

ビッグデータ同化



三好 建正

みよし たけまさ

理化学研究所 計算科学研究機構

Takemasa.Miyoshi@riken.jp



別所康太郎(気象衛星センター)、瀬古弘(気象研)、富田浩文(理研)、
佐藤晋介(NICT)、牛尾知雄(大阪大学)、石川裕(東京大学)



独立行政法人
科学技術振興機構
Japan Science and Technology Agency

データ同化

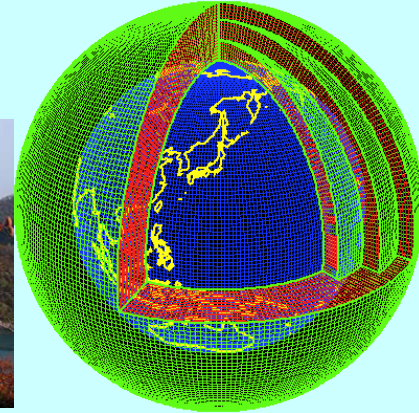
観測・実験データ



データ同化



シミュレーション



データ同化は、シミュレーションと現実世界を結びつけ、相乗効果を生み出す。

カオス同期 Chaos Synchronization

Master (drive) system

Slave (response) system

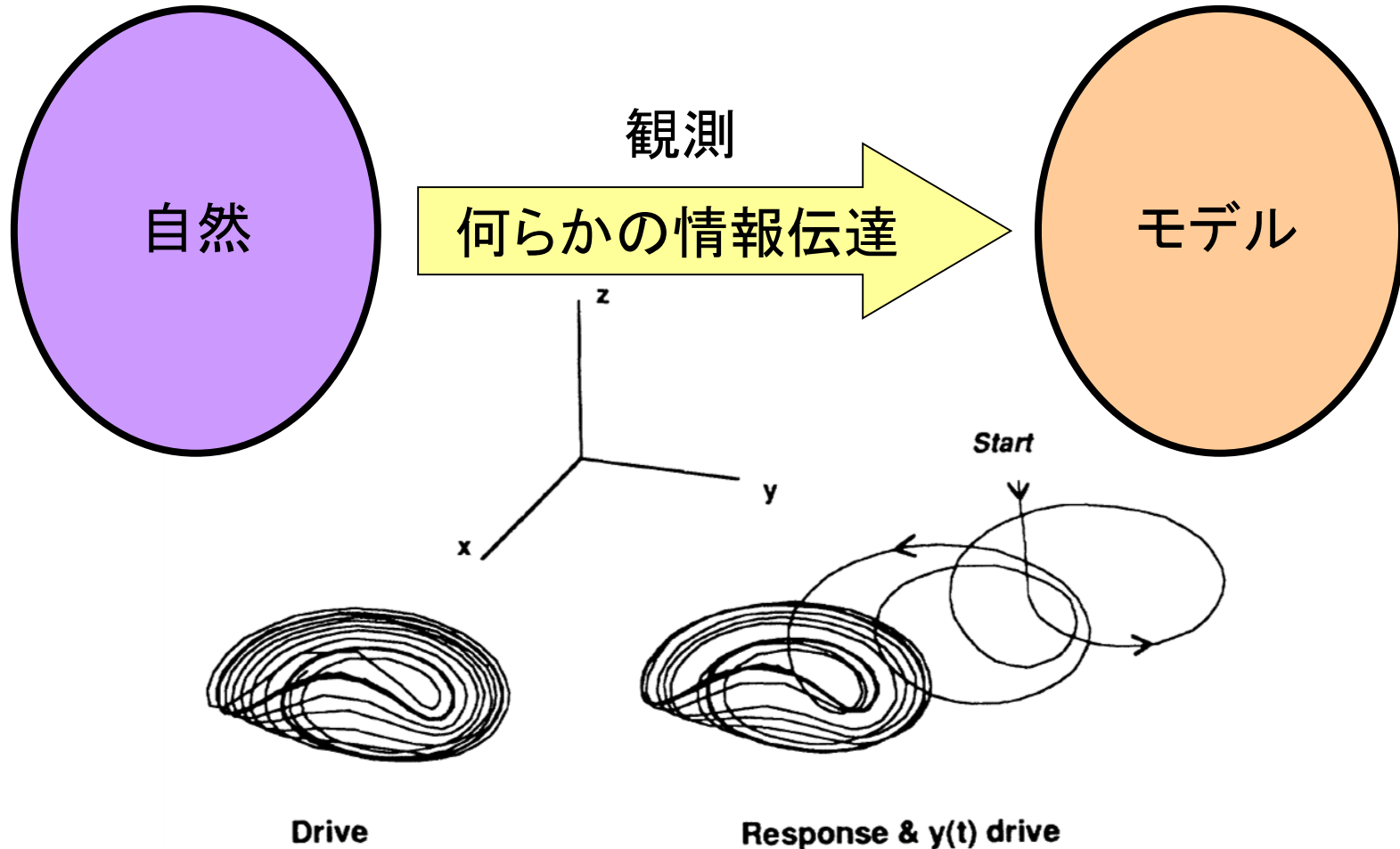


FIG. 1. The attractors for the Rössler drive system and the $(x'-z')$ response system and $y(t)$ drive variable.

カオス同期 Chaos Synchronization

Master (drive) system

Slave (response) system

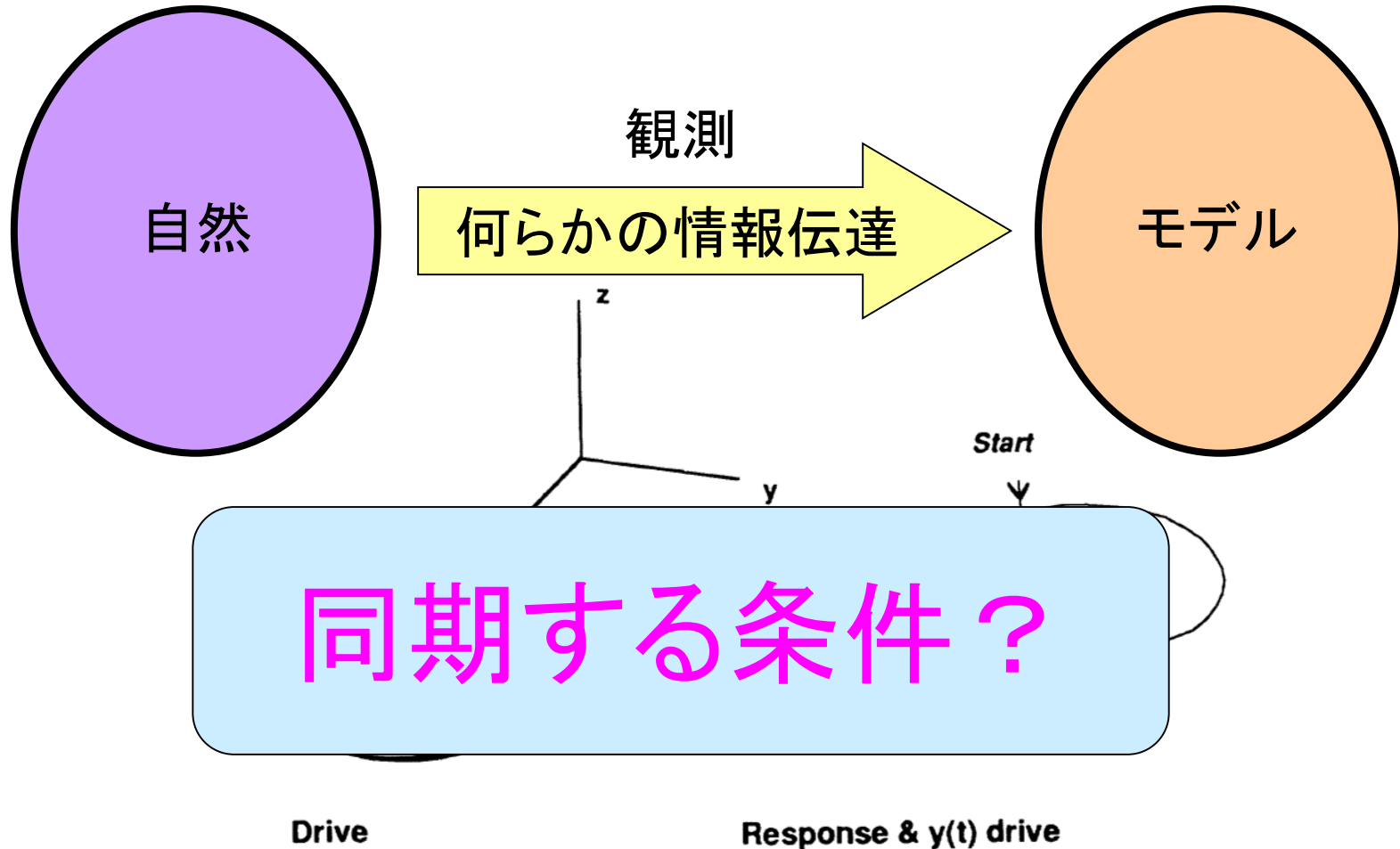
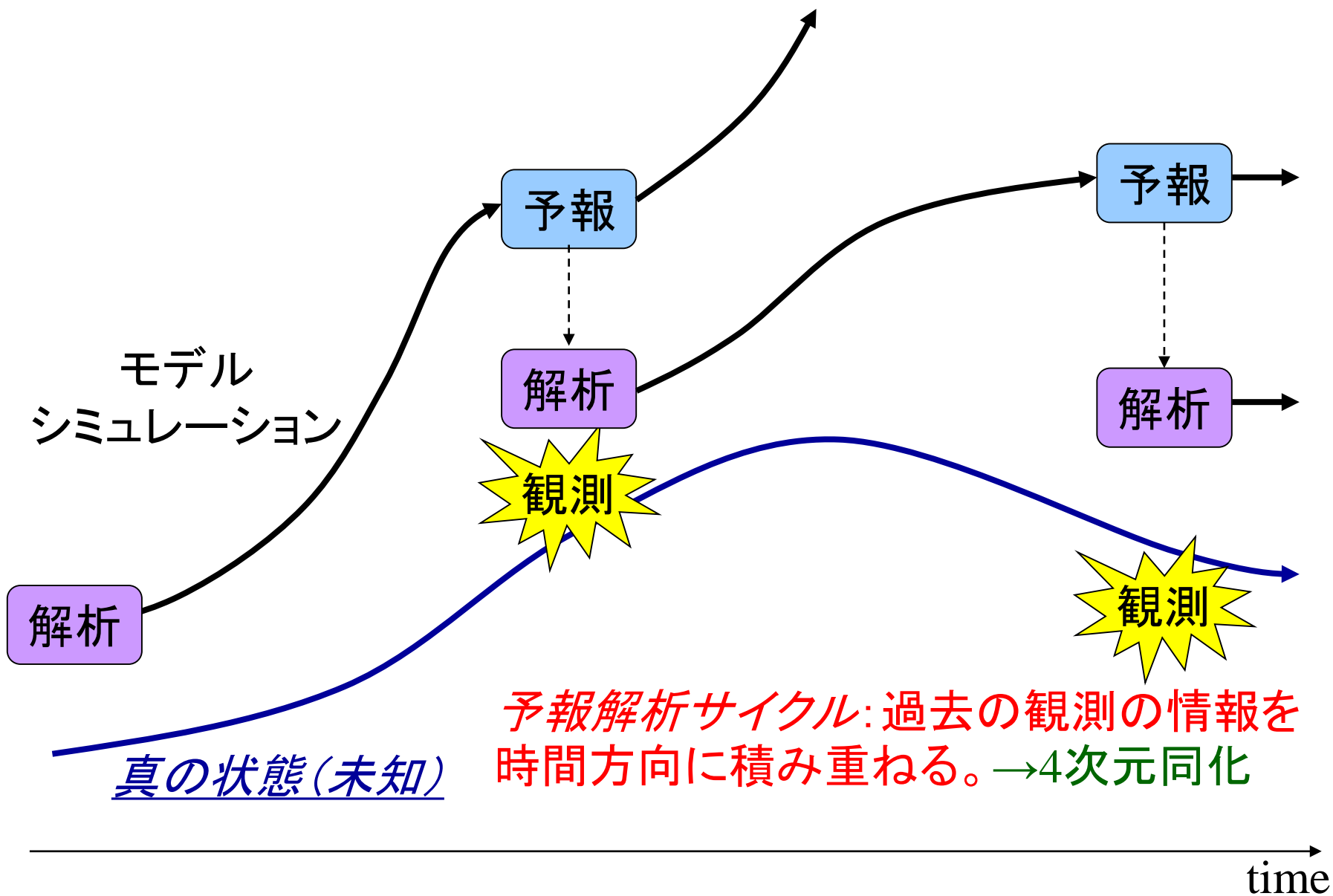


FIG. 1. The attractors for the Rössler drive system and the $(x'-z')$ response system and $y(t)$ drive variable.

データ同化の役割、方法、効果

- **役割**: シミュレーションと現実世界をつなぐ
- **方法**: 統計数理のアプローチ、データドリブン
- **効果**:
 - シミュレーションと現実世界のギャップを低減
 - シミュレーションの誤差評価
 - シミュレーション精度の向上
 - 初期値の改善
 - モデルパラメータの最適化
 - モデル誤差の推定・補正
 - 観測・実験システムの最適化
 - 観測・実験システムの効果を評価
 - 効果的な観測・実験システムの提言

数値天気予報のしくみ



Global Observing System

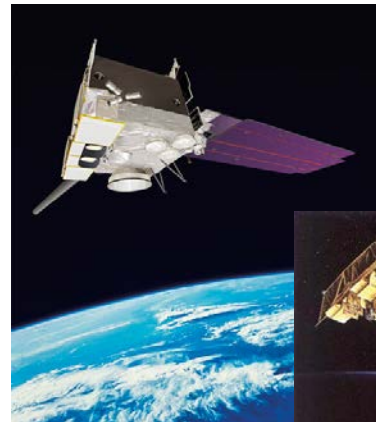
Radar



Aircraft



Satellite



Weather balloon



Ship



Surface station



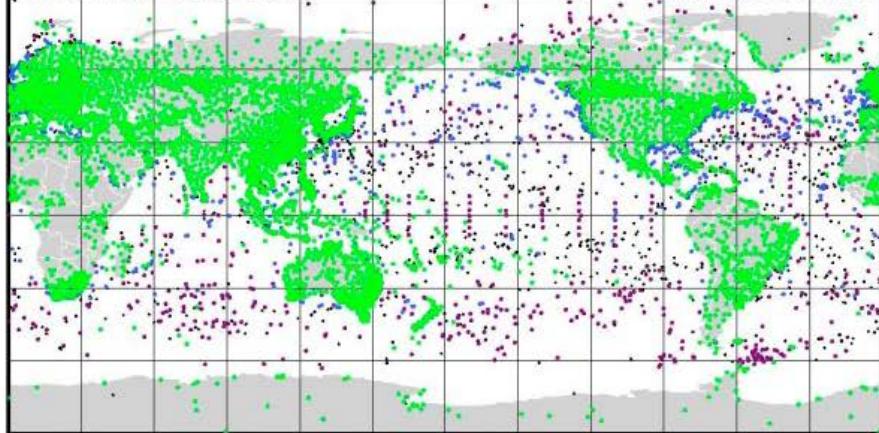
Buoy



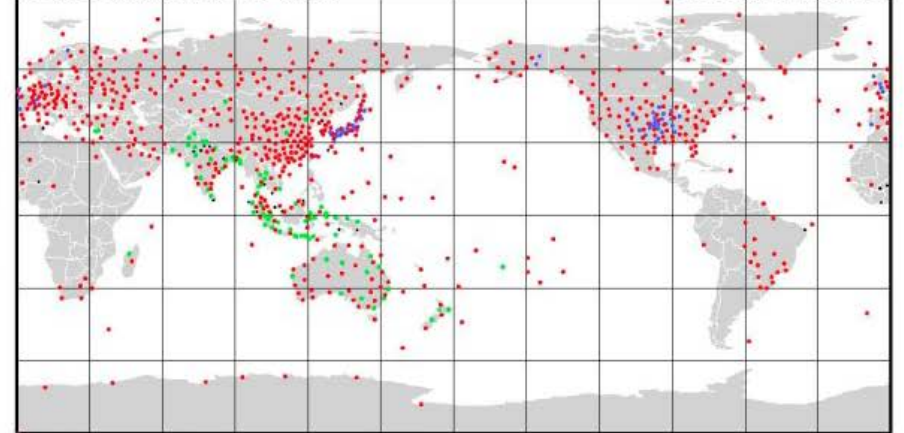
Collecting the data

JMA GLOBAL ANALYSIS – DATA COVERAGE MAP (Da00ps): 2009/04/22 00:00(UTC)

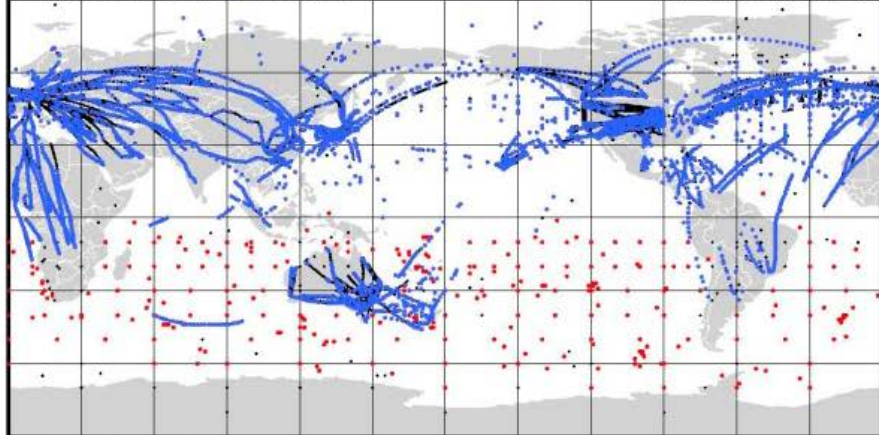
CONVENTIONAL SURF 2009/04/22 00:00(UTC)



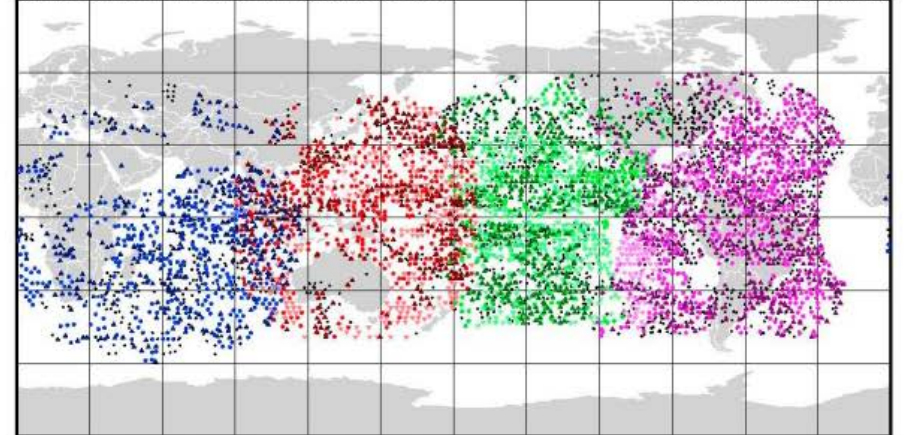
CONVENTIONAL UPPER 2009/04/22 00:00(UTC)



CONVENTIONAL OTHERS 2009/04/22 00:00(UTC)



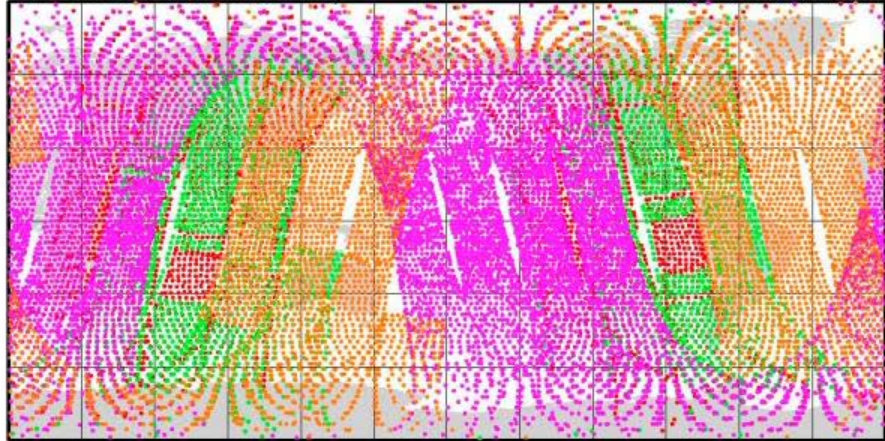
ATMOSPHERIC MOTION VECTOR 2009/04/22 00:00(UTC)



World's effort! (no border in the atmosphere)

Collecting the data

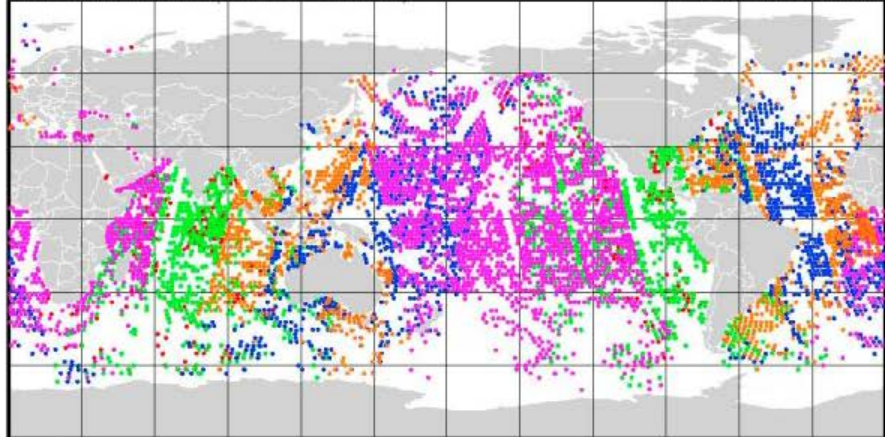
MW-SOUNDER(AMSU-A) 2009/04/22 00:00(UTC)



AMSU-A[●]: 18163 [●]NO_USE

NOAA-15 NOAA-16 NOAA-18 METOP-2

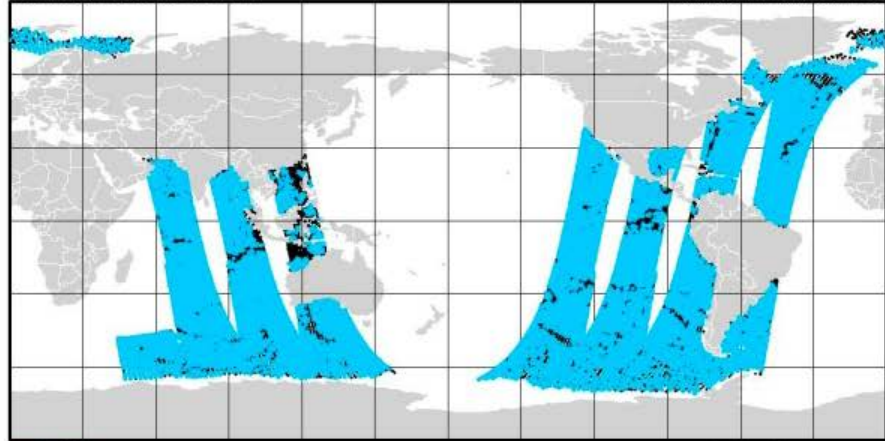
MW-SOUNDER(AMSU-B/MHS) 2009/04/22 00:00(UTC)



AMSU-B[●]: 4487 MHS[●]: 3452 [●]NO_USE

NOAA-15 NOAA-16 NOAA-17 NOAA-18 METOP-2

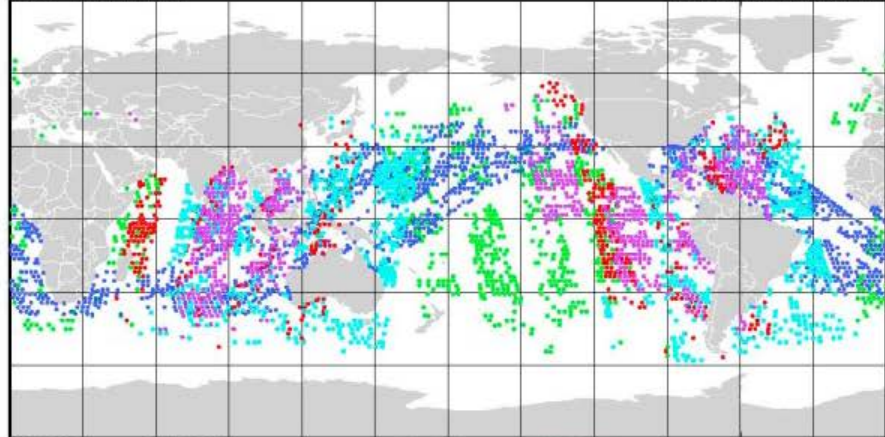
SCATTEROMETER 2009/04/22 00:00(UTC)



SCAT 12714

[●]NO_USE

MW-IMAGER 2009/04/22 00:00(UTC)



SSM/I 727 SSMIS 1719 TMI 1455 AMSR-E 810
DMSP13 DMSP16 DMSP17

1つ1つのデータを大切に！

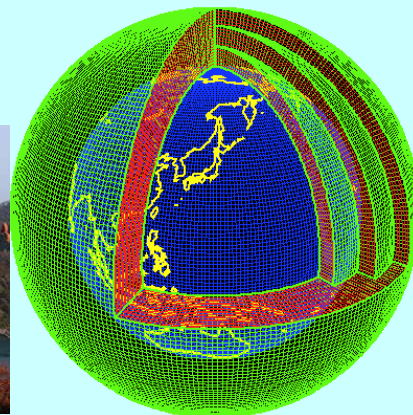
データ同化

観測データ



天気予報シミュレーション

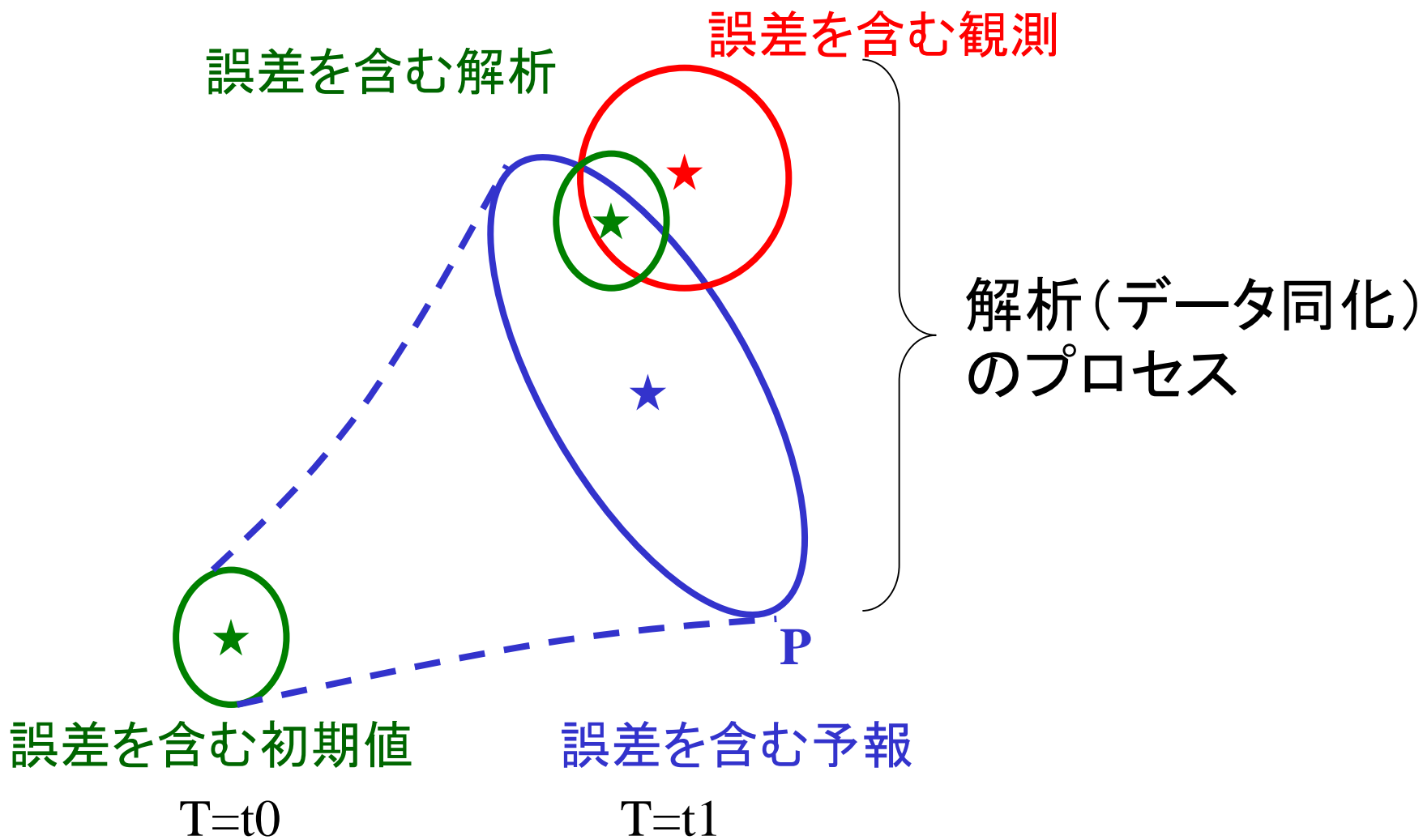
データ同化



データ同化は、シミュレーションと現実世界を結びつけ、相乗効果を生み出す。

双方の情報を最大限に抽出

確率論的な表現



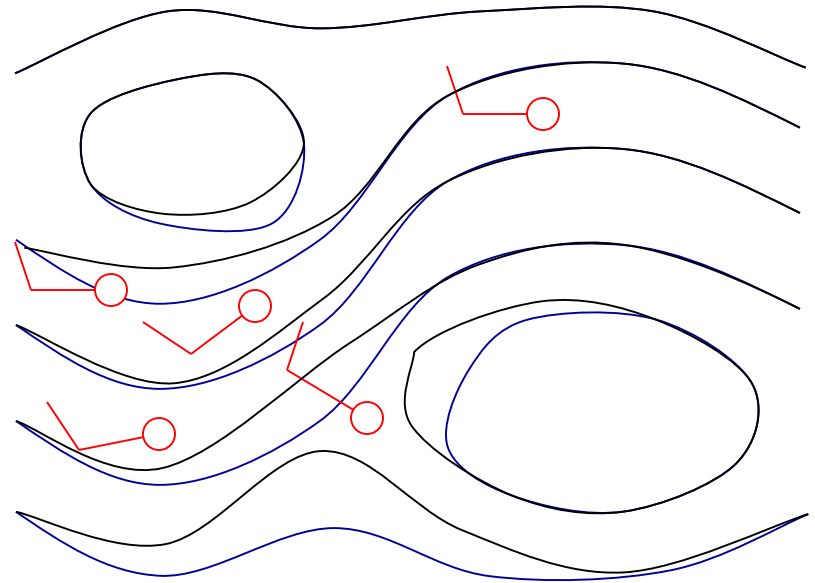
解析(データ同化)のプロセス

背景場を観測で修正する

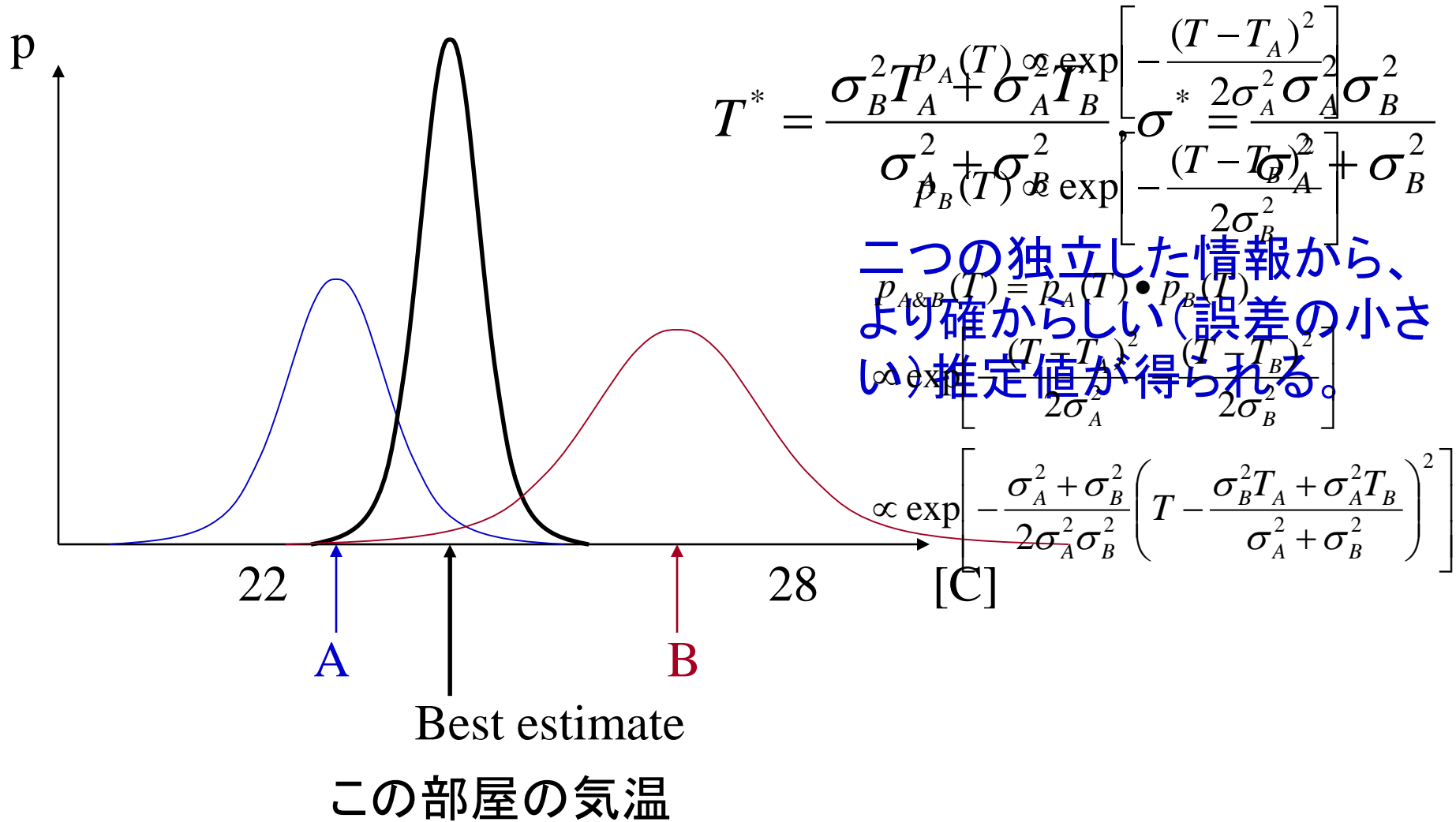
背景場も観測も、真の状態
(未知)のまわりに分布

↓
真の状態について、独立な
情報を提供する。

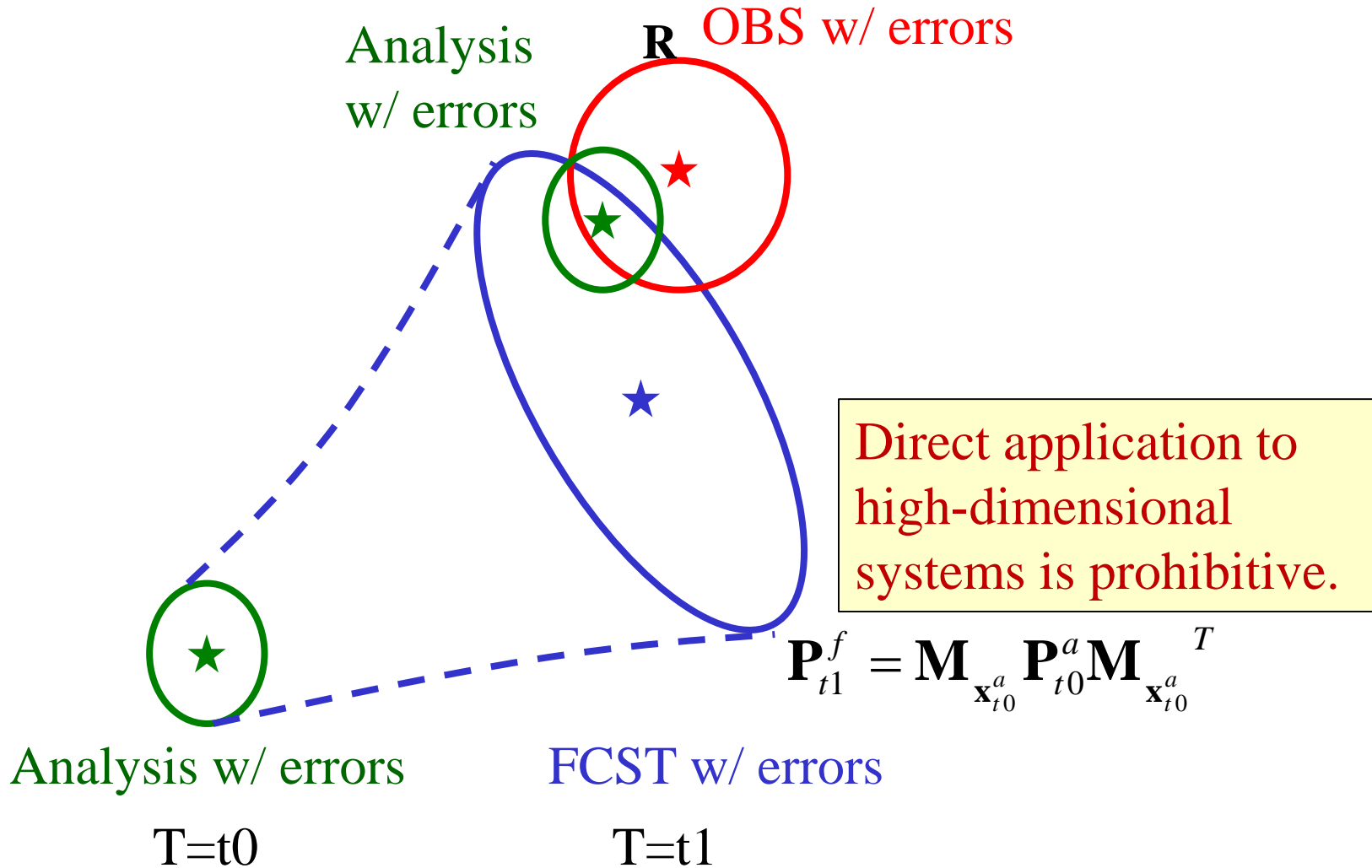
↓
二つの情報を組み合わせることで、より確からしい
解析値を得る。



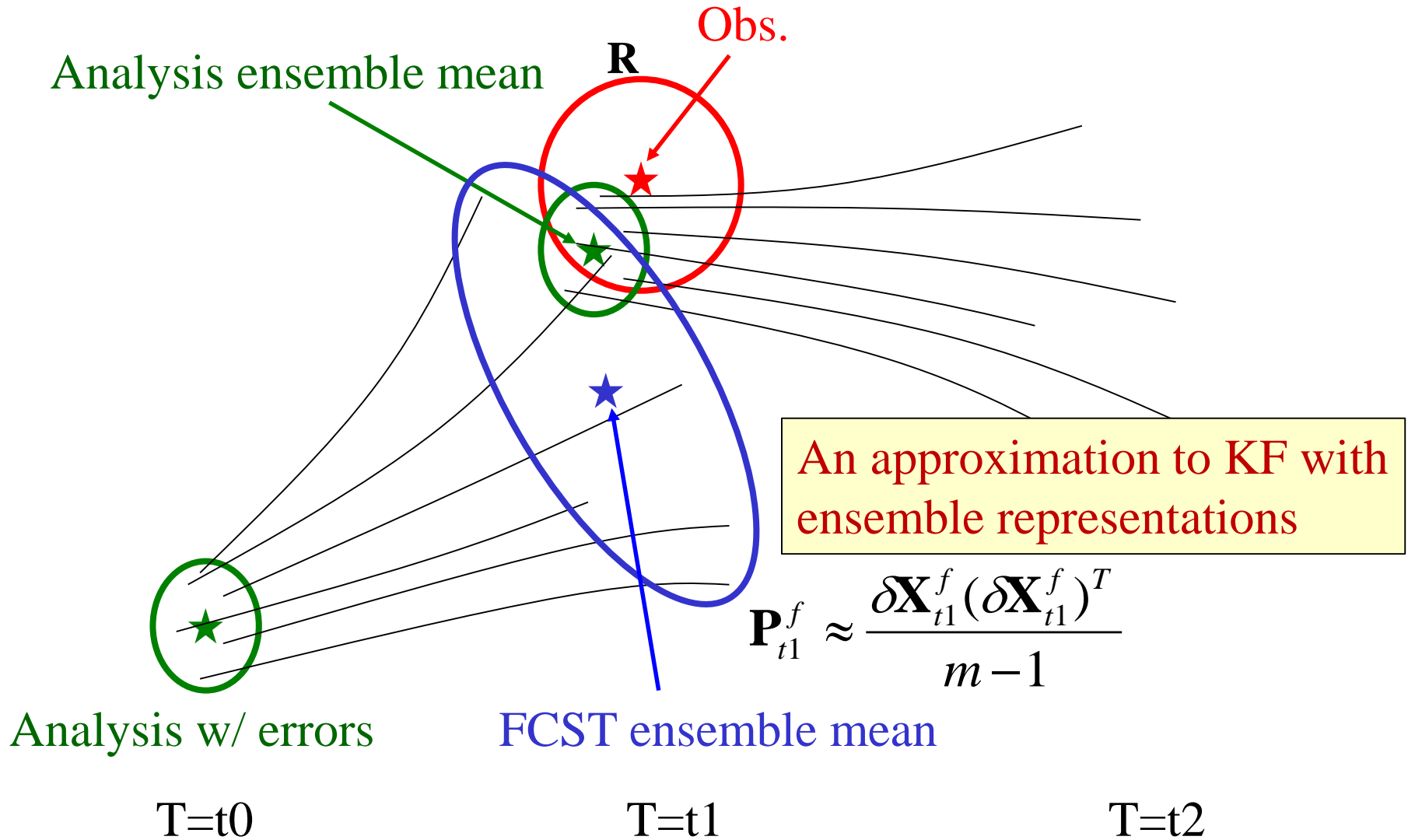
情報の組み合わせ(1次元の例)



Kalman Filter (KF)

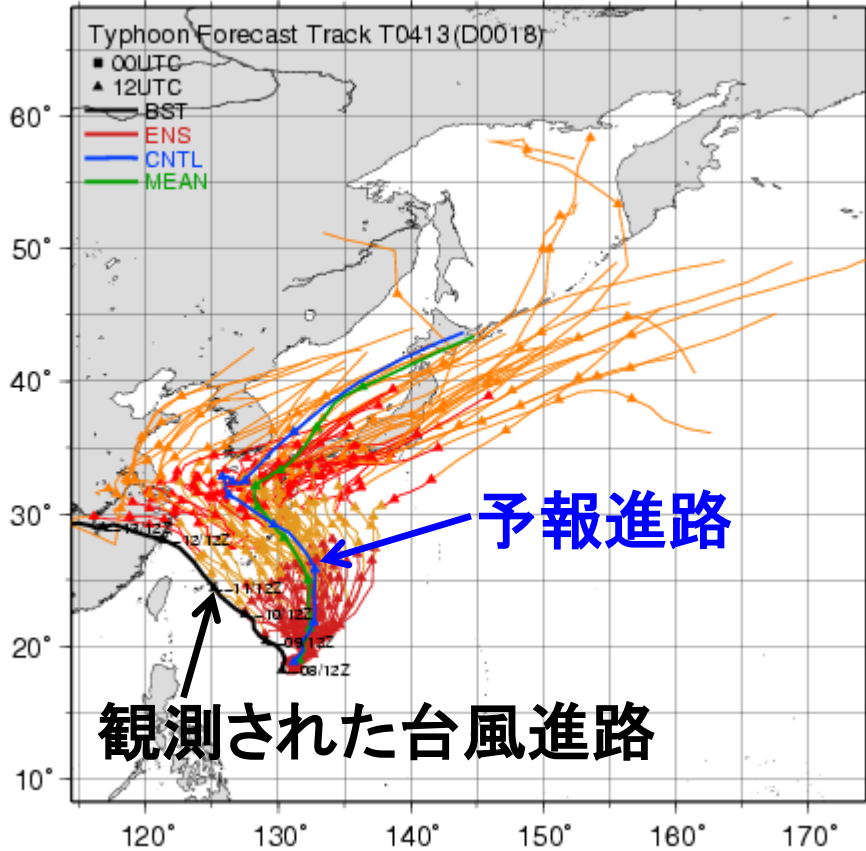


Ensemble Kalman Filter (EnKF)

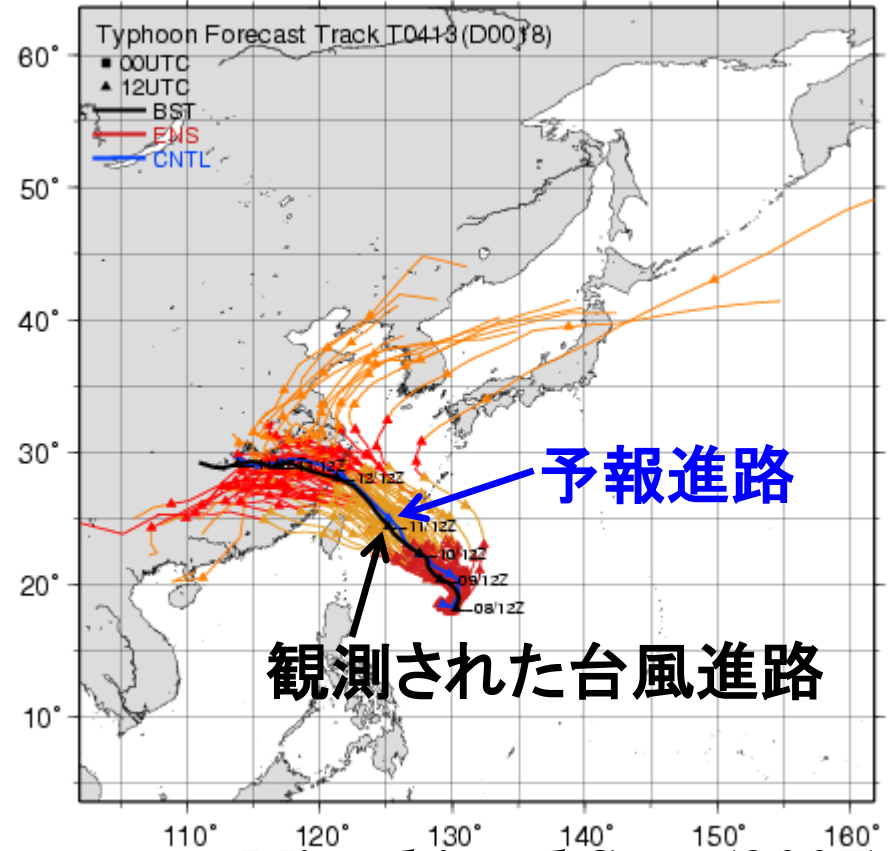


データ同化による台風予報の改善

気象庁の天気予報システム



新たに開発中のシステム



Miyoshi and Sato (2007)

同じ観測データと天気予報モデルを使用。

データ同化は重要！

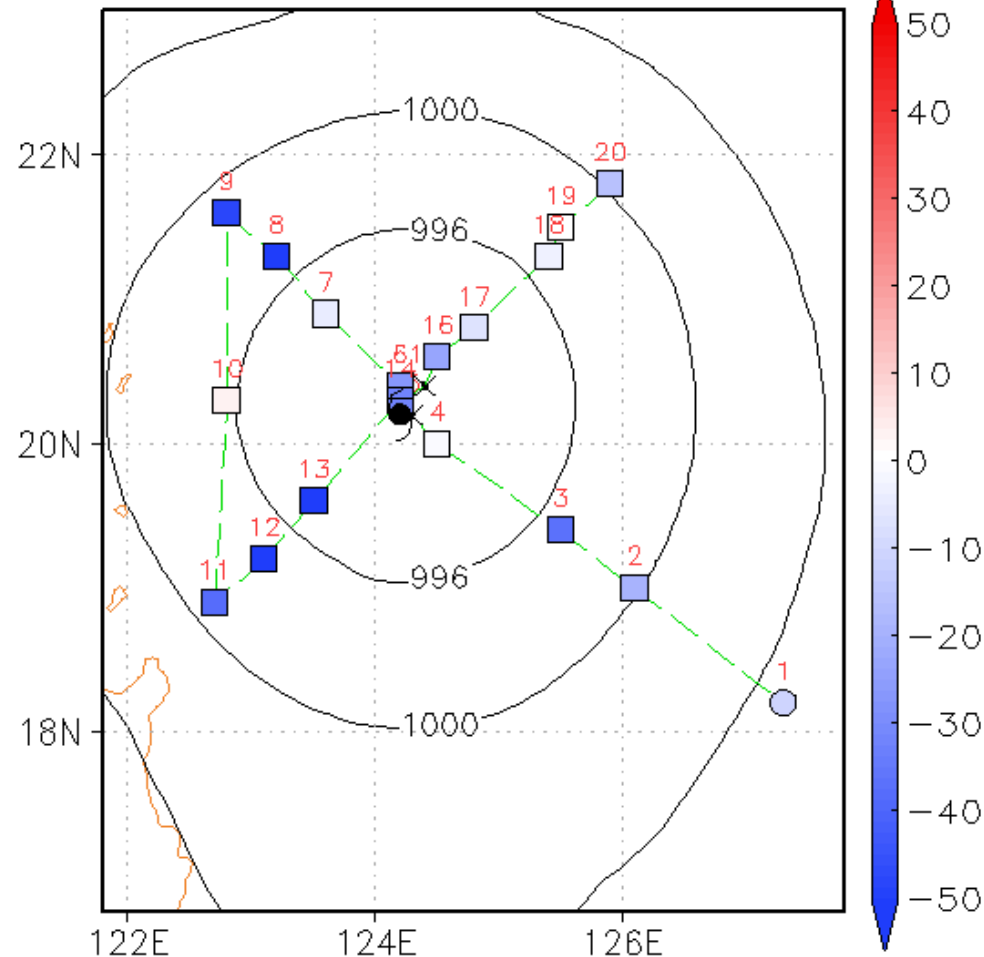
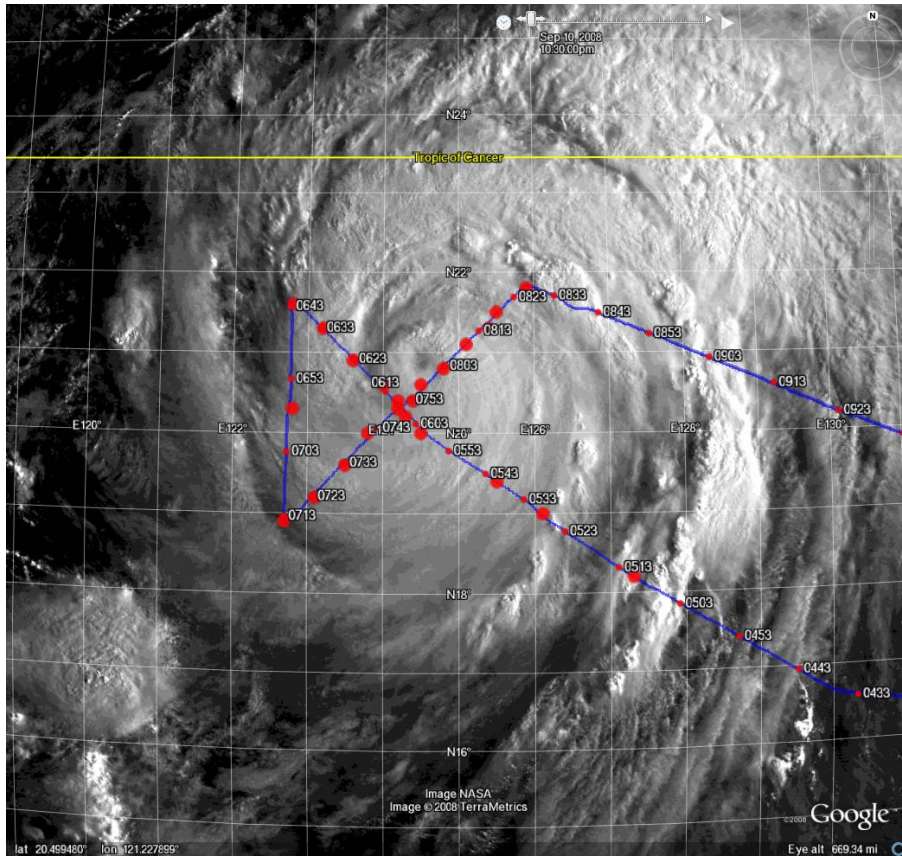
観測の“価値”を計算する

航空機観測についてその“価値”を計算した例

予報を悪化

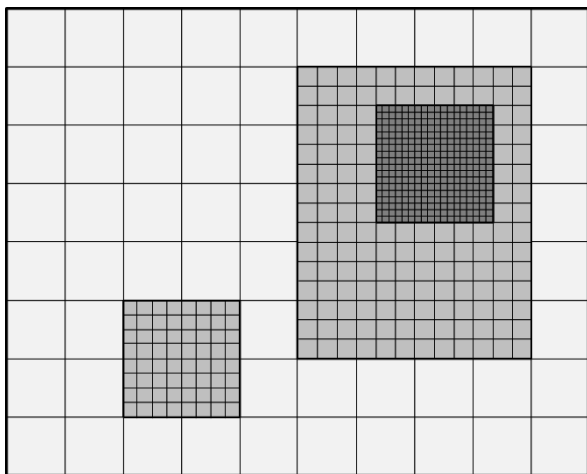
C-130

06Z10SEP2008



予報を改善

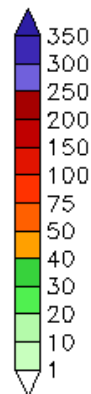
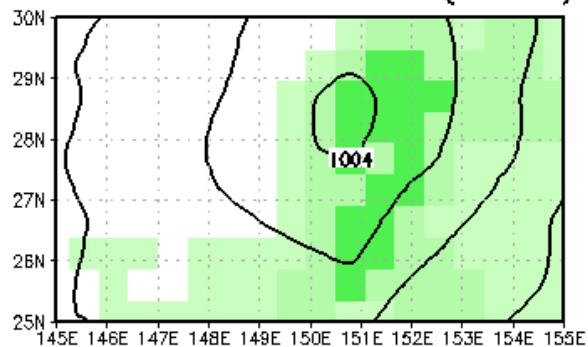
データ同化の高解像度化



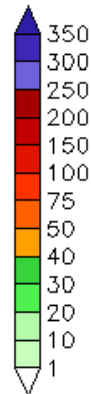
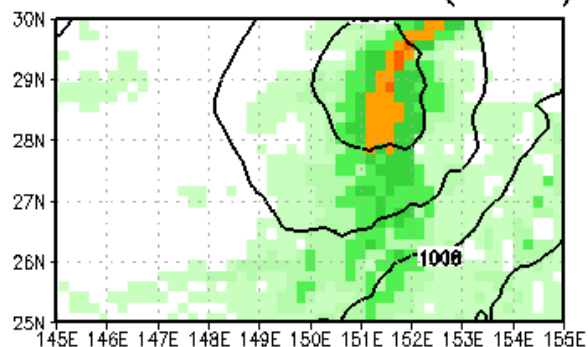
任意の格子構造に対応した
アルゴリズム研究

高解像度化

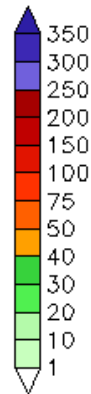
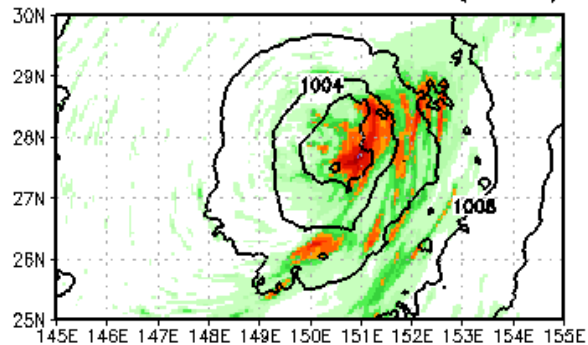
SLP and 6 hr RAIN (60KM)



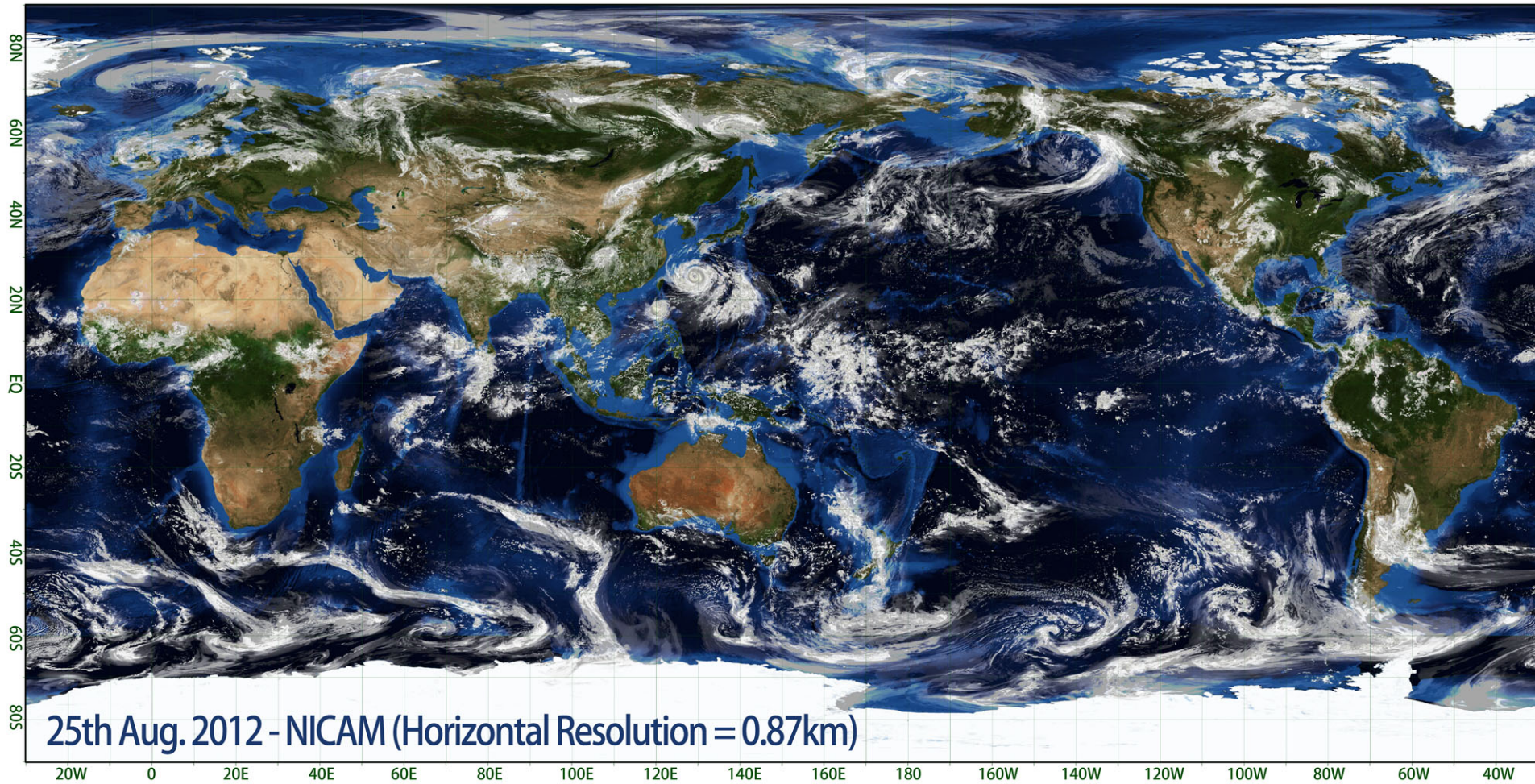
SLP and 6 hr RAIN (20KM)



SLP and 6 hr RAIN (5KM)



Global 870-m simulation



©JAMSTEC • AORI (SPIRE Field3), RIKEN/AICS
Visualized by Ryuji Yoshida

With more powerful computers...

- With an Exa-scale supercomputer (around 2020), we can run 100 members of global 870-m simulation.
- Or instead, we may have a larger ensemble at a lower resolution.

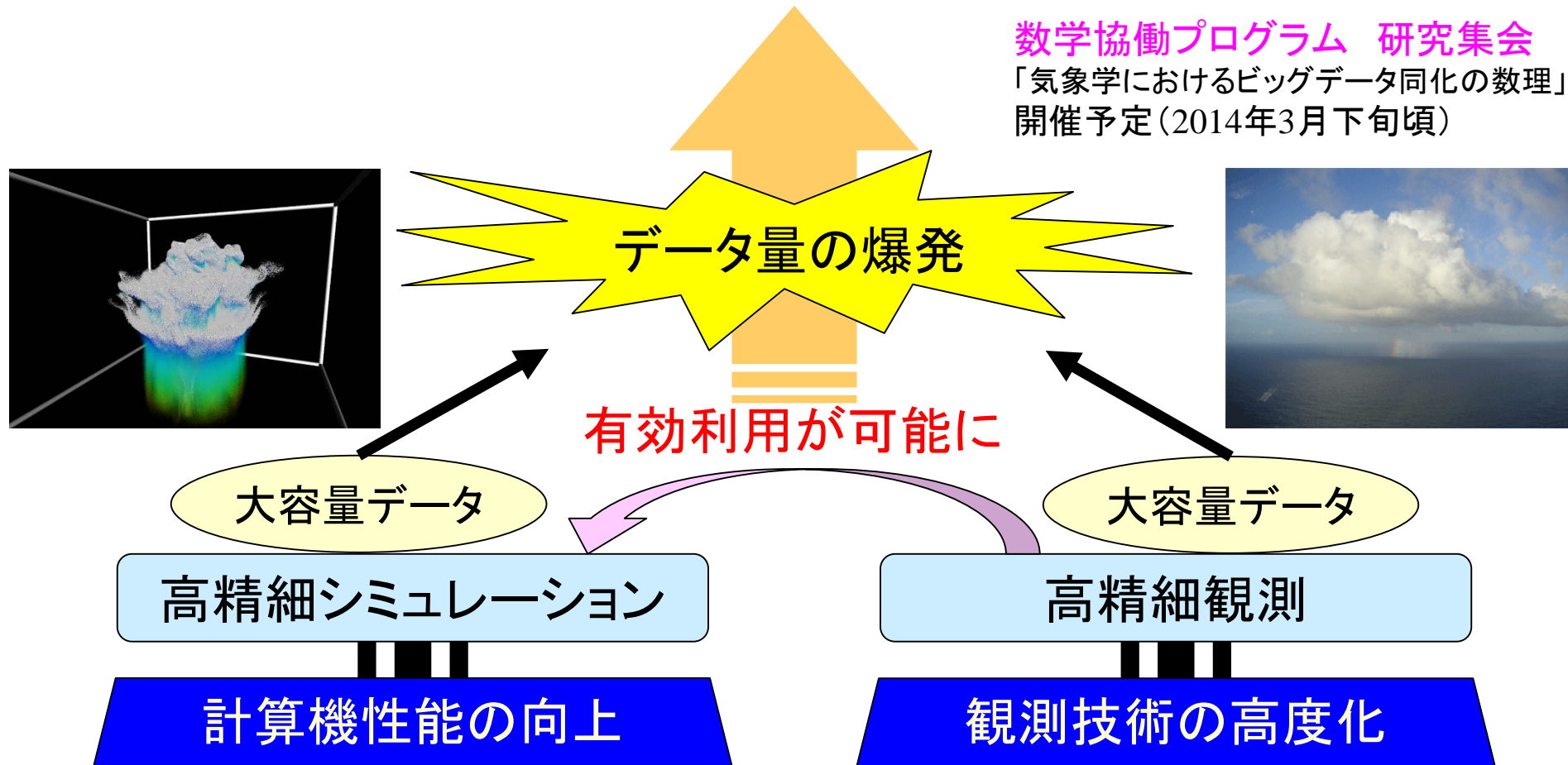


The Japanese 10-Peta-Flops K computer

今後20年を考える

“ビッグデータ同化”の時代へ

数学協働プログラム 研究集会
「気象学におけるビッグデータ同化の数理」
開催予定(2014年3月下旬頃)



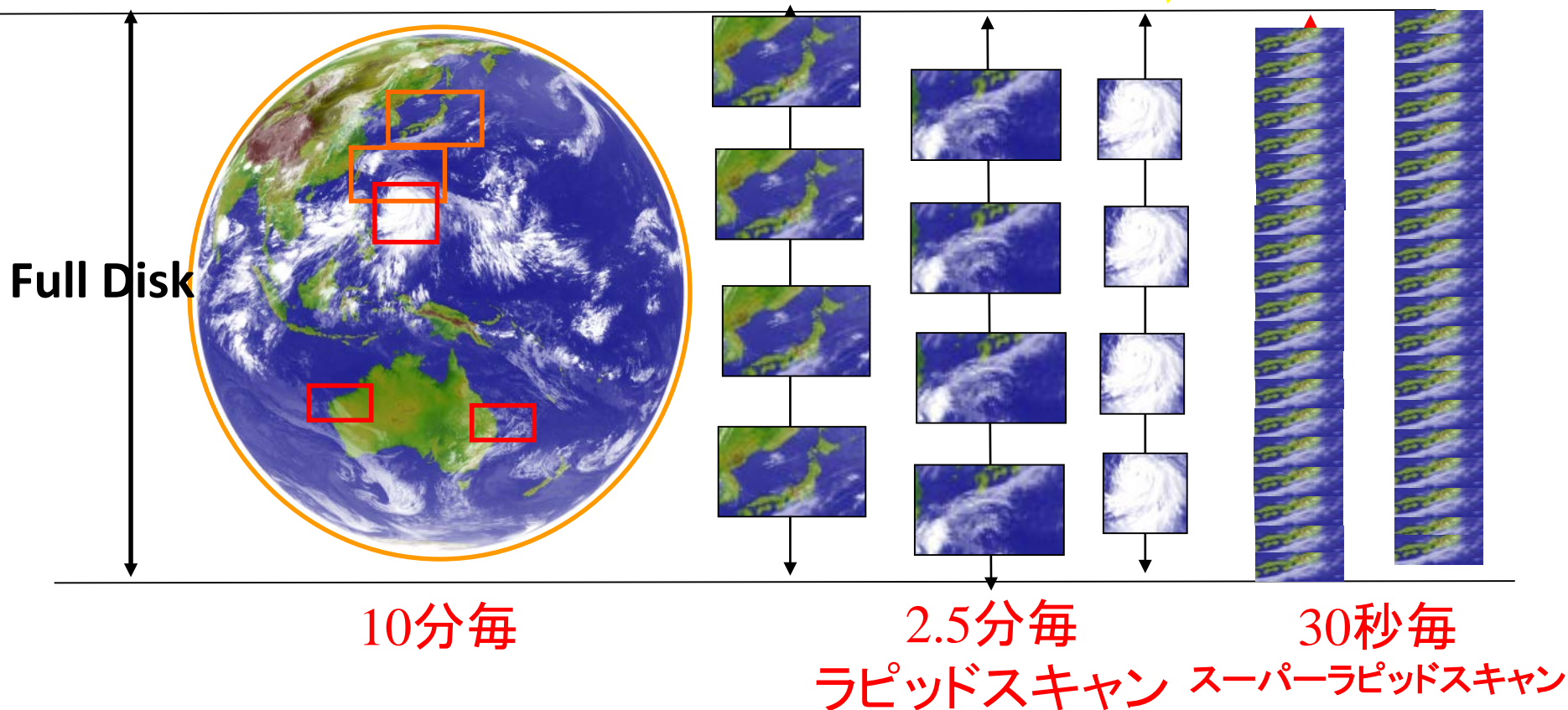
次世代静止気象衛星

ひまわり8号: H26打ち上げ予定

ひまわり9号: H28打ち上げ予定

(次世代衛星としては世界初)

30秒毎の狭領域撮像
スーパーラピッドスキャン



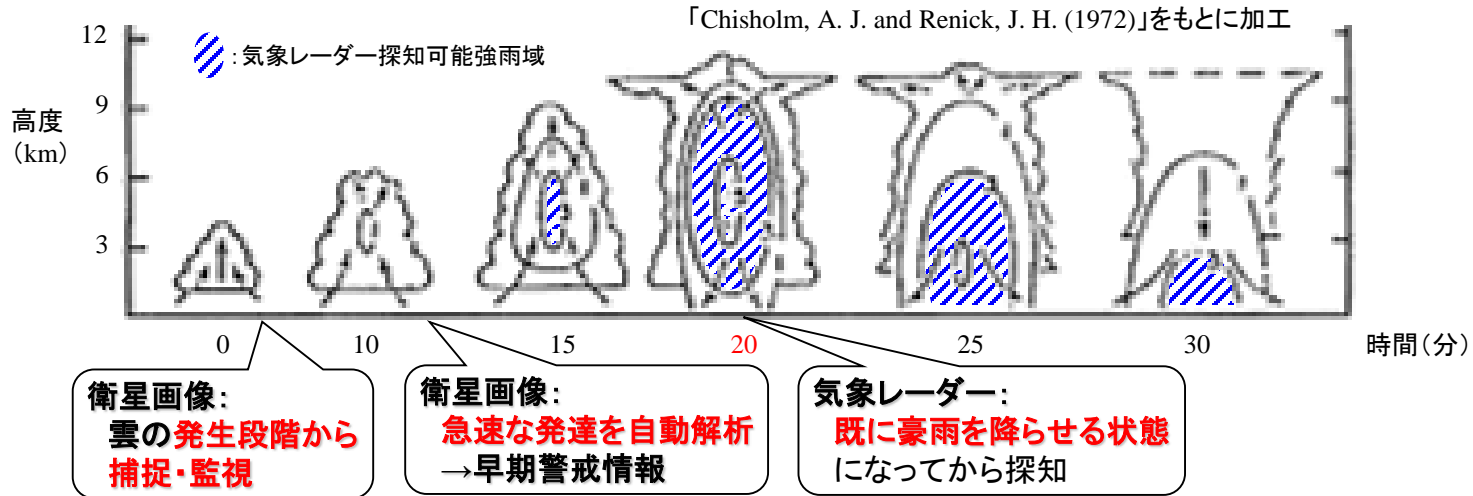
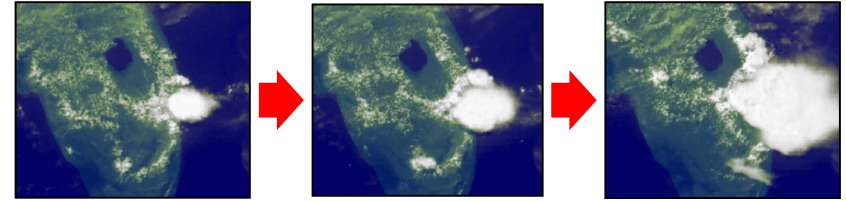
(気象庁より)

高解像度衛星画像による豪雨の検知

積乱雲の発達の早期検知

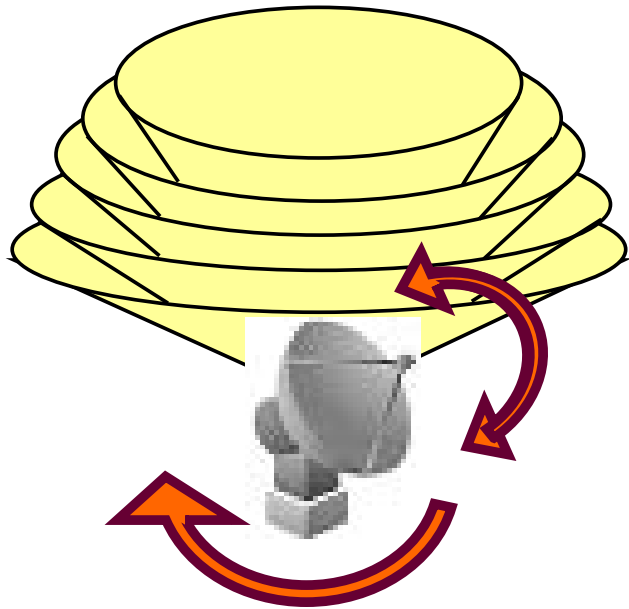
短時間間隔で取得される衛星画像を分析処理し、積乱雲を発生段階からレーダーよりもいち早く監視・検出し、集中豪雨や突風等の発生に対する早期警戒情報を提供する。

衛星で捉えた積乱雲の短時間連続画像



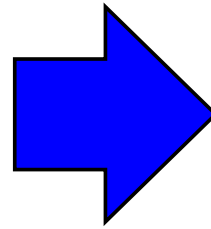
(気象庁より)

次世代型フェーズドアレイレーダー



パラボラアンテナによる
3次元立体観測(5~10分)

~15仰角

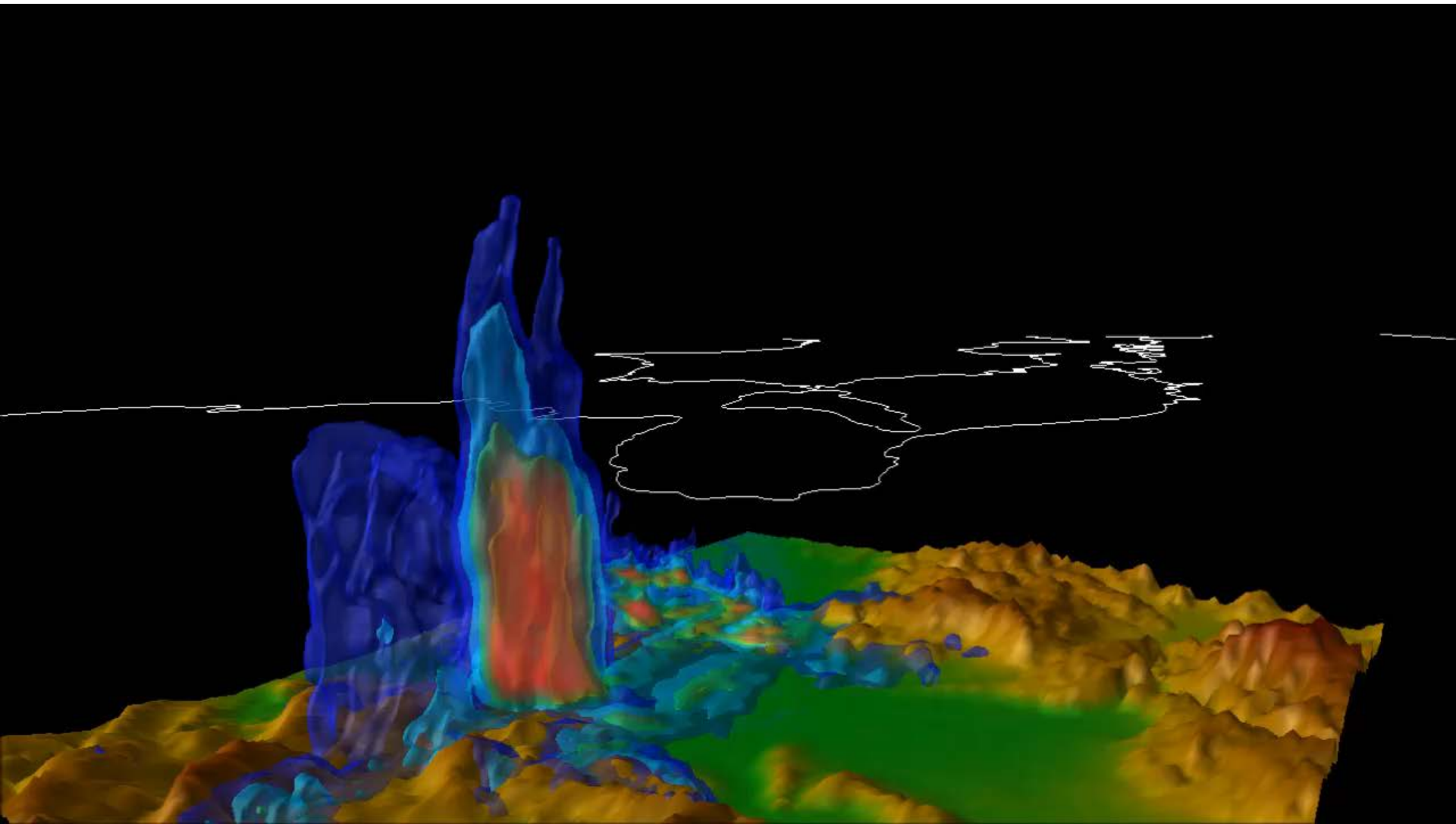


フェーズドアレイレーダーによる
3次元立体観測(10~30秒)

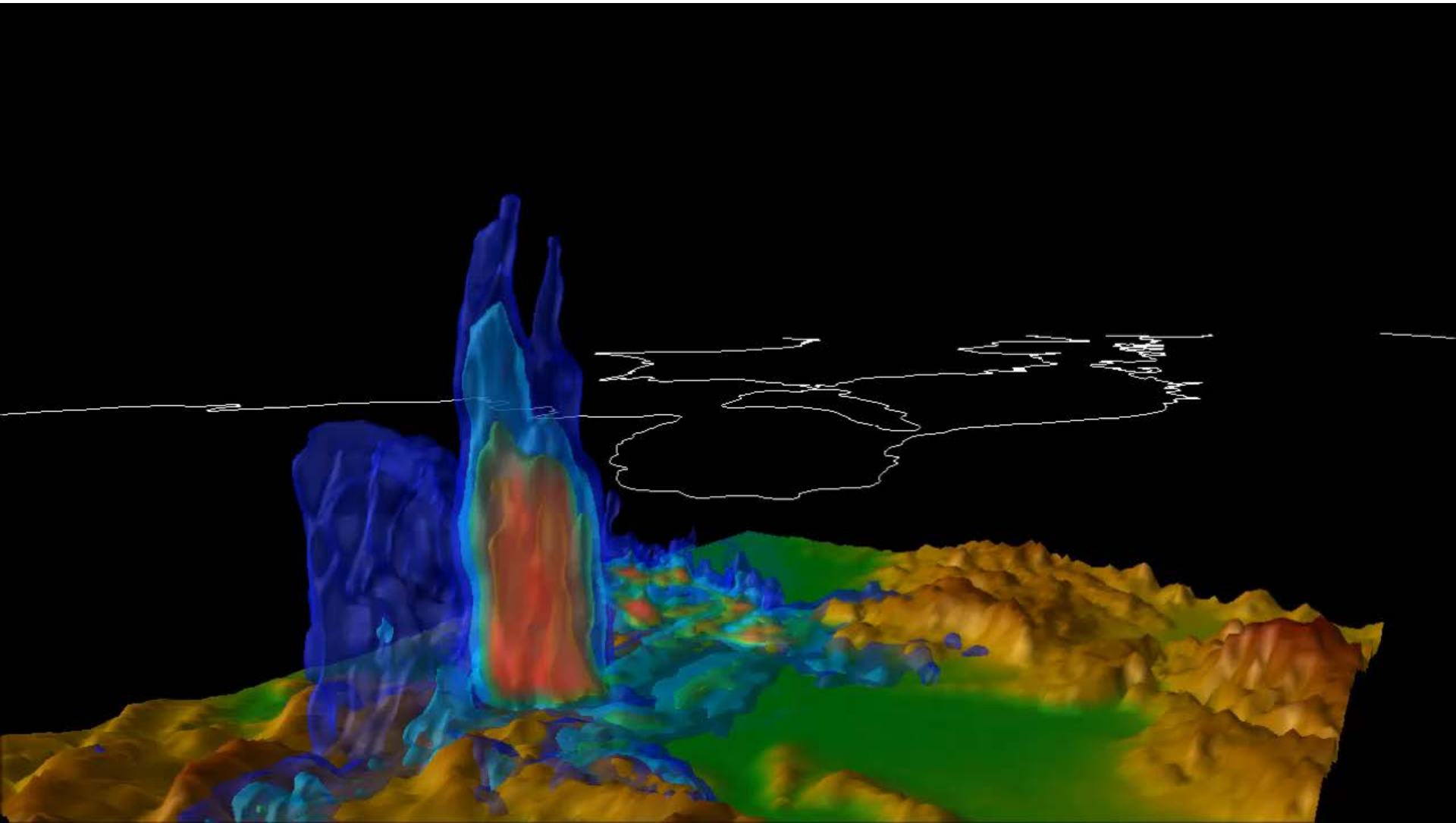
100仰角

- 次世代に普及する新しいレーダー技術。
- 現在日本では1基が稼働中(大阪大学吹田キャンパス)。
- H25年度中に、2基設置予定。そのうち1基は神戸市に設置。

従来のレーダー(5分毎)



フェーズドアレイレーダー(30秒毎)



データ開拓：(例)カメラ画像を利用できるか



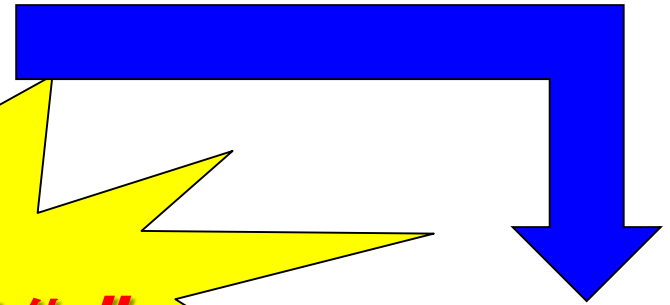
1. 情報抽出(天気、視程など) → 同化
(課題) 自動抽出技術
2. モデルから画像を作る(カメラの観測演算子) → 直接同化
(課題) 高精細3次元放射モデル(今後20年の課題か?)

“ビッグデータ同化”時代を先取り

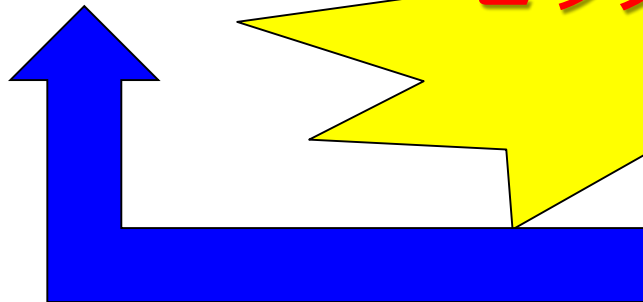
高精細シミュレーション



10年後の普及を見据えた
次世代技術のコラボレーション



“ビッグデータ同化”



シミュレーションの改善



実現象の次世代高精細観測

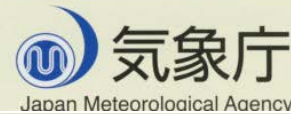
ビッグデータ同化によるゲリラ豪雨予測

●親水公園で水遊び



水位は **10分間で約1m30cm** も上昇

局地的大雨によって、平成20年7月28日、兵庫県神戸市灘区の都賀川が急激に増水し、河川内の親水公園で水遊びをしていた子供たちなどが流され、その内5人が亡くなった。（写真提供：神戸市）



増水直前



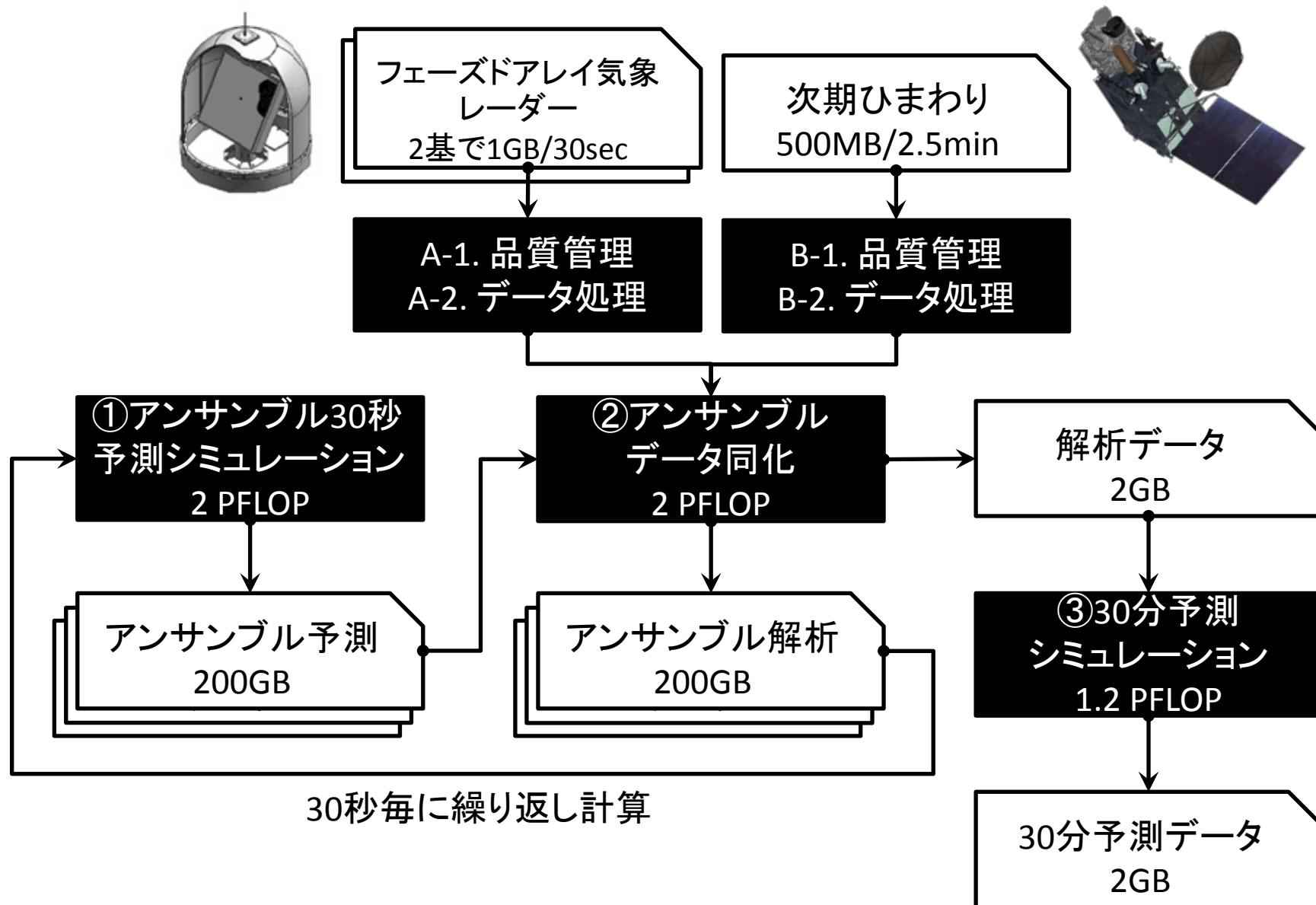
増水時

研究のねらい：高精細シミュレーションと次世代高精細観測のビッグデータ同化により、ゲリラ豪雨の30分予測に道筋を。

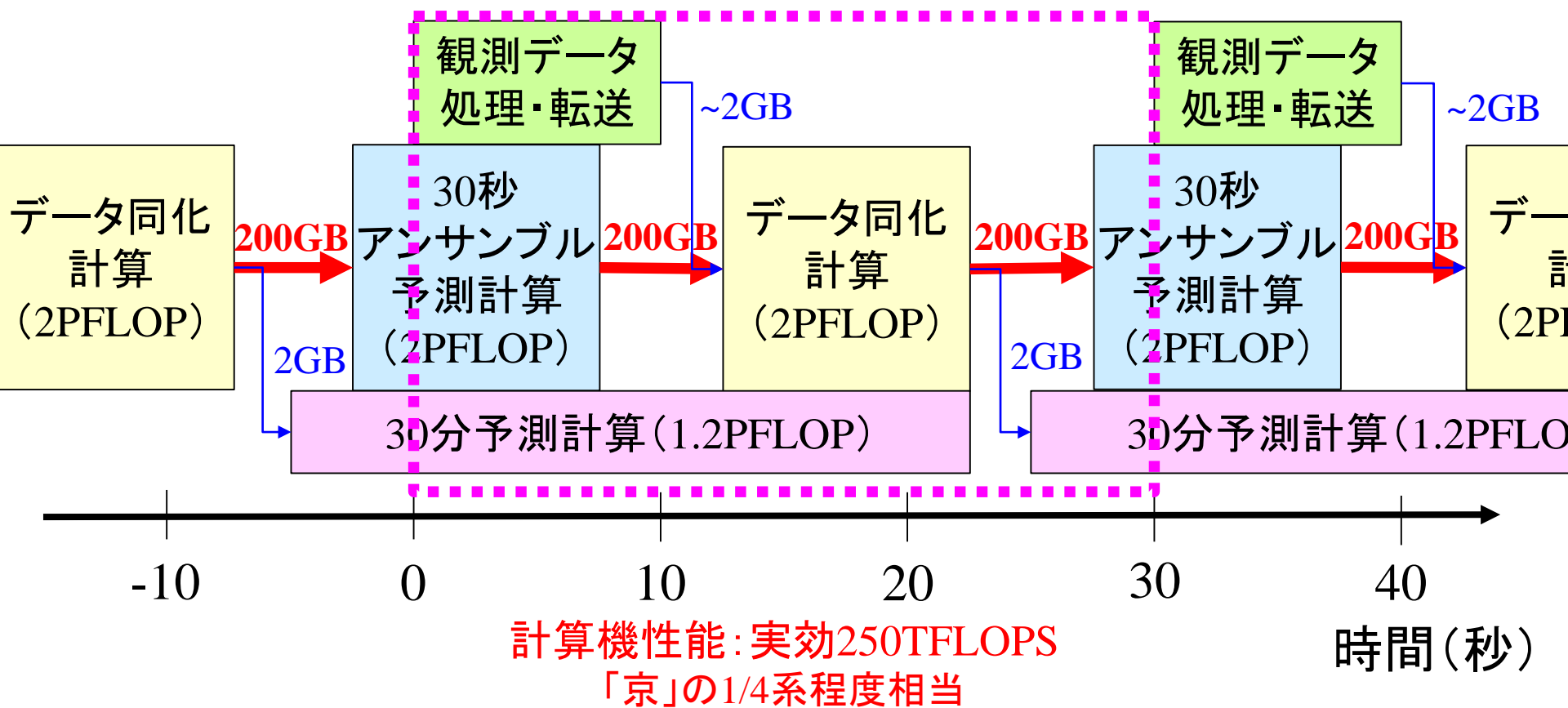
研究構想

- 30秒毎に更新する30分予報を行う画期的な天気予報システムを実証する
 - フェーズドアレイレーダー、次期衛星ひまわりの観測データを有効活用
 - ゲリラ豪雨の短時間予測 → 防災・減災に貢献
 - 「見えないものを見る」 → 気象学的発見へ

研究構想：システムフローチャート



タイムスケジュール(案)



従来のデータ同化システムとの違い

- 全球データ同化では6時間毎、メソデータ同化では3時間毎にデータ同化。
- ケタ違いの高頻度 → 桁違いに高速なデータ処理が不可欠。

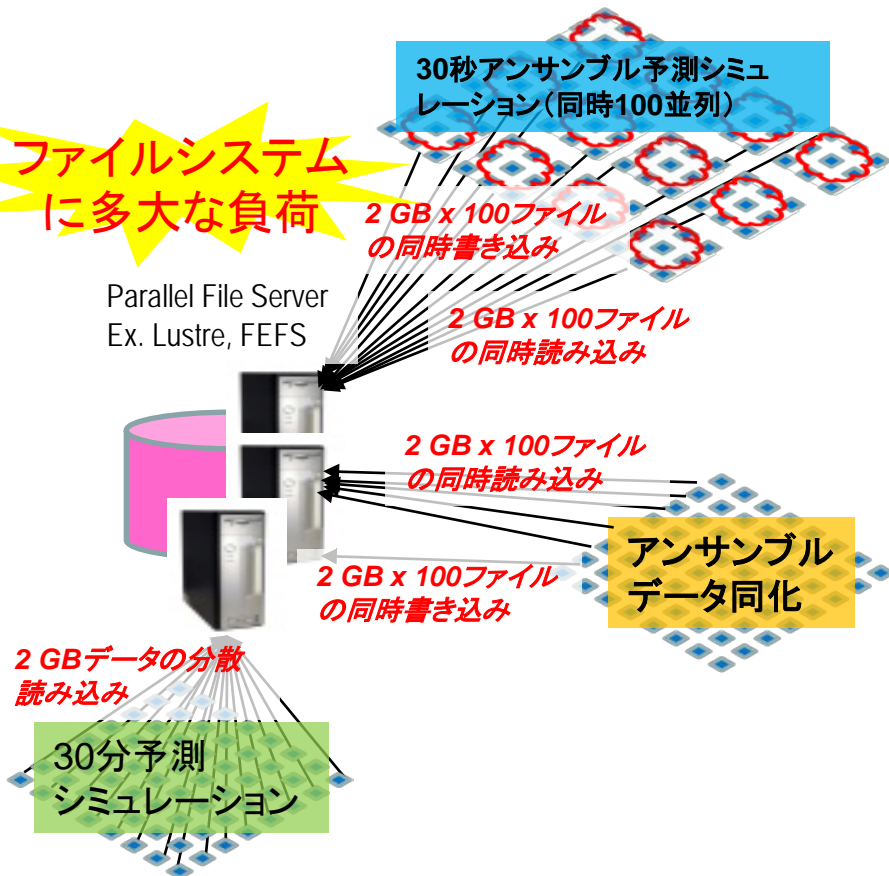
ビッグデータ同化の革新技術

- 計算機科学とのコラボレーション
 - ハードウェアとのコデザイン
- 高速I/Oに対応したデータ同化アルゴリズム
 - スパコンの200GB/sの超並列IO性能を生かすアルゴリズム開発
- 観測サイトにおける高速な品質管理・データ処理技術
 - 生データ取得から品質管理、データ処理までを10秒以内で終わるための高速処理技術

高速I/Oを実現するビッグデータ同化技術

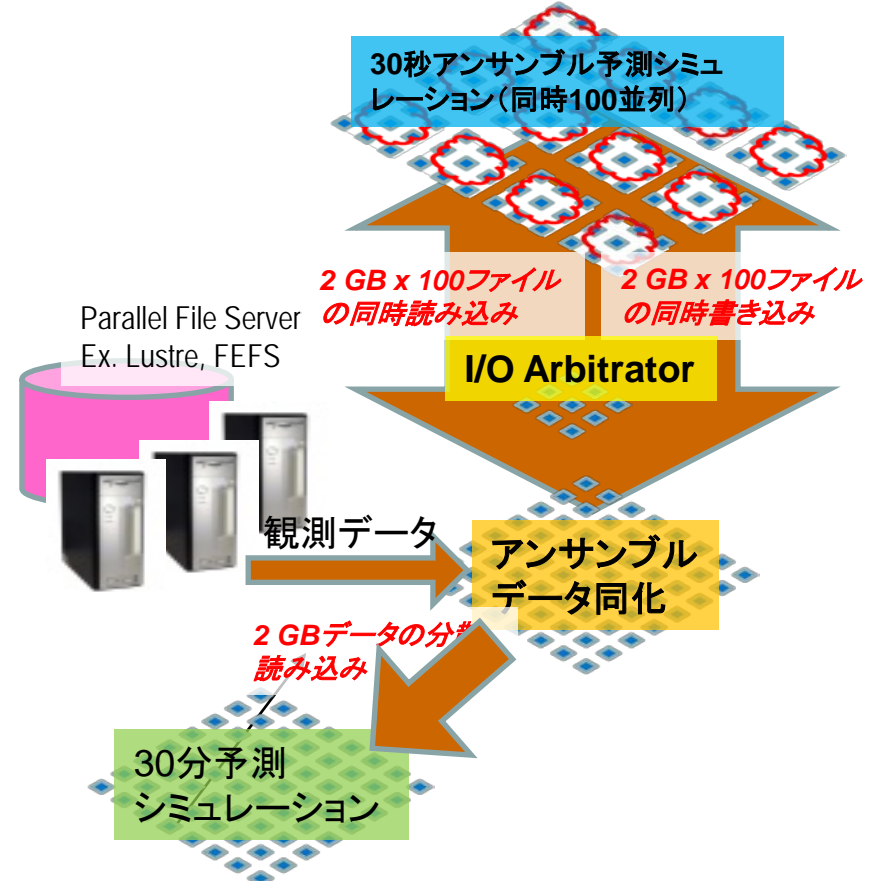
現状：ファイルシステム経由

- データ同化と予測シミュレーションのジョブ間データ転送がファイルシステムを介する



提案：ジョブ間並列通信機構

- ファイルシステムを介さず直接データ転送
- ファイルI/O APIを変更せず、ミドルウェアによってジョブ間の並列データ転送を実現
- I/O Arbitratorによるデータの集約配布



今こそ、ビッグデータ同化研究を

◆天の時

- ゲリラ豪雨予測は喫緊の社会的課題
- 次世代気象レーダー、京コンピュータの稼働、次期ひまわりの打ち上げ
- これら次世代技術により、サイエンス・ビッグデータの応用が可能に



京コンピュータ

- H24秋から稼働

◆地の利

- 我が国には10年後の普及を見据えた次世代技術が揃っている
 - 世界的に見ても非常に稀有なアドバンテージ
- 神戸に2基の次世代気象レーダー



次期衛星ひまわり

- H26打ち上げ予定

◆人の和

世界で活躍する各分野のエース研究者のコラボレーション



フェーズドアレイ気象レーダー

- H24夏から1基稼働
- H25年度中に追加設置

世界に先立って、幅広く応用可能な革新的基盤技術を創出
→ 本研究による「ビッグデータ同化」技術が世界の研究・実利用に貢献

「ビッグデータ同化」の技術革新の創出によるゲリラ豪雨予測の実証

(H25-30年度)

三好建正 (理化学研究所 チームリーダー)

CREST
戦略的創造研究推進事業
Core Research for Evolutional Science and Technology

研究領域「科学的発見・社会的課題解決に向けた各分野のビッグデータ活用促進のための次世代アプリケーション技術の創出・高度化」(H25-30年度)

研究の概要

データ同化は、シミュレーションと実世界のデータを融合し相乗効果をもたらす基盤技術です。本研究では、次世代の高精細シミュレーションと新型センサによる「ビッグデータ」を扱うための「ビッグデータ同化」の技術革新を創出し、ゲリラ豪雨予測に応用して、フェーズドアレイ気象レーダー、次期気象衛星ひまわり、京コンピュータという我が国が世界に誇る次世代技術を駆使して実証実験し、防災・減災に資するとともに、気象学的ブレークスルーをもたらします。



フェーズドアレイ気象レーダー



次期気象衛星ひまわり



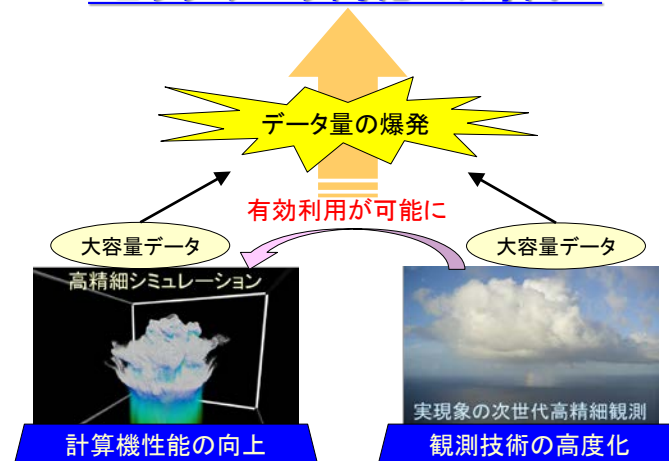
京コンピュータ

社会的・経済的・科学的課題と本研究による解決策 (提案の独創性、新規性等を含む)

- ・ゲリラ豪雨の短時間予測は、防災・減災の観点で重要な社会的課題
- ・積乱雲中の空気の流れの詳細構造など「見えないものを見る」ことの気象学的意義

これらの課題に対し、本研究では、30秒毎に更新するリードタイム30分の天気予報という従来では考えられない画期的なシステムを、フェーズドアレイ気象レーダー、次期衛星ひまわり、及び京コンピュータを駆使して実証実験します。本研究で目指すビッグデータ時代のデータ同化におけるリアルタイム処理は、現在のデータ同化技術の延長では到底実現し難いため、本研究では、ビッグデータを扱うデータ同化、すなわち「ビッグデータ同化」という技術革新を創出し、既存の技術では実利用が難しい次世代型センサによるビッグデータを有効利用することで、科学的発見を生むとともに、ビッグデータ利用の基盤技術を確立し、広く防災・減災に貢献します。

“ビッグデータ同化”の時代へ



研究成果により想定されるインパクト、将来像、イノベーション創出への寄与など

ゲリラ豪雨の短時間予測による防災・減災への貢献、気象学的発見へのブレークスルー、2020年代の現業天気予報に向けた提言
将来のエクストリーム・コンピューティング時代における幅広いシミュレーション分野で必須となるデータ同化基盤技術の創出

Mathematical Science on Big Data Assimilation in Meteorology

 **Date:** March 19 – 21, 2014

 **Venue:** Room 110 of Bldg. 3
Department of Mathematics
Kyoto University

 **Keynote Lecturer:** Professor Brian Hunt (University of Maryland)
"Ensemble Methods for Complex Models and Big Data"

 Invited Speakers (without honorifics):	 Organizing Committee:
Kazuyuki Nakamura (Meiji)	Takashi Sakajo (Kyoto)
Takemasa Miyoshi (RIKEN/AICS)	Takemasa Miyoshi (RIKEN/AICS)
Yasuaki Hiraoka (Kyushu)	Masaru Inatsu (Hokkaido)
Shinsuke Satoh (NICT)	Yoshitaka Saiki (Hitotsubashi)
Kotaro Bessho (JMA/MSU)	Naoto Nakano (Tohoku)

URL <http://www.wpi-aimr.tohoku.ac.jp/~nakano/cwm2013/>

 Organized by Department of Mathematics, Kyoto University
RIKEN Advanced Institute for Computational Science
 Supported by Mathematics-Meteorology Collaborative Research Team (MaeT)
JSPS Grant-in-Aid for Challenging Exploratory Research (25610028)
JST/CREST Miyoshi Team: Innovating "Big Data Assimilation" technology
for revolutionizing very-short-range severe weather prediction

