最小AIC(or BIC)は、よりパラメータ数が少ないモデルになり、分解能が低下する。このように、共分散性を正しく考慮しなければ、モデル自身の精度を落とすことになりかねない。

4. 最適な分潮群の探索方法

最も単純に最適な分潮群を選ぶには、各分潮群に対して総当たりで情報量基準(IC)を計算すればよい。ただし、解析対象となる分潮がK個あった場合、全組み合わせ数は $2^k - 1$ 個となるので分潮数が大きい場合は非常に多くの計算時間を要し、現実的ではない。そこで、多変量解析の変数選択方法の一つである後退消去法(田中ら, 1983)を応用した以下の方法を提案する。

- 1) 解析対象とする分潮がすべて含まれた状態にする。
- 2) モデルの調和定数とIC<sub>k</sub>を計算する(kは分潮数)。
- 3) 調和定数は変化させずに、分潮iだけを見捨ててIC<sub>k-i</sub>を計算する。これを解析中の全分潮に対して行う。
- 4) IC<sub>k-i</sub> < IC<sub>k</sub> であり、かつIC<sub>k-i</sub>が最も小さい分潮iを解析対象から外し、2)に戻る。IC<sub>k-i</sub> > IC<sub>k</sub> となる分潮がなくなるまで繰り返す。

上記の後退消去法では、解析対象となる分潮数がK個の場合、調和解析を行う回数はK回以下であり、現実的な計算時間で分潮群の選択を行うことができる。

この方法は分潮iを見捨ててICを算出したIC<sub>k-i</sub>と、実際に分潮iを解析対象から外して調和解析を行った場合のIC<sub>k-i</sub>は同じであるという仮定に基づいている。この検証は後ほど行う。

5. 数値実験

(1) 実験ケース

本研究で示した最尤法による調和解析と最適な分潮群の探索方法の検証を数値的に作成した仮想データを用いて行う。潮流データの調和定数は表-1の通りであり、u, vの平均値はゼロ、u, vの分散および相関係数はそれぞれ370.90, 473.74, -0.21とした。u, vはそれぞれ東方成分、北方成分流速を表す。データは1時間間隔で2006/1/1から1年分と2年分のデータを用意し、解析する情報量基準としてAICとBICを利用した場合の計4ケースを行った(表-2)。実験データの一部を図-1に示す。なお、この数値実験に用いた調和定数および誤差は後述する海洋短波レーダーの観測点(図-4のSt.A)の解析結果をもとに作成した。

(2) 実験結果

各実験ケースについて、前述の方法により最適な分潮群の探索を行った。初期分潮数は60個(気象庁(1999)に掲載されている分潮)からスタートした。分潮数と情報量基準の変化を図-2に示す。case 1Aは7分潮と、設定以外の分潮が分離された場合に最小AICとなり、過適合が

生じている。データ数が増えたcase 2Aの最小AICは10分潮と、さらに余分な分潮が増えている。これは上述したようにデータ数が増えたことによって式(14)右辺第1項の重みが大きくなり、相対的に説明変数減少の影響が小さくなってしまったためである。一方、BICを用いたcase 1B, 2Bはともに5分潮が最小BICであり、その5分潮は設定した分潮と同じであった。BICのケースではデータが増えたことによる悪影響は見られなかった。

実際の調和解析では、数千個あるいはそれ以上のデータ数を用いて解析を行うことも多い。したがって、情報量基準に関してはAICよりもBICを使用する方がより信頼性の高い結果になると考えられる。

表-1 数値実験データの調和定数

分潮	周期(h)	振幅(cm/s)		位相(deg)	
		u	v	u	v
M <sub>2</sub>	12.42	7.68	9.62	66.77	86.85
S <sub>1</sub>	24.00	4.83	4.83	157.81	70.41
S <sub>2</sub>	12.00	4.22	5.20	106.29	116.27
Sa	8765.82	1.62	4.23	16.78	64.83
Ssa	4382.90	0.52	1.93	264.85	252.03

表-2 実験ケース

case	データ期間	情報量基準	分散・相関係数
1A	1年(N=8760)	AIC	$\sigma_u^2 = 370.90,$ $\sigma_v^2 = 473.74,$ $\rho = -0.21$
1B	1年(N=8760)	BIC	
2A	2年(N=17520)	AIC	
2B	2年(N=17520)	BIC	

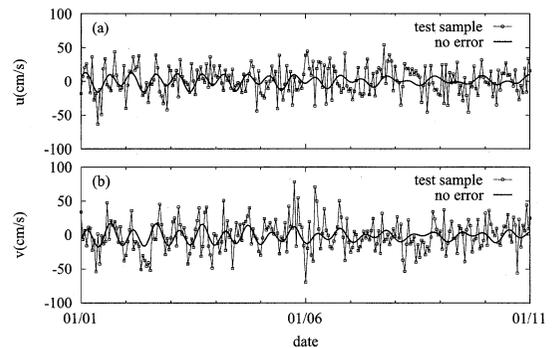


図-1 数値実験データの一部

最終的に得られた調和定数の比較を図-3に示す。どの実験ケースにおいても調和定数は真値とよく一致していた。データ期間が2年よりも1年の方が真値に近い分潮もあるが、全体的にはデータ期間が長くなると精度は向上した。

case 1Aと1Bおよびcase 2Aと2Bの結果は、最適モデルの分潮数が異なるものの、調和定数はほとんど同じであった。これは、ある分潮を調和解析の対象から外した場合も、他の分潮の調和定数には影響を与えないことを

意味している。従って、4節で用いた  $IC_{k,i} \approx IC_{k-1}$  という仮定は問題ないものと考えられる。

なお、ここでは示さないが、テストサンプルの分潮数を変えた場合や、誤差を変化させた場合の数値実験も行った。その場合でも、上記の数値実験と同様の結果が得られた。

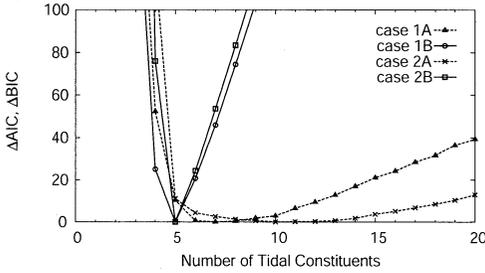


図-2 分潮数と情報量基準； $\Delta AIC$ は最小AICとの差， $\Delta BIC$ は最小BICとの差

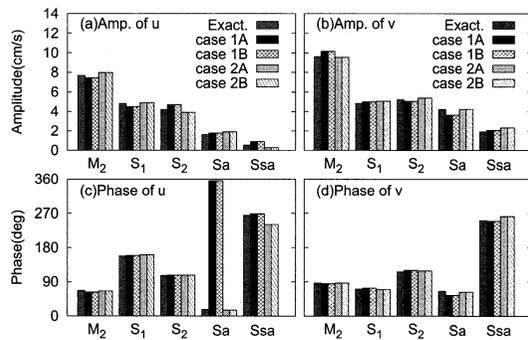


図-3 各実験ケースの調和定数

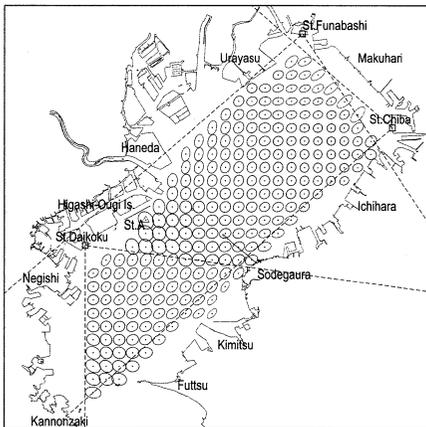


図-4 海洋短波レーダーの観測局，および観測地点の誤差楕円；点線は各観測局のレーダー観測範囲

## 6. 海洋短波レーダーへの適用例

### (1) 観測機器の諸元

東京湾ではHF帯の周波数を用いた海洋短波レーダー（以下、HFレーダー）で観測を行っており、そのリアルタイムデータはweb上で公開されている（国土交通省，2008）。アンテナを設置している観測局は千葉局、船橋局、大黒局の3ヵ所であり、その観測地点を図-4に示す（データ数が少ない地点は除く）。ほとんどの観測地点は2局からのレーダーで流速値を算出しており、大黒局の観測範囲である根岸沖付近は3局のレーダーを用いている。観測は全局同時にレーダー照射を毎正時に9分間行い、得られた流速の平均値を毎正時の流測値としている。

解析に用いたデータ期間は2006年2月18日～2008年3月31日であるが、場所によってデータ取得率は大幅に異なる。最もデータが少ない地点は634日相当、最もデータが多い地点は744日相当であり、平均して702日相当のデータが取得されている。本手法ではデータが欠測していても解析に問題はないが、データ数が少なくなる分だけ精度は低下する。

### (2) 調和解析結果

本手法によって得られた誤差分布を図-4に示す。誤差楕円が大きかった場所は東扇島付近、富津沖、幕張付近であり、これらの場所は観測局からのレーダー交差角度が90度から大きくずれている。基地局が2つの場合はレーダー交差角度が90度で最も誤差が小さくなるので（Nadaiら，1999）、これらの場所はレーダーによる観測誤差の影響が大きいと考えられる。富津沖よりも東扇島沖の誤差が大きい理由は、富津沖は3局のレーダーを使用しているためである。一方、誤差楕円が小さい場所は市原沖であり、この場所は千葉局と船橋局のレーダー交差角度が90度に近くなる場所である。このように、計算値と観測値の誤差はHFレーダーの交差角度によってほぼ説明できる。よって、調和解析結果から得られた誤差は、風の擾乱などに起因する誤差成分よりもHFレーダーに起因する誤差成分が支配的であることを示している。

$M_2$ 分潮の潮流楕円を図-5に示す。計算結果は富津沖で長軸角度の差異が大きいが、全体的には観測結果とほぼ一致している。富津沖での誤差は上述したようにHFレーダーの観測誤差による影響と考えられる。

HFレーダーの流速データから分解できた分潮数は8～21分潮であり、どの地点においても季節変動周期であるSa, Ssa分潮を分解することができた。Sa, Ssa分潮に平均流速を加えた流速は海流の季節変動成分と解釈することができる。図-6に夏期(8月1日)と冬期(2月1日)の季節変動成分流速を示す。夏期には湾奥部に時計回りの環流が存在しており、森川ら(1986)や中川ら(1997)による観測結果とほぼ一致する。一方、冬期の季節変動流速場は全域で湾口へ向かう流れとなった。これは中辻ら(1995)

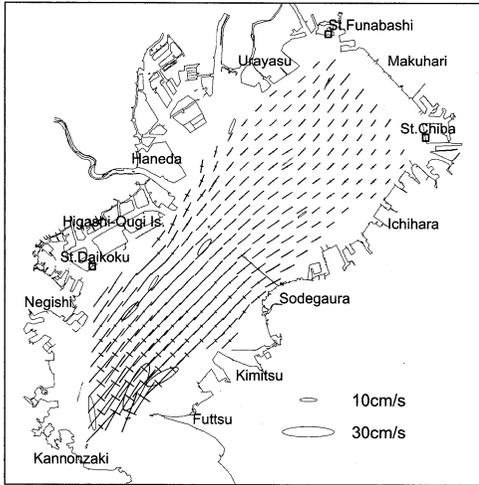


図-5 海洋短波レーダの流速データから求めたM<sub>2</sub>分潮の潮流  
楕円；直線は計算値，楕円は既往の観測値(日向，  
2005；森川ら，1986)

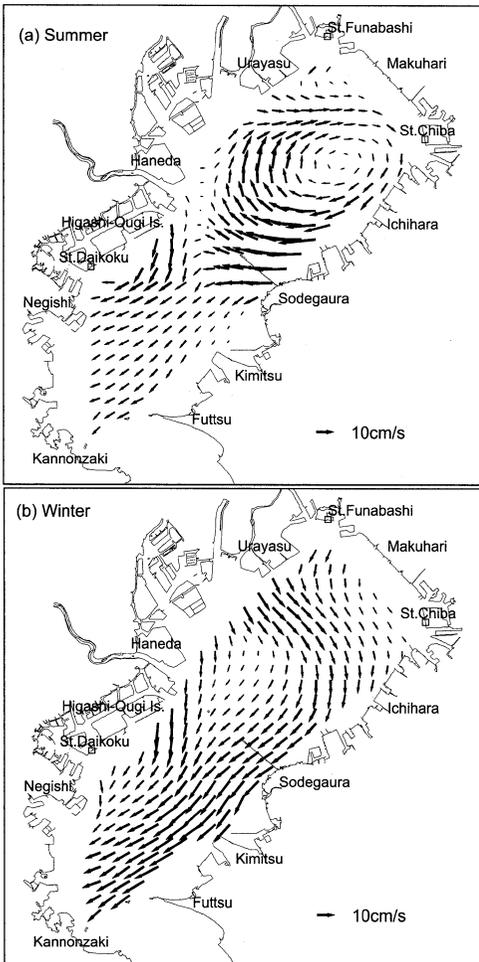


図-6 東京湾の季節変動成分流速；(a)夏期，(b)冬期

の計算結果に近く，季節風である北風の影響を強く受けていると考えられる。

### 7. まとめ

本研究では最尤法による潮流の調和解析方法を示した。また，情報量基準を利用した最適な分潮群の探索方法を提案し，数値実験により本手法の精度と特徴を考察した。本研究で得られた知見は以下の通りである。(1)多数のデータを使用して調和解析する場合は，AICよりもBICを情報量基準として採用する方が高い信頼性を得た。(2)本研究の方法をHFレーダーによる観測例に適用し，その妥当性を検討した。(3)HFレーダーの解析結果から東京湾表層の季節変動流速場を求め，夏期には時計回りの還流が発生し，冬期には還流が消滅する結果を得た。

今後の課題として，流速観測にはノイズ的なエラーが発生することが多く，観測値の重みを自動的に判定してよりロバストな推定ができる方法の開発が挙げられる。

謝辞：海洋短波レーダーに関する情報は関東地方整備局横浜港湾空港技術調査事務所から頂いた。港湾空港技術研究所の鈴木高二朗主任研究官からは有益な助言を頂いた。ここに記して謝意を表します。

### 参考文献

池谷毅(1993)：潮汐の調和解析への最小AIC推定法の適用，  
海岸工学論文集，第40巻，pp.216-220。  
気象庁(1999)：潮汐の推算と調和分析，海洋観測指針(第2部)，  
付録F，気象業務支援センター，pp.83-90。  
国土交通省関東地方整備局(2008)：東京湾環境情報センター，  
<http://www.tbeic.go.jp/>。  
田中豊・脇本和昌(1983)：多変量統計解析法，現代数学社，  
43。  
中川康之・八木宏・鯉淵幸生・日向博文・稲垣聡・小倉久子  
(1997)：東京湾内湾部における成層期の流動特性に関する  
現地観測，海岸工学論文集，第44巻，pp.401-405。  
中辻啓二・尹鐘星・白井正興・村岡浩爾(1995)：東京湾にお  
ける残差流系に関する三次元数値実験，海岸工学論文集，  
第42巻，pp.386-390。  
日向博文(2005)：HFレーダーによる東京湾のM<sub>2</sub>潮流観測，  
国総研資料，212号，17 p。  
村上和男(1981)：最小自乗法による潮汐・潮流の調和解析と  
その精度，港研資料，369号，38 p。  
森川雅行・村上和男(1986)：東京湾における長周期流れの変  
動特性，港研資料，550号，50 p。  
Akaike, H. (1974) : A new look at the statistical model identi-  
fication, IEEE Trans. Autom. Contr., AC-19, pp.716-723.  
Anderson, T. W. (1958) : An Introduction to Multivariate  
Statistical Analysis, John Wiley & Sons, Inc., 374 p.  
Burnham, K. P. and David R. A. (2002) : Model Selection and  
Multimodel Inference. 2nd ed., Springer, 488 p.  
Nadai, A., H. Kuroiwa, M. Mizutori and S. Sakai (1999) :  
Measurement of Ocean Surface Currents by the CRL HF  
Ocean Surface Rader of FMCW Type. Part 2. Current  
Vector, J. Oceanogra., Vol.55, pp.13-30.  
Schwarz, G. (1978) : Estimating the dimension of a model,  
Annals of Statistics, Vol.6, pp.461-464.