

# Support Vector Machine を用いた極成層圏雲表面積の推定

池田 奈生<sup>†</sup> 芦田 尚美<sup>††</sup> 林田 佐智子<sup>†</sup>

オゾンホールとして広く知られている化学的オゾン破壊は、冬季北極域の高度約 20 km 付近で起こる。この化学的オゾン破壊量は、気温が約  $-78$  度以下になると発生する極成層圏雲 (Polar Stratospheric Clouds: PSC) の表面積に大きく依存する。したがって、化学的オゾン破壊量の定量的推定のためには PSC 表面積を精度良く推定する必要がある。そこで本研究では光化学ボックスモデルを用いて、PSC 表面積と  $\text{ClONO}_2$  との関係を理論的に決定し、これをサポートベクターマシン (Support Vector Machine: SVM) に入力して学習させる。そして観測された  $\text{ClONO}_2$  や気温などの条件を入力することで PSC 表面積を推定するシステムを構築する。

## Estimate of the Surface Area Density of Polar Stratospheric Clouds by Using the Support Vector Machine

NAO IKEDA,<sup>†</sup> NAOMI ASHIDA<sup>††</sup> and SACHIKO HAYASHIDA<sup>†</sup>

Chemical ozone destruction occurs at around 20 km altitude in the Arctic winter, and the amount of ozone loss depends on the surface area density of Polar Stratospheric Clouds (PSC) which appear at the temperatures below about  $-78^\circ\text{C}$ . Therefore, a precise estimate of the PSC surface area density is highly required to investigate the amount of ozone loss quantitatively. In this study, we determine the relations among the surface area density of PSCs and other key parameters such as  $\text{ClONO}_2$  concentration by using a chemical box model. The data of those key parameters obtained from the model are input into Support Vector Machine. Then we construct a system to estimate PSC surface area density by using the Support Vector Machines.

### 1. はじめに

冬季極域での大規模な化学的オゾン破壊は、一般にオゾンホール現象として知られている。これは冬季極域に発生する PSC 上で、オゾンを破壊しない  $\text{ClONO}_2$  や  $\text{HCl}$  といった不活性塩素が、オゾンを破壊する  $\text{Cl}$  や  $\text{ClO}$  といった活性塩素に多量に変換され、オゾンと反応するためである。したがって、PSC はオゾン破壊に大きな影響を及ぼす。

PSC 上での反応は PSC の表面積に大きく依存する。したがって PSC の表面積は化学的オゾン破壊量を理解するうえで重要なパラメータである。しかしながら冬季北極では、年ごとに PSC 発生量が大きく異なり、さらに PSC の観測事例も少ないため、見積りが難しい。

そこで本研究では、PSC に大気中の存在量が大きく依存する化学種である  $\text{ClONO}_2$  の観測値と、 $\text{ClONO}_2$  に関連するパラメータから PSC 表面積を推定するシステムの構築を目的とする。本論文では光化学ボックスモデルを用いて PSC 表面積と  $\text{ClONO}_2$  との関係を理論的に決定し、これを SVM に入力して学習させる。そして観測された  $\text{ClONO}_2$  や気温などの条件を入力することで PSC 表面積を推定するシステムを構築する。SVM は Vapnik ら<sup>1),2)</sup> によって提案された 2 クラス分類器で、少ない学習データで特徴の次元を増やすことが可能であり、本研究で行うようなシステムの構築に適している。以上のように構築されたシステムについての評価・検討を行う。

以下、本論文の構成は次のとおりである。まず、2 章では本研究の背景について述べる。3 章では作成したデータセットについて述べ、4 章では SVM を用いた PSC 表面積推定システムの構築方法と結果について述べる。5 章では構築したシステムについての考察を行い、6 章でまとめと今後の課題を述べる。

<sup>†</sup> 奈良女子大学大学院人間文化研究科  
Graduate School of Humanities and Sciences, Nara  
Women's University

<sup>††</sup> 奈良女子大学理学部  
Faculty of Science, Nara Women's University

## 2. 背景

ここでは本研究の背景として、化学的オゾン破壊過程、PSC（極成層圏雲）について述べる。

### 2.1 化学的オゾン破壊過程

冬の極域での大規模な化学的オゾン破壊の主な原因は、大気中に多量に放出された Cl や ClO などの活性塩素が、オゾンと反応するためである。この活性塩素は次のようなプロセスを通して大気中に多量に放出される。

冬の極域では西風が強くなり、極渦という風の壁を形成する。これにより、極渦の内側と外側の空気の交換が妨げられる。さらに日照が少なくなることから極渦内の気温が大きく低下する。そして気温が約  $-78^{\circ}\text{C}$  まで低下すると、PSC が発生する。PSC は高度 20 km 付近に発生するため、日没後も太陽光がわずかにあたり、虹色に反射する。この色がアコヤ貝の内側の色彩に似ていることから、真珠母雲とも呼ばれている。この雲上では、オゾンを破壊しない HCl や  $\text{ClONO}_2$  といった不活性塩素が、不均一反応によって活性塩素へと変換される。不均一反応とは、異なった相間で起こる反応のことである。ここでは、図 1 に示すように、大気中の  $\text{ClONO}_2$ （図中の A 分子）が、液滴または固体の PSC 内に存在する分子と反応することである。この結果、大気中に多量の活性塩素が放出される。

### 2.2 PSC（極成層圏雲）

不均一反応の反応速度は粒子表面積に比例する。このため、PSC の表面積は化学的オゾン破壊量を理解するうえで重要なパラメータである。PSC 表面積は直接観測することはできず、光学的厚さの観測より間接的に導かれるものである。PSC は冬季南極域・北極域で存在が確認されており、冬季南極域では、気温が PSC 生成気温より十分に低下するため、PSC は空

間・時間的にほぼ一様に分布する。しかしながら冬季北極域では、南極よりも気温が高く、PSC の分布は空間・時間的に非一様である。加えて冬季北極域では年ごとに気温変化が大きく異なり、さらに PSC は観測事例が少ないため、PSC 表面積の定量的な見積りは難しい。

化学的オゾン破壊量をシミュレーションで推定するには、PSC 表面積の見積りが重要である。冬季北極域では、これまでさまざまなモデルを用いて化学的オゾン破壊量の推定が行われた<sup>3)</sup>。これらの結果では、比較的暖かく PSC の発生が少ない年のオゾン破壊量の推定結果は、観測結果から推定されるオゾン破壊量とよく一致していた。しかしながら気温が低く PSC の発生が多い年は、モデル結果が観測値よりも低く、モデルがオゾン破壊量を過小評価していた。この原因として、気温のバイアスの存在が指摘されている<sup>3)</sup>。PSC は気温が約  $-78^{\circ}\text{C}$  以下になると発生するが、この気温付近では、気温の変動に対して表面積の変動が非常に大きい。したがって、気温バイアスが存在すると PSC 表面積の推定に誤差が生じ、その結果、モデルが化学的オゾン破壊量を過小評価していると考えられる。したがって、モデル計算で定量的に化学的オゾン破壊量を推定するためには、詳細な PSC 表面積データが必要である。

1996 年に打ち上げられた日本の衛星「みどり」に搭載された改良型大気周縁赤外分光計 (Improved Limb Atmospheric Spectrometer: ILAS) は、世界で初めて冬季北極域での  $\text{ClONO}_2$  の長期間連続観測を行った<sup>4)</sup>。大気中の  $\text{ClONO}_2$  量は PSC の発生に大きく依存するため、 $\text{ClONO}_2$  量の情報を使えばより現実に近い PSC 表面積の推定が可能となる。ILAS が観測を行った 1997 年の 1 月から 3 月の北極域は、気温が低く、PSC が多数存在していたことが Hayashida ら<sup>5)</sup> によって報告されている。また、WMO<sup>3)</sup> は、この年の北極域でのモデル計算結果がオゾン減少量を過小評価していたことを報告している。したがってこの年の PSC 表面積を推定することにより、モデル計算によるオゾン破壊量の推定精度の向上が期待される。

## 3. 学習データセット作成

この章では、SVM に与えるデータセットの作成に必要な、 $\text{ClONO}_2$  観測データ、シミュレーションモデル、モデル計算・データ作成方法について説明する。

### 3.1 $\text{ClONO}_2$ 観測データ

$\text{ClONO}_2$  は地上観測、飛行機観測、衛星観測などの方法で観測が行われてきた。地上観測・飛行機観測

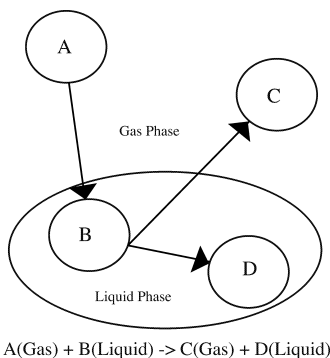


図 1 不均一反応の模式図

Fig. 1 Schematic diagram of a heterogeneous reaction.

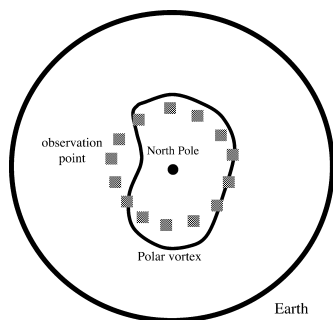


図2 ILASの観測位置の模式図．北極点を中心としており、■が観測位置を示す

Fig.2 Schematic diagram of ILAS observation point.

は時間・空間的に限られた範囲でしか観測ができないため、オゾン破壊開始から終了までの長期間、広域でCIONO<sub>2</sub>の時間変化を観測することはできない．一方、衛星観測は長期間・広域で観測を行うことが可能であるため、オゾン破壊開始から終了までの長期間、CIONO<sub>2</sub>の変動を解析することが可能である．

ILAS センサは冬季北極域でのCIONO<sub>2</sub>の長期観測に初めて成功した<sup>4)</sup>．ILASは極軌道衛星である「みどり」に搭載された太陽掩蔽法のセンサで、1996年11月から1997年6月まで南北高緯度でO<sub>3</sub>、HNO<sub>3</sub>、N<sub>2</sub>O、CIONO<sub>2</sub>、H<sub>2</sub>O、エアロゾルなどの大気微量成分の高度分布を観測した．図2にILASの観測位置の概略を示す．観測は1日約14回行われ、観測データは極渦内外を網羅している．

本研究では最新のVersion 7のデータを使用する．Version 7以前のデータでは、PSC存在時に、観測されたスペクトルからガス種のデータをうまく導出することができなかった．しかしながら、Version 7ではデータ導出のアルゴリズムが改善され、PSC発生時でもガス種のデータを導出することが可能となった<sup>6)</sup>．この観測データの中から北極域の1月から3月、オゾン破壊量が最も多い温位475 K（高度約20 km）の極渦内のデータを使用する．温位とは、ある気圧、気温状態での空気を断熱的に地表面気圧まで持ってきたときの温度のことである．一般に温位は位置・高度が変化しても保存されるため、空気は同じ温位面上を移動すると考えてよい．使用するCIONO<sub>2</sub>観測値を図3に示す．ここで、この高度では18日以降に極渦が形成されたと定義されているため、18日以前のデータは使用しない．CIONO<sub>2</sub>は一度不均一反応が起こると大きく減少し、その後増加傾向を示す．このとき不均一反応が再度起こると敏感に反応して減少する．図3では20日付近と40–60日付近で不均一反応が起こっている．また、CIONO<sub>2</sub>はばらついた値をとるが、こ

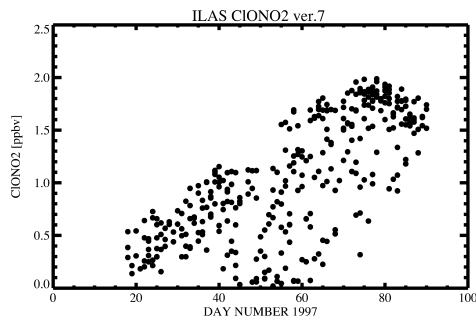


図3 ILASが観測した温位475 K、極渦内での1997年1–3月のCIONO<sub>2</sub>の観測値．縦軸はCIONO<sub>2</sub> [ppbv]、横軸は1997年1月1日を1とした日数をそれぞれ表す

Fig.3 CIONO<sub>2</sub> mixing ratio measured by ILAS.

れは観測値ごとに、観測される以前に経験した不均一反応の規模が異なるためである．本研究ではこの観測値を使用する．

### 3.2 シミュレーションモデル

次にCIONO<sub>2</sub>とPSCの関係を理論的に決定するためのシミュレーションモデルについて説明する．流跡線計算モデルは、観測された空気塊が移動した位置を計算するモデルである．また、光化学ボックスモデルは大気中の化学種の濃度の時間変化を計算するモデルである．

#### 3.2.1 流跡線計算モデル

ILASは大気中の空気塊を観測し、その中に含まれている大気微量成分の濃度を観測している．ILASで観測された空気塊は、空気塊ごとに異なる軌跡により輸送されたものである．したがって空気塊がこれまでに経験した光化学反応を再現するためには、観測値ごとに空気塊が移動してきた場所を推定し、その位置で光化学ボックスモデルを計算させなければならない．そこで本研究では空気塊の輸送経路を推定する流跡線計算を行うために、宇宙航空研究開発機構/地球観測研究センターが開発した流跡線計算モデル、EORC-TAM (Earth Observation Research Center Trajectory Analysis Model<sup>7)</sup>)を使用する．このモデルでは入力に初期地点の緯度、経度、高度、時刻を与えると、指定時刻までの緯度、経度、高度、気温を1時間ごとに出力する．ここで使用されている気象場のデータはEuropean Centre for Medium-Range Weather Forecasts (ECMWF)の全球客観解析データである．また本研究で行う計算方法は475 K温位面に沿って観測日時から7日前まで遡る後方流跡線計算である．

#### 3.2.2 光化学ボックスモデル

次に光化学ボックスモデルについて説明する．光化学ボックスモデルとは各化学種の濃度の時間発展を、

連立微分方程式として表し、時間積分して計算するものである。このモデルは主に化学反応プロセスのシミュレーションなどに使用される。本研究では入力に各化学種の初期値、緯度、経度、高度の時間変化を与え、指定時間後の各化学種の濃度を出力するモデルとして使用する。本研究で使用する光化学ボックスモデルは Kagawa ら<sup>8)</sup> において使用されたモデルを基礎として構築する。このモデルは成層圏の化学反応を表現できるように 59 種類の化学種、101 種の気相反応、48 種の光解離反応、7 種の不均一反応を含む。反応速度定数、光解離定数は Sander ら<sup>9)</sup> を使用する。多重散乱を含めた各高度での太陽放射量は、アメリカの The National Center for Atmospheric Research / Atmospheric Chemistry Division (NCAR/ACD) が無料配布している TUV (Tropospheric Ultraviolet and Visible Radiation Model) モデル<sup>10)</sup> の Version 4.4 の結果を使用する。TUV の計算に必要なオゾンの高度分布は ILAS の 1 月の月平均値を与える。本モデルでは、液滴の PSC (Supercooled Ternary Solution (STS): 三成分過冷却液滴) を仮定する。大気中のエアロゾルは、主に  $\text{H}_2\text{SO}_4$  を含んだ液滴である。ここで、エアロゾルとは大気中に浮遊する微粒子のことである。気温が約  $-80^\circ\text{C}$  まで低下すると、この液滴は大気中の  $\text{HNO}_3$  を取り込んで急激に成長し、PSC となる。この成長過程を表現するため、Carslaw ら<sup>11)</sup> が提唱した熱力学平衡モデルを使用する。このモデルは飛行機観測の結果によって十分検証されており、このモデルを用いることによりエアロゾルから PSC への変化過程を理論的に求めることが可能となる。各化学種の初期値は Kagawa ら<sup>8)</sup> で用いられた初期値を使用する。

### 3.3 データセット作成

これまでに説明したモデルを使用し、次の手順で計算を行う。

- (1) EORC-TAM を用いて、すべての ILAS 観測地点より温位 475 K で、観測日時から観測 7 日前までの流跡線を計算する。
- (2) 観測の 7 日前から観測日まで、EORC-TAM の出力位置情報にそって光化学ボックスモデル計算を行う。このとき気温バイアスを考慮して EORC-TAM で出力された気温履歴と、この気温履歴プラスマイナス  $5^\circ\text{C}$  の間を  $1^\circ\text{C}$  刻みで計 11 通り計算する。1997 年の 1 月 1 日から 3 月 31 日まで、極渦内では計 311 点の観測結果が存在する。したがって、311 (観測地点数)  $\times$  11 (気温履歴) の計 3421 本のモデル計算を

行う。

- (3) 観測イベントごとに、 $\text{ClONO}_2$  観測値と最も近いモデル計算結果を選択する。
- (4) 選択したモデル計算結果から、 $\text{ClONO}_2$  に関連する以下の 7 つのパラメータを用いてデータセットを作成する。
  - 選択したパラメータは
    - (a) 7 日間の PSC 表面積積分値
    - (b)  $\text{ClONO}_2$  観測値
    - (c) 1997 年 1 月 1 日を 1 とする日数
    - (d) 7 日間の積分日照時間
    - (e) 7 日間の  $\text{ClO}$  積分量
    - (f) 7 日間の  $\text{NO}_2$  積分量
    - (g) 気温のバイアス

の 7 つである。ここで、 $\text{ClO}$  積分量、 $\text{NO}_2$  積分量は  $\text{ClONO}_2$  の生成に関わるパラメータである。そして積分日照時間は  $\text{NO}_2$  の生成と  $\text{ClONO}_2$  の光解離による消滅に関わるパラメータである。そして気温のバイアスは客観解析データ (ECMWF) の気温履歴と  $\text{ClONO}_2$  をモデルで推定したときの気温履歴の差、日数は季節パラメータである。次章で構築する SVM にこれらのデータを与える。このとき、(a) を出力教師データ、(b)–(g) を入力データとする。

## 4. PSC 表面積推定システムの構築と検証

### 4.1 サポートベクターマシン

PSC 表面積の推定を行うにあたり、関数近似問題と多クラス分類問題という 2 種類の問題の解法が考えられる。関数近似問題として定式化を行う場合、分散が大きくなるという問題が生じる可能性がある。そこで本研究では多クラス分類問題として取り扱う。

SVM は多次元の特徴ベクトルを、2 つのクラスに識別する 2 値クラスの特徴ベクトルである。識別器は  $f(x) = \langle w \cdot x \rangle + b$  で表され、 $f(x)$  の値によってクラスに分類する。ここで  $w$  は重みベクトル、 $x$  は入力ベクトル、 $b$  はバイアスである。 $f(x) = 0$  を満たす面を超平面と呼ぶ。これを複数組み合わせることにより、多クラス分類問題への応用が可能である。また、SVM の特徴としてマージン最大化があげられる。マージンとは超平面と特徴ベクトル間の最小距離である。このマージンを最大化することによって学習データの識別精度を高めている。これにより、学習データが少なくても特徴ベクトルの次元を増加させることが可能である。さらに、非線形問題への対応はカーネル関数を用いて実現されている。本論文では、LIBSVM<sup>12)</sup> を用

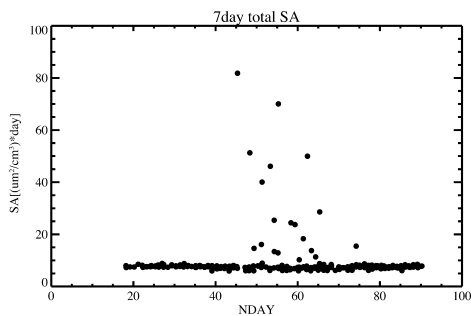


図 4 7 日間の PSC 表面積積分値  $[(\mu\text{m}^2/\text{cm}^3) * \text{day}]$ . 横軸が 1997 年 1 月 1 日を 1 とした日数, 縦軸が PSC 表面積の積分値をそれぞれ表す

Fig. 4 7 day's total PSC surface area.

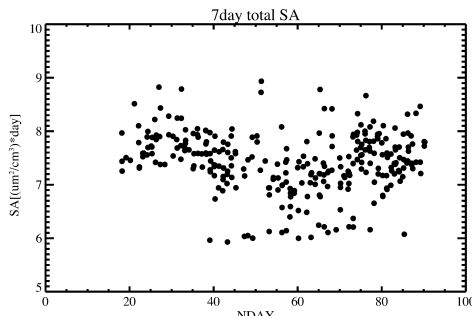


図 5 図 4 のうち, PSC 表面積積分値が 5-10  $[(\mu\text{m}^2/\text{cm}^3) * \text{day}]$  の範囲を拡大した図

Fig. 5 7 day's total PSC surface area, Y axis scale is 5-10  $[(\mu\text{m}^2/\text{cm}^3) * \text{day}]$ .

いて SVM の実装を行う. LIBSVM は台湾国立大学の Lin らによって作成された SVM のライブラリで, シンプルで使いやすいツールとして提供されている. カーネル関数は, クラスと属性の関係が非線形でも対応可能な Radial Basis Function (RBF) を使用する.

#### 4.2 PSC 表面積積分値データ

本研究でクラス分類するパラメータは PSC 表面積積分値である. 図 4 に, 作成したデータセットの PSC 表面積積分値を示す. この図は, 観測された空気塊が過去 7 日間にどれだけの PSC 表面積を経験したかを示している. 図 4 より大部分のデータが  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  以下であることが分かる. ここで,  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  以上の値はオゾン破壊に大きく影響を及ぼす PSC であることを示している. 一方, 図 5 に PSC 表面積積分値が  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  以下の範囲の分布を示す. この分布では, 表面積積分値が  $6 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  付近の値は PSC ではなく, バックグラウンドに存在する粒子の表面積を示している. また,  $7-10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  付近のデータは, 短時間だけ PSC が存在し, オゾン破壊に影響を及ぼした可能性があることを示している. 観測値の合計 311 点のうち, PSC の表面積積分値が  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  以上は 19 点,  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  未満の点は 292 点であった.

#### 4.3 PSC 表面積推定システムの構築

本節では, SVM を用いた PSC 表面積推定システムを構築し, この結果を示す. PSC 表面積積分値データの分布は 4.2 節で示したように表面積積分値が  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  を境に明らかに特徴が異なるため, 階層的なシステムの構築を試みた. 図 6 にこのシステムの概略図を示す.

まず PSC 表面積積分値が  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  以上のクラスと  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  未満のクラスに分

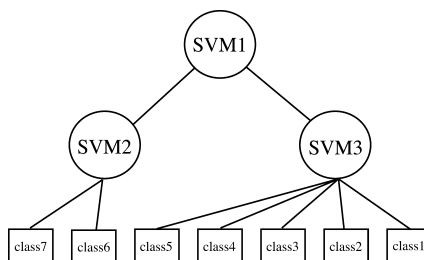


図 6 階層的 SVM の概略図

Fig. 6 Schematic diagram of SVM system.

表 1 クラスごとの PSC 表面積積分値 (SA) の範囲  
Table 1 Total PSC Surface Area range of each class.

class name	SA $[(\mu\text{m}^2/\text{cm}^3) * \text{day}]$
class 1	$SA \leq 6.5$
class 2	$6.5 \leq SA < 7.1$
class 3	$7.1 \leq SA < 7.7$
class 4	$7.7 \leq SA < 8.3$
class 5	$8.3 \leq SA < 10$
class 6	$10 \leq SA < 40$
class 7	$40 \leq SA$

類する SVM を作成した. この SVM を「SVM1」とする. 次に PSC 表面積積分値が  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  以上のクラスに分類されたデータをさらに分類する SVM を作成した. この SVM を「SVM2」とする. ここでのクラス数はデータが少量であることから, クラス分類が可能かどうかを検討するために 2 クラス (クラス 6, 7) での分類を試みた. 一方, PSC 表面積積分値が  $10 (\mu\text{m}^2/\text{cm}^3) * \text{day}$  未満のデータは 300 点近く存在する. PSC 表面積積分値の有効桁数の関係より, データが存在する範囲を等間隔で 5 クラス (クラス 1-5) に分類することを試みた. この SVM を「SVM3」とする. 表 1 に SVM2 と SVM3 の, クラス分類時の PSC 表面積積分値の範囲を示す.

これらの SVM を作成するにあたり, 次の条件で実

表 2 実験を行ったトレーニングデータ数とテストデータ数の平均  
Table 2 Number of training data and test data for each experiment.

experiment	SVM	training	test
experiment 1	SVM1	200	111
	SVM2	13	5.6
	SVM3	187	103.8
experiment 2	SVM1	160	151
	SVM2	10	7.3
	SVM3	150	140.8

表 3 実験 1 の正答率の平均と分散，正答数/誤答数の平均値  
Table 3 Mean, deviation and average number of correct /error answer of experiment 1.

SVM	mean [%]	deviation [%]	correct /error
SVM1	98.56	1.01	109.4/1.6
SVM2	45.71	172.95	3.1/3.7
SVM3	81.09	1.67	84.5/19.7

表 4 実験 1 の SVM1, SVM2 の各クラスの正答数/誤答数の平均値  
Table 4 Average number of correct /error answer of SVM1 and SVM2 (experiment 1).

SVM	class	correct	error
SVM1	class 1-5	103.8	1.2
	class 6-7	5.6	0.4
SVM2	class 6	3	0.7
	class 7	0.1	1.8

験を行った。

- (1) 311 個の全データのうち，トレーニングデータとテストデータの個数を表 2 に示すように 2 通り（実験 1，実験 2）の実験を行った。
- (2) 各実験では，10 種類のトレーニングデータ集合・テストデータ集合のペアを作成し，10 回の実験を行った。各データ集合は，各クラスに含まれるデータ数の割合が，全データを対象としたときに各クラスに含まれるデータ数の割合と同じになるようにした。

表 3 に実験 1 の結果を示す。また，SVM1 と SVM2 の各クラスごとの正答数，誤答数を表 4 に示す。この結果より，全体を 2 クラスに分類する SVM1 では 10 回の実験の平均値が 98.56% とシステムを構築するのに十分な結果が得られた。次に，PSC 表面積が  $10(\mu\text{m}^2/\text{cm}^3) \cdot \text{day}$  以上の部分を 2 クラスに分類する SVM2 では，10 回の実験の平均正答率が 45.71% となり，識別が難しいことが分かった。一方，PSC 表面積積分値  $10(\mu\text{m}^2/\text{cm}^3) \cdot \text{day}$  以下を 5 クラスに分類する SVM3 では，10 回の実験の平均正答率が 81.09% となった。したがって SVM1 と SVM3 を用いた階層的な PSC 表面積推定システムは，構築が可能である

表 5 実験 2 の正答率の平均と分散，正答数/誤答数の平均値  
Table 5 Mean, deviation and average number of correct /error answer of experiment 2.

SVM	mean [%]	deviation [%]	correct /error
SVM1	98.08	1.09	148.1/2.8
SVM2	57.67	146.39	4.9/3.7
SVM3	77.67	7.04	110.7/31.8

表 6 実験 2 の SVM1, SVM2 の各クラスの正答数/誤答数の平均値  
Table 6 Average number of correct /error answer of SVM1 and SVM2 (experiment 2).

SVM	class	correct	error
SVM1	class 1-5	140.8	1.2
	class 6-7	7.3	1.7
SVM2	class 6	3.8	1.9
	class 7	1.1	0.6

ことが分かった。

次に，実験 2 の結果を表 5 に示す。また，SVM1 と SVM2 の各クラスごとの正答数，誤答数を表 6 に示す。SVM1 ではシステムを構築するのに十分な結果が得られたが，SVM2 と SVM3 では正答率が 80% に満たず，トレーニングデータが 160 個ではシステム構築は難しいことが分かった。

今回の実験を通して，1 回の学習をするのに要した時間は，AMD Sempron 1.833 GHz，メモリ 256 MB の Windows 2000 上で数分程度であった。

## 5. 考 察

CIONO<sub>2</sub> と PSC 表面積の関係をシミュレーションモデルを用いて理論的に決定し，これを SVM に与えることによって PSC 表面積推定システムの構築を試みた。PSC 表面積データの分布に，明らかに異なる 2 つの特徴が見られたため，2 層構造システムの構築を行った。システム上層では，PSC 表面積分布に見られる 2 つの特徴をクラス分けする SVM を作成した。また下層では，上層で分離した 2 つの集合に対して，さらにクラス分けを行う 2 つの SVM を作成した。そして全データ数 311 個をトレーニングデータとテストデータに分け，SVM の学習とテストを行った。

その結果，上層ではシステムを構築するのに十分な正答率が得られ，PSC 表面積分布に見られる 2 つの特徴を十分に識別できることが分かった。一方，下層のシステムでは使用できるデータ数に偏りが存在した。それぞれの SVM が使用できるデータの個数は 19 個と 292 個であった。そこで 19 個のデータを用いる SVM では，分類可能かどうかを検討する目的から 2 つのクラスに分類した。また，292 個のデータを用い

る SVM では、PSC 表面積の有効桁数の観点から、5 つのクラスに分類する実験を行った。結果、トレーニングデータを 200 個使用した場合、5 つのクラスに分類する SVM は正答率 81.09% を示し、5 クラスでの識別が可能であることが分かった。一方、2 つのクラスに分類する SVM は 45.71% の正答率しか得ることができず、識別が難しいことが分かった。したがって本論文で使用したデータセットでは、上層の SVM と下層の 5 クラス分類の SVM を用いて、PSC 表面積推定システムの構築が可能であることが分かった。

下層の 2 クラス分類 SVM のテスト実験の結果、十分な正答率を得られなかった原因としてデータ個数が少ないことがあげられる。下層 2 クラスで分類される PSC 表面積積分値は、気温が大きく低下し PSC が大規模に発生した場合である。したがって、今後 ClONO<sub>2</sub> 観測が増加し、大規模な PSC 発生時の観測データが増加すれば、下層 2 クラス分類の SVM も多クラス分類でのシステム構築が可能であると考えられる。化学的オゾン破壊過程の解析には、PSC 表面積が大きい場合の推定値が重要であるため、本システムの実用化に向けてデータ数の増加が必要である。

本論文で構築した 7 クラスに分類するシステムは「オゾン破壊に大きな影響を及ぼす PSC」、「オゾン破壊に影響を及ぼした可能性がある PSC」、「オゾン破壊に影響しない、バックグラウンドに存在するエアロゾル」の 3 種類を識別することが可能であることが分かった。一般に PSC 表面積データを必要とした解析を行う場合、PSC 表面積は直接観測できないパラメータであるため、すべての観測データに対して 3 章で示したような計算を行う必要がある。本論文で構築したシステムを用いることにより、PSC 表面積以外のデータから PSC 表面積積分値が大きいと判断されたものを重点的に解析することが可能となり、解析時間を短縮できる。ゆえに本研究で構築したシステムは PSC データを用いた解析を行ううえで有用であると考えられる。

将来 ClONO<sub>2</sub> 観測が増加すれば、クラス数を増やすことによって PSC 表面積を具体的な数値として推定が可能になると考えられる。この場合、PSC 表面積積分値が  $10(\mu\text{m}^2/\text{cm}^3) \cdot \text{day}$  以上の部分では  $10(\mu\text{m}^2/\text{cm}^3) \cdot \text{day}$ 、 $10(\mu\text{m}^2/\text{cm}^3) \cdot \text{day}$  未満では  $1(\mu\text{m}^2/\text{cm}^3) \cdot \text{day}$  の間隔でクラス分類を行えば、有効な PSC 表面積を推定することができる。

## 6. ま と め

本研究での最終的な目的は、SVM を用いて ClONO<sub>2</sub>

観測値とこれに関わるパラメータから、PSC 表面積を推定するシステムを構築することである。本論文ではこの研究の第 1 歩として、光化学ボックスモデルを用いて PSC 表面積と ClONO<sub>2</sub> との関係を理論的に決定し、これを SVM に学習させることによって、観測された ClONO<sub>2</sub> や気温などの条件から PSC 表面積を推定するシステムの構築を試みた。システム構築にあたり、PSC 表面積に明らかに異なる 2 つの特徴が見られたため、2 層構造システムの構築を行った。

その結果、PSC 表面積データ分布の明らかな 2 つの特徴については、システムを構築するのに十分な精度で識別することが可能であった。このことは、オゾン破壊に大きな影響を及ぼす PSC を高い精度で判別できたことを示しており、大気化学の分野への貢献が可能である。また下層の SVM においてもデータ数が十分に確保できる場合は識別可能であった。しかしながら、データ数が少ない部分では識別が困難であった。今後、十分なデータ数が確保できる場合には構築が可能であると考えられる。

今後の課題として以下のことが考えられる。1 点目は観測値の誤差の評価である。本論文では、ClONO<sub>2</sub> 観測値の誤差の評価を行わずにシステム構築を行った。精度の高いシステムを構築するためには、今後誤差をどのように考慮するか検討する必要がある。2 点目はモデル計算に使用した大気中の各化学種の初期値である。今回は Kagawa ら<sup>8)</sup> で作成された気候値を使用した。このほかに、観測値の平均値などを用いることによって、モデル計算の確度を高めることが可能である。以上の課題を検討していくことによって、実用的な PSC 表面積推定システムを構築することができると考えられる。

謝辞 流跡線ツールを提供して下さった宇宙航空研究開発機構/地球観測研究センターの皆様、ILAS のデータを提供して下さった ILAS サイエンスチームの皆様、特に Version 7 データの提供に対し、国立環境研究所、ILAS チームの皆様にご感謝いたします。TUV は NCAR/ACD で提供いただきました。

## 参 考 文 献

- 1) Vapnik, V. and Lerner, A.: Pattern recognition using generalized portrait method, *Automation and Remote Control*, Vol.24 (1963).
- 2) Vapnik, V.: *The Nature of Statistical Learning Theory*, Springer-Verlag (1995).
- 3) World Meteorological Organization: *Scientific assessment of ozone depletion*, Report No.47 (2003).

- 4) Nakajima, H., Sugita, T., Irie, H., Saitoh, N., Kanzawa, H., Oelhaf, G., Wetzell, G., Toon, G.C., Sen, B., Blavier, J.-F., Traub, W.A., Jucks, K., Johnson, D.G., Yokota, T. and Sasano, Y.: Measurements of  $ClONO_2$  by the Improved Limb Atmospheric Spectrometer (ILAS) in high-latitude stratosphere: New products using version 6.1 data processing algorithm, *J. Geophys. Res.*, Vol.111, p.D11S09 (2006).
- 5) Hayashida, S., Saitoh, N., Kagawa, A., Yokota, T., Suzuki, M., Nakajima, H. and Sasano, Y.: Arctic polar stratospheric clouds observed with the Improved Limb Atmospheric Spectrometer during winter 1996/1997, *J. Geophys. Res.*, Vol.105, pp.24, 715–24, 730 (2000).
- 6) Oshchepkov, S., Sasano, Y., Yokota, T., Nakajima, H., Uemura, N., Saitoh, N., Sugita, T. and Matsuda, H.: ILAS data processing for stratospheric gas and aerosol retrievals with aerosol physical modeling: Methodology and validation of gas retrievals, *J. Geophys. Res.*, Vol.111, p. D02307 (2006).
- 7) Matsuzono, T., Sano, T. and Ogawa, T.: Development of the Trajectory Analysis Model (EORC-TAM), EORC Technical report (1998).
- 8) Kagawa, A. and Hayashida, S.: Analysis of ozone loss in the Arctic stratosphere during the late winter and spring of 1997 using the Chemical Species Mapping on Trajectories (CSMT) technique, *J. Geophys. Res.*, Vol.108, No.D22, p.4698 (2003).
- 9) Sander, S.P., Friedl, R.R., Golden, D.M., Kurylo, M.J., Huie, R.E., Orkin, V.L., Moortgat, G.K., Ravishankara, A.R., Kolb, C.E. and Molina, M.J.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Jet Propulsion Laboratory (2003).
- 10) Madronich, S.: Photodissociation in the atmosphere: 1. Actinic flux and the effect of ground reflections and clouds, *J. Geophys. Res.*, Vol.92, pp.9740–9752 (1987).
- 11) Carslaw, K.S., Luo, B. and Peter, T.: An analytic expression for the composition of aqueous  $HNO_3 - H_2SO_4$  stratospheric aerosols including gas phase removal of  $HNO_3$ , *Geophys. Res. Lett.*, Vol.22, No.14, pp.1877–1880 (1995).
- 12) Chang, C.-C. and Lin, C.-J.: *LIBSVM: A library for support vector machines* (2001). Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

(平成 18 年 8 月 17 日受付)

(平成 18 年 10 月 6 日再受付)

(平成 18 年 11 月 24 日再々受付)

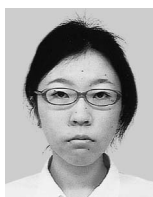
(平成 18 年 12 月 11 日採録)



池田 奈生 (学生会員)

昭和 53 年生。平成 14 年奈良女子大学理学部情報科学科卒業。平成 16 年同大学院人間文化研究科情報科学専攻博士前期課程修了。現在同研究科複合現象科学専攻博士後期課程在

学中。



芦田 尚美

昭和 60 年生。平成 17 年国立明石工業高等専門学校卒業。準学士 (工学) を同校より取得。平成 17 年奈良女子大学理学部情報科学科進学、

在学中。パターン認識に関する研究に従事。



林田佐智子 (正会員)

昭和 32 年生、昭和 55 年京都大学理学部卒業、昭和 60 年名古屋大学大学院理学研究科満期退学、同年理学博士取得。同年国立環境研究所 (旧国立公害研究所) 研究員、平成 6 年奈良女子大学理学部情報科学科助教授、平成 13 年同教授。衛星リモートセンシングによる成層圏オゾン関連研究、大気微量成分解析等の研究に従事。SPARC 国際科学運営委員、日本学術会議特任連携会員、気象庁気候問題懇談会委員等を兼務。米国地球物理連合学会会員、日本気象学会理事、リモートセンシング学会会員ほか。