

地球水循環データアーカイブシステムにおける 異種データ相互解析機能の実装

根本 利弘[†] 小池 俊雄[‡] 喜連川 優[†]

[†] 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

[‡] 東京大学大学院工学系研究科 〒113-8656 東京都文京区本郷 7-3-1

E-mail: [†] {nemoto,kitsure}@tkl.iis.u-tokyo.ac.jp [‡] tkoike@hydra.t.u-tokyo.ac.jp

あらまし

地上観測, 衛星観測, シミュレーションにより作成された地球水循環データの相互比較, 統合などの解析処理を行うためのアーカイブシステムとそのユーザインタフェースについて述べる. 地球水循環データは, データの種類やデータ作成機関により異なる次元, 座標系, データフォーマットを有するが, 本システムは, 利用者にデータの差異を意識させることなく, 相互に利用することを可能としている. アーカイブシステム内における多様な地球水循環データの統一的表現法, 相互比較するための座標軸, 時間軸の変換法, 複数データ比較のためのデータ表示を行うユーザインタフェースについて説明する.

キーワード 科学データ, アーカイブシステム

Implementation of Mutual Analysis Functions for Different Kinds of Data on Global Water Cycle Data Archiving System

Toshihiro NEMOTO[†] Toshio KOIKE[‡] and Masaru KITSUREGAWA[†]

[†] Institute of Industrial Science, The University of Tokyo 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

[‡] School of Engineering, The University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656 Japan

E-mail: [†] {nemoto,kitsure}@tkl.iis.u-tokyo.ac.jp, [‡] tkoike@hydra.t.u-tokyo.ac.jp

Abstract We describe the data archive system and its user interface for comparing and analyzing various global water cycle data generated by ground observations, satellite observations and simulation for weather forecasting. The global water cycle data have different dimensions, coordinates and data format depending on their type and their supplier. The data archive system we are now constructing can offer the users the environment to use the data together without considering the difference among the data. In this paper, we explain the unified data representation of the various global water cycle data in the system, the methods for alignment and arrangement of the special and the temporal axes and the user interface to visualize the data to compare them.

Keyword Scientific data, Archiving system

1. はじめに

地球の表面には莫大な量の水が存在している. これらの水は海面・陸面から蒸発して大気中に放出され, 大気中を移動し, 上空で雲となり, 雨・雪として地表へ降り, 再び蒸発し, というように循環している. 近年, 水不足, 豪雨災害, 水質汚染, 生態系の破壊など, 水に関わる問題が世界各地で広がり, これらに起因する食糧難や伝染病が蔓延するなど, その影響は開発途上国においてますます拡大しているが, これらの問題の背景には, 急激な人口増加による水需要の増大や都市開発, 産業発展などの社会的要因に加えて, 水循環の変動が大きな要因となっている. すなわち, 地球水

循環の変動メカニズムを理解し, 予測精度を向上させることは, 水危機の有力な解決策のひとつである. このような背景の下, 国連関係機関, 国際科学計画, 地球観測衛星委員会らにより構成される統合地球観測戦略 (IGOS : Integrated Global Observing Strategy) では, 水循環テーマの第一ステップとして統合地球水循環強化観測期間 (CEOP : Coordinated Enhanced Observing Period) プロジェクトを開始し, 2002年10月から地球水循環統合化データセットの作成を行っている. このデータセットは, 地上観測データ, 衛星観測データ, 数値気象予報モデル出力データにより構成されるが, これらは, 次元・時空間分解能・座標系・精度・フォ

フォーマットの異なるファイルであり、これらのデータを統合的に利用可能とし、水循環変動の理解・予測へ役立てるシステムを構築することが求められている。

本稿では、CEOP プロジェクトにより作成される地上観測データ・衛星観測データ・数値気象予報モデル出力データをデータベース化し、アーカイブするシステムにおいて、これら地球水循環データを相互に利用し、解析を行うことを可能とするための機能、およびそれらの機能のユーザインタフェースへの実装について述べる。まず、アーカイブの対象とする地球水循環データについて説明を行う。その後、水循環関連データを相互利用し解析を行うために必要とされる機能を示し、それらの機能を実装したユーザインタフェースを紹介する。

2. アーカイブ対象データ

CEOP プロジェクトにおいて作成され、アーカイブ対象となる地球水循環データについて説明を行う。

2.1. 地上観測データ

地上観測データは、世界 36 ヶ所のリファレンスサイトにおいて観測された時系列データである。地上観測データは、リファレンスサイトに設置された各種センサにより観測された物理量の時系列データであり、地表面データ、土壌データ、上層大気データに分類される。地表面データは、地表面付近において観測された気温、気圧、湿度、降水量、放射量など、土壌データは、地下数 cm～数 m における温度や湿度など、上層大気データは、ラジオゾンデなどにより観測される上空の気温、気圧、湿度などの時系列データである。CEOP プロジェクトにおいては観測対象となる物理量は定められているが、各リファレンスサイトにおいてすべての物理量が観測されるわけではなく、観測される物理量の種類、センサの地表からの高度等はリファレンスサイトごとに異なる。また、1 つの物理量を複数の地点あるいは高度において測定を行うリファレンスサイトもある。

地上観測データは、各リファレンスサイトにおいてクオリティチェックがなされ、各値に対してクオリティフラグが付与されている。また、観測時間間隔はリファレンスサイトごとに異なるが、地上観測データ配布担当機関である UCAR (University Corporation for Atmospheric Research) により、1 時間ごとのデータに変換され、CEOP 統一フォーマットにより提供される。データ量は年間数百 MB 程度である。

2.2. 数値気象予報モデル出力

数値気象予報モデル出力は、日本の気象庁を含む世界 11 の気象予報機関による、全球予報モデルあるいは同化システムからの出力データである。数値気象予報

モデル出力は、グリッドデータおよび MOLTS (Model Output Location Time Series) の 2 つに分類される。

グリッドデータは、全球の 2 次元あるいは 3 次元のデータであり、各格子は、気温、気圧、湿度などの予報値を持つ。CEOP プロジェクトにおいては、出力されるべき予報値が定められているが、各機関におけるモデルの違いにより、必ずしもすべての予報値が出力されず、また、同じ物理量を表す予報値であっても、瞬時値であるか平均値であるかなどの違いがある。CEOP プロジェクトによって定められた予報値に加え、それら以外の他の予報値を提供している機関もある。予報機関ごとに使用する数値モデルや予報システムの運用方法が異なるため、出力される 3 次元格子の座標系、時空間分解能は、予報機関ごとに異なる。例えば、水平方向では、緯度、経度を等間隔に分割した正方向格子の他に、緯度、経度の間隔が不等なガウス格子と呼ばれる座標系が用いられ、鉛直方向では、等気圧面や、モデルでのレイヤそのものなどが用いられる。また、数値気象予報モデル出力においては、シミュレーションの初期値としてどの時刻の観測値を用いるか、すなわちシミュレーションの開始時刻をいつにするかにより、同じ時刻に対する数値気象予報モデル出力でも複数のデータが存在する。グリッドデータは、各予報機関において Grib と呼ばれるフォーマットで、数値気象予報モデル配布担当機関である MPI (Max Plank Institute) に転送され、提供される。データ量は年間数十 TB である。

MOLTS は、36 ヶ所のリファレンスサイトを含む世界 41 地点に対する数値気象予報モデルの出力である。リファレンスサイトにおける現地観測データとのより詳細な比較目的として作成され、グリッドデータよりも時間分解能は高い。グリッドデータ同様、予報機関ごとに提供される予報値の数、種類、鉛直方向の座標系、分解能が異なる。また、同時刻の予報値も複数存在する。MOLTS は、そのフォーマットは統一されておらず、予報機関ごとにデータフォーマットが異なる。

2.3. 衛星観測データ

衛星 TRMM, DMSP, NOAA, GMS, Terra, Aqua など、CEOP プロジェクトの観測期間に運用されている衛星による地表面観測データである。これらの衛星データは NASA, ESA, 宇宙航空研究開発機構などの衛星データ提供機関により幾何補正が行われ、世界 36 ヶ所のリファレンスサイトを中心とする 250km 四方の領域、5 ヶ所のモンスーン領域、全球領域の 3 種類の領域の等緯度経度のデータに再配列されて提供される。空間分解能は、衛星に搭載されているセンサが有する分解能とほぼ等しくなるように定められているため、衛星、センサごとに提供されるデータの分解能は異な

る。また、提供される物理量データの種類は、衛星に搭載されているセンサにより異なるが、輝度温度、アルベドなどの低次物理量のみならず、海面温度、土壌水分量等の高次物理量データも提供される。衛星データの大部分は地表面上の物理量の2次元データであるが、センサによっては鉛直方向の物理量の分布を測定、生成が可能なものがあり、この場合は3次元データとなる。データフォーマットは衛星データ提供機関ごとに異なる。データ量は年間数百TBと想定されている。

3. 地球水循環データアーカイブシステム

本節では、現在構築を行っている地球水循環データアーカイブシステムの構成について述べる。

3.1. システム構成

図1に地球水循環データアーカイブシステムの構成を示す。システムはクライアント-サーバモデルにより実現しており、クライアント-サーバ間の通信はHTTPに基づくSOAPを採用し、ウェブサービスとして実装している。地球水循環データの利用者は世界各地に散在し、その計算機・ネットワーク環境は様々である。ファイアウォールを越えてアクセスする必要がある場合も多いが、HTTPを採用することで、より多くの利用者への対応が容易となる。また、ウェブサービスとして提供することにより、利用者が自ら作成したアプリケーションで直接水循環関連データへアクセスすることも可能となる。一方、地球水循環関連データはサイズが極めて大きく、1つ1つの物理量を忠実にXML化して通信を行うことは極めて非効率的である。このため、データの転送に関しては、データ列の各値をIEEE754で表現して圧縮することによりバイナリ配列化して通信を行い、転送時間の短縮を図っている。

3.2. データサーバ

サーバでは、システムへのログイン、指定した条件に基づくデータの検索、検索されたデータの時間軸、空間軸の変換、検索されたデータに対する演算や統計量の計算などのメソッドを提供する。

サーバ側においては、1地点における物理量であるためデータ量が比較的小さい現地観測データ、および数値気象予報モデル出力におけるMOLTSデータは、メタデータとともにデータそのものをDBMSによって管理するが、2次元、3次元的な広がりを持つ数値気象予報モデル出力におけるグリッドデータ、および衛星観測データは、データ自体をディスクアレイ、あるいはテープライブラリシステムを利用した階層ファイルシステム上のファイルとして格納し、DBMSにおいてはそのインデックスとメタデータのみを管理する。数値気象予報モデル出力のグリッドデータ、衛星観測

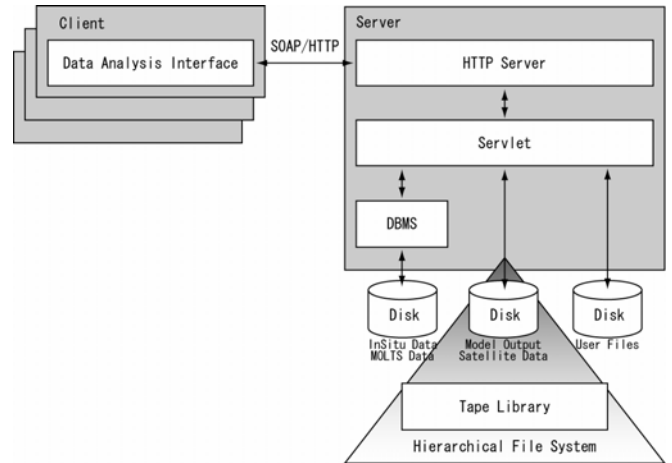


図1 システム構成

データはそのサイズが大きく、すべてのデータをディスク上に格納するのは困難であるため、階層ファイルシステム上のファイルとする必要があるが、DBMSによって提供されているBLOBは階層ファイルシステム上のファイルには必ずしも最適化されておらず、十分な性能が発揮されないことが考えられるためである。

利用者により検索のメソッドが実行されると、データサーバはDBMS、あるいは外部ファイルから検索条件に合うデータを切り出し、内部表現形式に変換し、サーバ内の利用者データ領域に保存する。利用者は、この利用者データ領域内の切り出されたデータに対して、各軸の変換や演算、統計量の計算等のメソッドを実行する。メソッドの実行結果として生成されるデータもまた、新しいファイルとしてデータサーバ内の利用者データ領域に保存される。この新たに作成されたデータも内部表現形式によって表現されており、利用者は検索から直接作成されたデータと検索結果データに対して処理を加えて作成されたデータを区別することができなく利用することができる。

3.3. データ検索・解析インタフェース(クライアント)

データ検索・解析インタフェースは、データサーバのウェブサービスに対するクライアントアプリケーションとして利用者側の計算機上で動作し、利用者の要求に応じて、データサーバのメソッドを実行する。データ検索・解析インタフェースはJAVAにより記述されている。利用者は世界各地に散在しており、そのコンピュータ環境は様々であるが、データ検索・解析インタフェースは、JAVAの実行環境があり、ネットワークが接続されていれば、他の特別なハードウェア、ソフトウェアは必要としない。

データの解析処理は、平均値や最大・最小値の計算などの新たなデータを作成する処理に関しては、データ検索・解析インタフェースがデータサーバのメソッドを実行し、処理自体はデータサーバで行われる。一

No.	Label	Dimension	Data	Location	Period	Creation Time
1			0In-Situ/Meteorolo	Reference Site/N	2001/09/05 00:00	Tue Mar 01 16:13...
2			0In-Situ/Surface	Reference Site/N	2001/09/05 00:00	Tue Mar 01 16:14...
3			2Satellite/GMS-S-V	Reference Site/N	2001/09/05 00:00	Tue Mar 01 16:14...
4			3Satellite/TRMM P	Reference Site/N	2001/09/05 00:00	Tue Mar 01 16:15...
5			2Model Output/EC	Global	2001/09/05 00:00	Tue Mar 01 16:17...
6			2Model Output/EC	Global	2001/09/05 00:00	Tue Mar 01 16:19...
7			2Model Output/EC	Global	2001/09/05 00:00	Tue Mar 01 16:21...
8			2Model Output/EC	Global	2001/09/05 00:00	Tue Mar 01 16:22...
9			0MOLTS/JMA/BH	Reference Site/LI	2003/03/01 00:00	Wed Mar 30 14:53...
10			0MOLTS/UKMO/J3	Reference Site/E	2002/10/01 00:00	Fri Apr 01 14:35.2
11			3Model Output/NC	Global	2002/10/01 00:00	Wed Apr 06 10:42...
12			2Model Output/NC	Global	2002/10/01 00:00	Fri Apr 08 17:30.3
13			2Model Output/CP	Global	2001/07/01 00:00	Fri Apr 08 17:40.2
14			2Masked/Satellite	Reference Site/N	2001/09/05 00:00	Wed Apr 20 16:02...
15			0Processed/Zonal	Reference Site/N	2001/09/05 00:00	Wed Apr 20 16:04...
16			0Processed/Diurn	Reference Site/C	2002/10/14 00:00	Mon May 02 14:00...
17			1In-Situ/Meteorolo	Reference Site/E	2002/10/01 00:00	Thu May 19 14:08...
18			0MOLTS/JMA/BH	Reference Site/E	2002/10/01 00:00	Thu May 19 14:09...
19			0Processed/Diurn	Reference Site/E	2002/10/01 00:00	Thu May 19 14:14...
20			1Processed/Diurn	Reference Site/E	2002/10/01 00:00	Thu May 19 14:14...

図 2 データ管理ウィンドウ

方、データの表示は、データ検索・解析インタフェースがデータ転送メソッドを実行し、データサーバからデータ検索・解析インタフェースへ物理量データが転送され、データ検索・解析インタフェース上にてグラフ、画像化が行われる。データサーバ上でグラフ化、画像化を行い、その画像をクライアントへ転送する方法も考えられるが、この場合にはグラフ化、画像化パラメータを変更するたびに画像を転送する必要が生じること、グラフ、画像上の指定された点の物理量の値の表示要求に対応するためには、画像データに加えて、結局物理量自体の値もクライアント側で必要となることによる。同様に、相関係数、回帰係数を表示する処理も、データ検索・解析インタフェースがデータサーバにデータを要求し、データ検索・解析インタフェース側で計算を行い表示する。

図 2 はデータ検索・解析インタフェースにおいて、データサーバの利用者領域上のデータを管理するためのウィンドウである。検索、処理されたデータの簡単な情報が表形式により表示されている。このウィンドウから、より詳細な情報の表示や、不要なデータの削除を行うことが可能である。また、統計量の計算や演算、表示の対象とするデータの選択も、このウィンドウを通じて行う。操作の対象とする 1 つ以上のデータを表から選択し、上部のメニューバーから操作を選択することにより、必要なメソッドがサーバ上で実行され、処理が行われる。

4. 異種データ相互利用とその実装

4.1. 統一的データ表現

地球水循環データは、データごとに次元が異なるが、地球水循環データアーカイブシステム内部においては、検索され切り出されたデータをすべて、時間、鉛直方向、南北方向、東西方向の 4 つの軸を有する 4 次元配

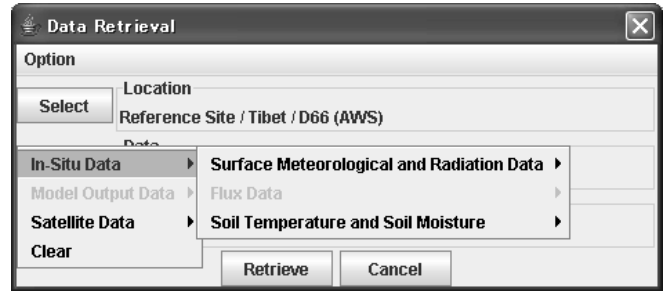


図 3 データ検索条件入力ウィンドウ

列として表現し、この 4 次元配列にデータの情報、および 4 つの各軸の目盛値、および各軸の情報を付加して表現する。例えばリファレンスサイトにおける地上気温の時系列データは、時間軸方向のみ n のサイズを持ち、鉛直軸、南北軸、東西軸のサイズが 1 である、 $n \times 1 \times 1 \times 1$ の配列となる。配列の各要素には、物理量を表す数値に加え、そのデータのクオリティを表すフラグが付与される。データの情報には、データの種類、単位、データが検索された際に指定された条件、データ作成時刻等により構成され、また、軸情報は、軸の種類や目盛値の単位等により構成される。各軸の目盛値は、範囲の指定も可能としている。これは、その値が示す物理量が、ある範囲の総和や平均値といった明確な 1 点ではなく、ある範囲における値である場合にも対応するためである。この表現法はアーカイブ対象となる地球水循環データのフォーマットをほぼ包含しており、この表現法をシステム内部で用いることにより、データの種別を区別することなく、統一的に検索されたデータを扱うことが可能となる。このデータ表現は利用者には隠蔽されており、利用者はデータ表現については意識をする必要はない。

4.2. データ検索

CEOP プロジェクトにおける地球水循環データは、データの種別（現地観測データ、衛星観測データ、数値気象予報モデル出力、さらにはその中の気温、気圧、土壌水分量など）、時刻、位置（全球、5 ヶ所のモンスーン領域、36 ヶ所のリファレンスサイト）の 3 つのパラメータによって決定される。地球水循環データアーカイブシステムでは、これらの 3 つのパラメータをメニュー形式で指定することにより検索を行う。利用者はこれら 3 つのパラメータを指定することにより、すべての地球水循環データを同様の手順で検索することが可能である。データの種別によってデータが存在する期間や位置が異なるが、データが存在するパラメータの組み合わせをデータサーバにおいてデータベース化しており、データ検索時には利用者がデータ検索・解析インタフェースにおいてあるパラメータを指定すると、指定されたパラメータのデータの存在に関する

問い合わせメソッドを実行して、データの存在するパラメータの組み合わせを取得する。この情報に基づき、他のパラメータの選択メニュー内のデータが存在しない項目の文字の輝度を下げ、データが存在しないパラメータの組み合わせが利用者に分かるようにしている。

4.3. フォーマット変換

地球水循環データは、データごとにフォーマットが異なる。現地観測データでは、すべてのリファレンスサイトのデータで統一のフォーマットが採用されているものの、CEOP独自形式であり、また、MOLTSデータではデータ提供数値予報機関ごとに異なる形式のフォーマットが採用されている。利用者が既存のアプリケーションで使用する場合には、そのアプリケーションが対応しているフォーマットに変換する必要があり、また、利用者が自ら作成したプログラムで利用する場合にも、利用しやすいフォーマットに統一されている方が好ましい。

地球水循環データアーカイブシステムを用いて解析を行う場合には、データサーバがそれぞれのフォーマットを自動的に判断して読み込むため、利用者はデータフォーマットの差異を考慮する必要はない。一方、利用者が自身のプログラムや地球水循環データアーカイブシステム以外の解析ツールで地球水循環データを利用するために、地球水循環データアーカイブシステムでは、ASCIIテキストファイルもしくは気象学、海洋学などで一般的に使用されているNetCDFフォーマットに変換してデータ検索・解析インタフェースが動作するクライアント側のマシンに保存する機能を提供する。データ検索・解析インタフェースは、データの転送メソッドを実行し、データサーバより統一されたデータ表現形式の保存対象データを受け取り、データ検索・解析インタフェースが要求されたフォーマットに変換して、クライアント側マシンに保存する。

4.4. 単位の変換

地球水循環データは、データの種類やデータ提供機関ごとに、同じ物理量が異なる単位で与えられている場合がある。これら単位の異なるデータを直接比較するためには、統一された単位に変換する必要があるが、地球水循環データアーカイブシステムでは、地球水循環データを扱う上でよく用いられる単位の変換は、データ検索・解析インタフェースでメニューを選択することのみで実行できる。また、メニューに登録されていない単位の変換に関しても、利用者が変換式を入力することができる。データ検索・解析インタフェースが単位変換メソッドを発行することにより、データサーバは対象データの単位の変換を行い、利用者領域に新たなデータとしてファイルを作成する。

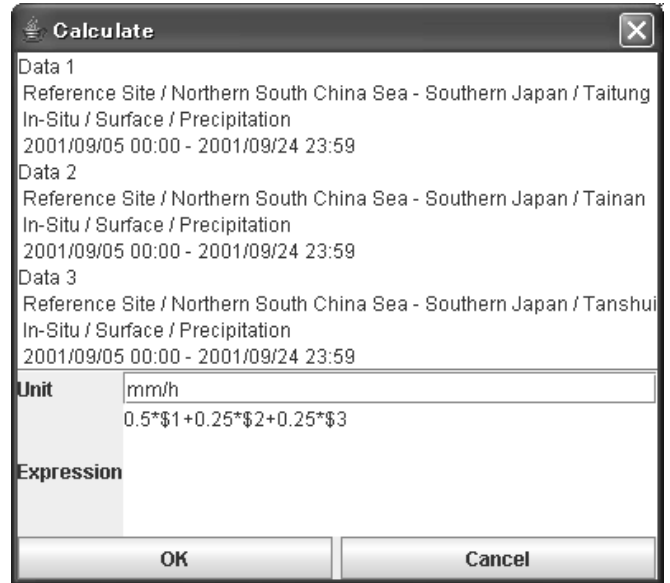


図4 演算式入力ウインドウ

4.5. 時空間軸変換・整合

地球水循環データは、データごとに時空間分解能や座標系が異なり、そのままでは相互に比較することはできない。このために、時空間の分解能や座標系を変換し、データ間の時空間軸を整合させる機能が必要とされる。地球水循環データアーカイブシステムにおけるデータサーバは、最近隣値、線形補間値、最大値、最小値、平均値によって再サンプリングする方法により、利用者が定義する新たな時間軸に既存データを整合させるメソッドを提供する。また、既存のあるデータの時空間軸に他のデータを整合させるメソッドも提供し、この場合には、最近隣法、線形補間法が適用可能である。これらの時空間軸変換・整合処理時には、データが瞬時値のようにある1点の物理値ではなく、ある範囲における集約値である場合も自動的に考慮される。例えば、3時間毎の平均降水量に関して、0時の値として0～3時までの降水量の平均値が記載されているデータの場合、これを1時間毎のデータに変換する際には、単純に最近隣法を適用すると2時のデータは3時として記載されている値となるが、実際には0時に記載されている値をとらねばならない。地球水循環データアーカイブでは、変換元データのメタデータに従い、自動的に適切な値を選択する。データ検索・解析インタフェースより、時空間軸変換・整合メソッドが発行されると、データサーバは利用者領域に新たなデータとして、時空間軸変換・整合を行ったデータを作成する。

4.6. 演算・解析機能

データ解析を進めるためには様々な演算が必要とされる。地球水循環データアーカイブシステムでは、複数のデータの相関係数、回帰係数を、散布図の表示

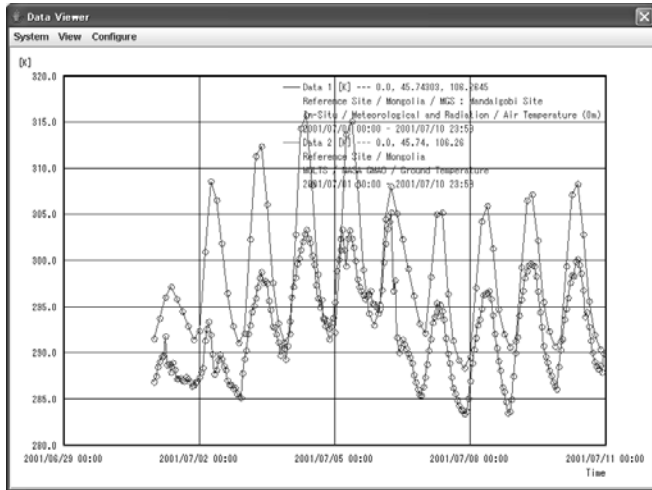


図 5 2次元グラフ表示

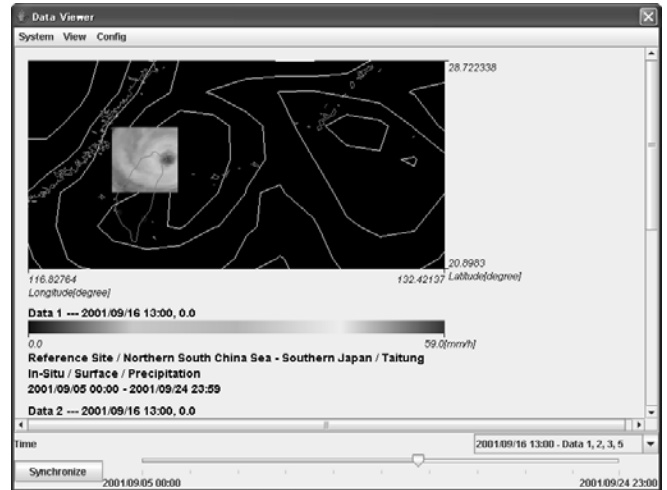


図 8 重ね合わせ表示

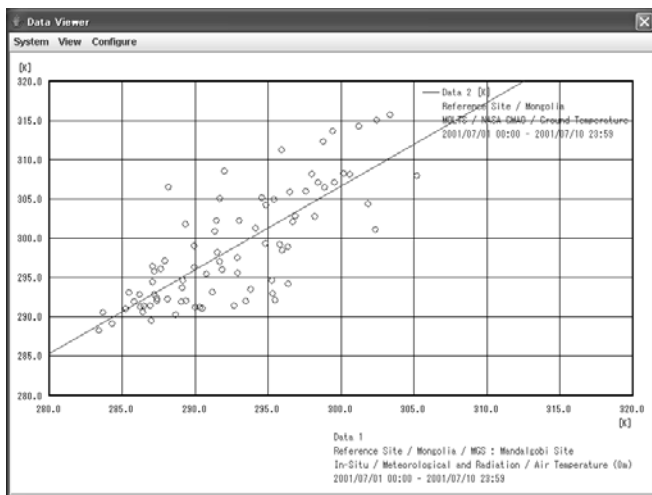


図 6 散布図表示

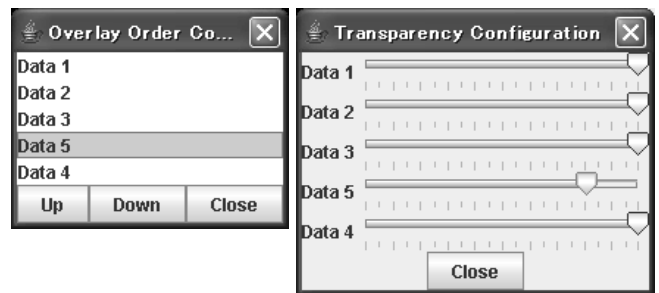


図 9 重ね合わせ順序 (左) と透明度 (右) の指定

較対象となる2つの物理量の差分や比率の計算など、様々な演算が可能である。演算メソッドが実行されると、データサーバは指定された演算を実行し、演算結果は利用者領域に新たなデータとして作成される。図4の例では、3地点における降水量の加重平均を求めている。

4.7. データ表示機能

地球観測データの解析を行うためには、様々な種類のデータを目視して比較ができるよう、データの可視化機能が必要である。データ検索・解析インタフェースがデータ転送メソッドを実行しデータサーバより物理量データを取得し、データ検索・解析インタフェースがグラフ画像、ビットマップ画像を作成して、表示を行う。

4.7.1. 複数データのグラフ表示

データ検索・解析インタフェースは、データを比較するために、複数のデータを1つのグラフ上に表示する機能を提供している。グラフのx軸、y軸には、時間軸、鉛直軸、南北軸、東西軸、物理量の値を任意に指定可能である。データ比較を容易にするため、注目領域を拡大する機能が提供しており、また、表示対象としたデータが異なる物理量、時空間次元を有する場合でも、表示可能である。この場合、物理量に割り当てられた軸の目盛りの表示は、グラフのメニューから

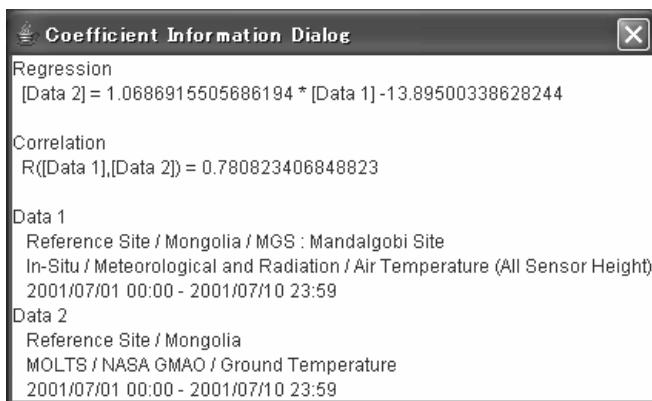


図 7 回帰式・相関係数表示

ウィンドウのメニューの係数表示コマンドを選択することにより表示することができる、また、データ管理ウィンドウにおいて複数のデータを選択してメニューから演算操作を選択することで、図4のようなウィンドウが表示され、利用者が四則演算や算術関数によって構成される任意の数式を入力することにより、直接提供されていない物理量の他の物理量からの導出、比

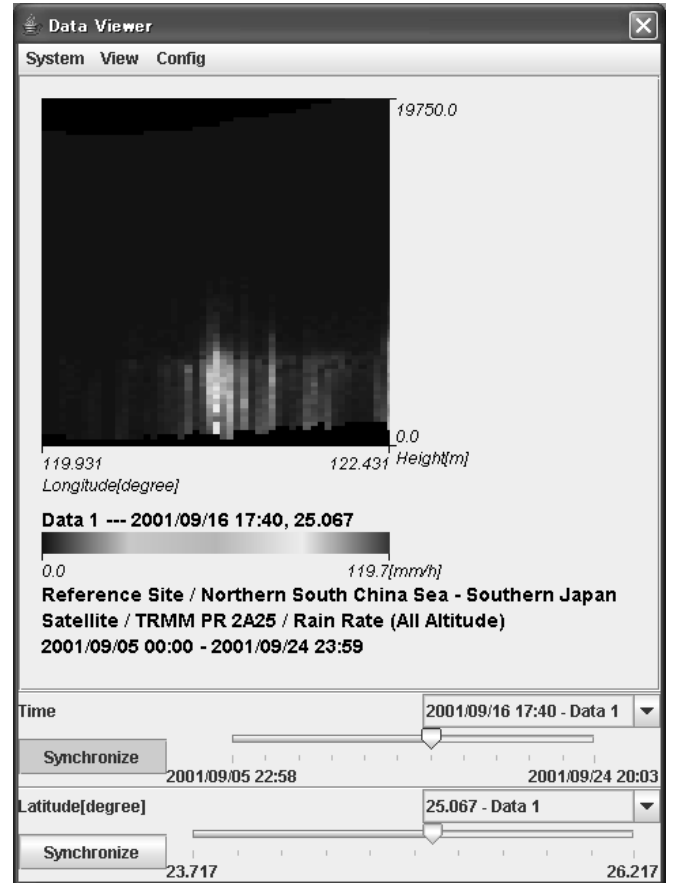


図 10 複数ウィンドウの同期

変更することができる。x 軸, y 軸以外の軸は, 画面内にスライダーが表示され, このスライダーを操作することにより表示対象値を変更できる。図 5 は地上観測における気温とモデルによる気温の 2 つのデータを同時に 1 つのグラフ上に表示した例である。

4.7.2. 散布図表示

複数データの関連を表示するために, データ検索・解析インタフェースには, 散布図を作成し, 表示する機能が実装されている。散布図は複数のデータの対応をとり, それぞれの物理値を軸として作成されるが, 異なる時間, 高度, 位置のデータを比較する場合もあり, 4 つの軸に関して, いずれをデータ対応の対象とするかの選択ができる。例えば, 時間, 位置のみをデータ対応の対象とした場合, 高度はデータ対応の対象とはせず, 異なる高度であっても対応しているとみなされるため, 異なる高さのデータの分布状況を見ることができる。また, 散布図には回帰直線も同時に表示され, さらにこのウィンドウのメニューから相関係数・回帰係数の表示を選択することで, 図 7 のようにその数値が表示される。

4.7.3. 画像の重ね合わせ

データ検索・解析インタフェースには, データを 2 次元ビットマップ画像として表示する機能が実装され

ている。複数のデータを選択して, それらを重ね合わせて表示することが可能である。図 8 は数値気象予報モデル出力のグリッドデータと衛星観測データを重ね合わせて表示している。数値気象予報もモデル出力データは等位線表示を指定している。グラフの場合と同様, x 軸, y 軸にはデータの任意の軸を指定することが可能であり, 図 8 の例では, x 軸に東西軸 (経度), y 軸に南北軸 (緯度) を指定している。また, 図 8 で画像化の対象としたデータは時系列データであり, 時間軸のサイズが 2 以上であるが, このようにグラフ化, 画像化の際に x 軸, y 軸として指定しない軸でそのサイズが 2 以上である場合には, ウィンドウ上にスライダーが表示され, 任意にそれらの軸の値を変更することができる。

画像化パラメータとして, カラーリング (カラーチャート, グレイスケール) の選択, 等位線, 海岸線, 地点名の表示の有無等を指定することができる。複数のデータを重ね合わせて表示した場合, 個々のデータに対して画像化パラメータを指定することができ, また, 図 9 のウィンドウで指定することにより, 重ね合わせる画像の順序の変更, およびそれぞれの画像に対する透明度のスライダーで操作することが可能である。

4.7.4. 複数グラフ・画像の同期

グラフ表示, 画像表示においては, 時間軸, 鉛直軸, 東西軸, 南北軸の4つの軸のうち, x軸, y軸に指定された以外の軸のサイズが2以上の場合, スライドバーが表示される. このとき, 複数のグラフ表示ウィンドウ, 画像表示ウィンドウのスライドバーの同期ボタンをセットすることにより, それらの軸を同期させることを可能とした. すなわち, 1つのグラフ表示ウィンドウ, あるいは画像表示ウィンドウのスライドバーを操作することで, 同期が指定されている他のグラフ表示ウィンドウ, 画像表示ウィンドウのスライドバーも同期して, 値が変更される. 同期対象となっている軸の値は, 最近隣法により選択される. 同期ボタンが押されると, データ検索・解析インタフェースは同期ボタンが押されているウィンドウの同期対象軸間の対応テーブルを作成する. スライドバーによりあるウィンドウの値が変更されると, データ検索・解析インタフェースはこの対応テーブルを参照し, 他のウィンドウの軸の値を変更する. 図10は, 衛星による3次元の降雨強度分布データについて, 水平面での分布を表示した画像(右)と鉛直方向の分布を表示した画像(左)について, 時間軸を同期させた例である. 一方の時間軸を変化させることで他方のウィンドウの時間軸も変化し, 同時刻の水平・鉛直分布が表示される.

5. おわりに

本稿では, 現在構築を行っている地球水循環データアーカイブシステムにおいて, 異種データを相互に利用し, 解析を行うための機能について述べた. まず, アーカイブ対象となるデータについてまとめ, その後, 構築中のシステムの構成を説明するとともに, 異種データの相互利用のための機能およびそれらの機能を実装したデータ検索・解析インタフェースを紹介した. 利用者が異種データを相互に利用し, 解析を行う際に利用される基本的な機能の実装はほぼ完了したが, より高度な機能も要求されており, 今後, 実装をさらに進め, 機能の充実を図る予定である.

文 献

- [1] T. Nemoto, E. Ikoma and M. Kitsuregawa, "Design of data server for CEOP data," Proc. 2nd Asia Pacific Association of Hydrology and Water Resources Conference, no.2, pp.558-565, Singapore, July 2004.
- [2] 根本利弘, 小池俊雄, 喜連川優, "地球水循環データアーカイブシステムとユーザインタフェースの構築", 電子情報通信学会技術報告, vol.105, no.173, pp.49-54, July 2005.
- [3] E. Stolte, C. Praun, G. Alonso and T. Gross, "Scientific Data Repositories – Designing for a Moving Target," Proc. ACM SIGMOD International Conference on Management of Data, pp.349-360,

San Diego, California, USA, June 2003.

- [4] T. Koike, "The Coordinated Enhanced Observing Period – an initial step for integrated global water cycle observation," WMO Bulletin, vol.53, no.2, pp.115-121, April 2004.