

## 次世代気象衛星ひまわり 8 号と機械学習技術を用いた アジア・オセアニア地域における地表面温度観測システムの開発

山本 雄平

千葉大学 環境リモートセンシング研究センター

### 【概要】

地表面温度は、陸面の様々な環境における熱収支・水収支過程に係る物理量であり、水資源・農業気象研究、植生動態・砂漠化の監視、都市気候研究等において重要な役割を果たしている。本研究では、我が国の次世代静止軌道衛星「ひまわり 8 号」データを用いた新たな地表面温度推定手法を試みた。まず、従来の地表面温度推定式である半経験的手法と、今回の提案手法である機械学習（ランダムフォレスト回帰）を用いた推定手法の推定精度を比較した。その結果、半経験的手法の方が安定した推定精度をもつことが分かった。これは、地表面温度の推定に重要となる地表面射出率の推定段階で、機械学習法に大きなエラーが生じたためと考えられる。特に東アジアにおいて雲が多く生成される（つまり観測値にノイズが多く混入する）環境では、学習段階では想定されていない観測値によって解が一定の値をとってしまうことが分かった。世界の中でもとりわけ雲量の多い東アジアや東南アジア地域においては、機械学習による推定は困難であるという結論を得た。次に、東アジアにのみ焦点を当てていた従来の半経験的手法をオーストラリアもカバーできるように改良した。本報告書では、後半で行った推定手法の拡張にあたって行ったことについて述べていく。

### 【教師データの再構築】

東アジアに焦点を当てた地表面温度の推定アルゴリズムは、 $60^{\circ}\text{C}$ 以上の超高温な地表面状態が考慮されておらず、また高温時には大気の水蒸気量も非線形に増加するような教師データが用いられていた。つまり、オーストラリア大陸で生じうる「超高温・乾燥状態」にそのまま従来アルゴリズムを適用すると大きく過小推定されてしまう問題があった。そこで、地表面温度  $60^{\circ}\text{C}$ – $70^{\circ}\text{C}$  のケース、可降水量  $1.0\text{g}/\text{cm}^2$ – $5.0\text{g}/\text{cm}^2$  のケースを追加し、超高温・乾燥時にも適用できるように推定アルゴリズムの係数を再計算した。改良後の出力例として、図 1 に 2018 年 12 月に発生したオーストラリア熱波時の地表面温度分布を示す。従来手法の限界値である  $55^{\circ}\text{C}$  以上の高温域が推定されていることがわかる。

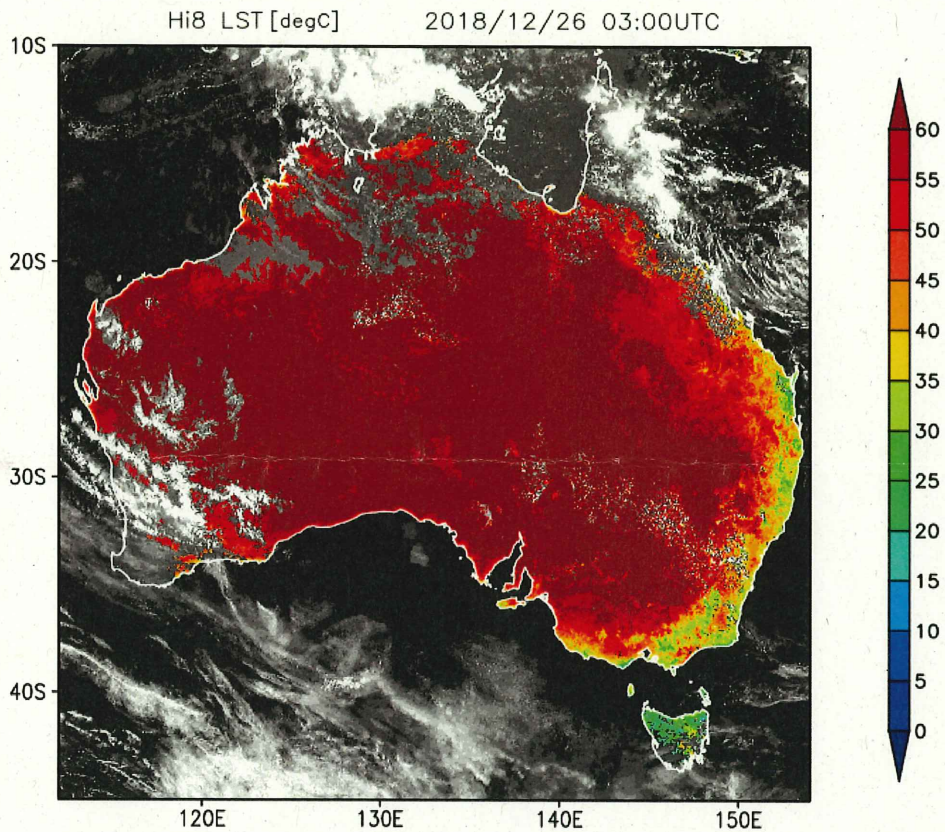


図1. 改良手法により推定された2018年12月26日03:00UTCにおける地表面温度分布(°C)。背景のグレースケールの画像は、ひまわり8号のバンド3で観測された雲反射率画像である。

### 【雲域検出アルゴリズムの構築】

地表面温度は、地表面から宇宙空間へ射出される熱赤外放射量から推定されるわけであるが、地表面と人工衛星センサとの間に雲が存在すると、赤外放射が遮られてしまい、雲の上端からの放射量が観測される(即ち雲頂面温度らしきものが推定されてしまう)。そのため、雲域を正しく検出し、推定対象から除外することが推定精度に非常に重要である(図1は雲域検出処理後の画像)。雲域の検出には、陸域と雲域との反射率特性や温度特性の違いを利用する。例えば、ひまわり8号のバンド3で観測される反射率は、陸域は低く雲域は高いという性質があり、バンド14で観測される輝度温度は、陸域で高く雲域で低いという性質がある。これを利用し、バンド3で「ある閾値」よりも低い反射率が観測されると「晴」、高い反射率が観測されると「雲」と判別し、バンド14で「ある閾値」よりも高い輝度温度が観測されると「晴」、低い輝度温度が観測されると「雲」と判別する。これらの閾値データを作成するには、数年分の晴れた状態の反射率と輝度温度の情報が必要となる。また、オーストラリア大陸は東アジア域とは陸面特性が大きく異なり、陸面の反射率が非常に高く、輝度温度は極端に高い。そのため、東アジア仕様の閾値データをそのまま適用すると、反射率特性を利用した検出テストは厳し

すぎて常に「雲」と判別してしまい、輝度温度特性を利用した検出テストは緩すぎて常に「晴」と判別されてしまう。そこで本研究では、2015年7月（ひまわり8号の運用開始時期）から2020年6月までの5年分の観測データを用いて閾値データを作成した。閾値データの例を図2に示す。オーストラリア大陸の反射率と輝度温度の閾値は、他の地域と比べて高く設定されているようすが分かる。

新しい閾値データを用いたオーストラリアにおける雲域検出精度は、現時点では東アジアのそれよりも高められてはいない。検出精度を高めるために、時間分解能の向上（閾値データの数を増やす）や他のバンド観測値を用いた検出テストの追加など、今後さらなる改良を行っていく。

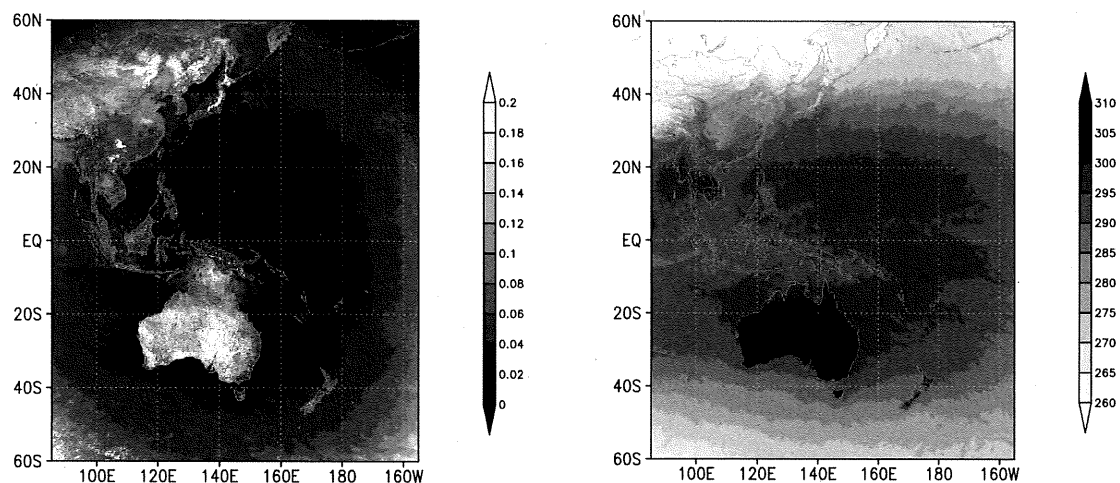


図2. 1月の03:00UTCにおけるバンド3反射率(左)、バンド14輝度温度(右)の閾値データ。

### 【おわりに】

閾値データは月別に作成され、各月で1時間おきに24時間分作成されます（即ち24×12個の閾値データ）。一つの閾値データを作成するには、576MBのひまわり8号のスナップショット画像を1ヶ月間（約30日）×5年分処理します（=約86.4GB）。これを24×12個分、バンド3とバンド14のものを作成するので、合計で約50TBもの容量が必要となります。本研究助成を受けられたことにより、今回の大規模データの計算を行うことができました。エスペック株式会社および研究助成事務所の方々に厚く御礼申し上げます。