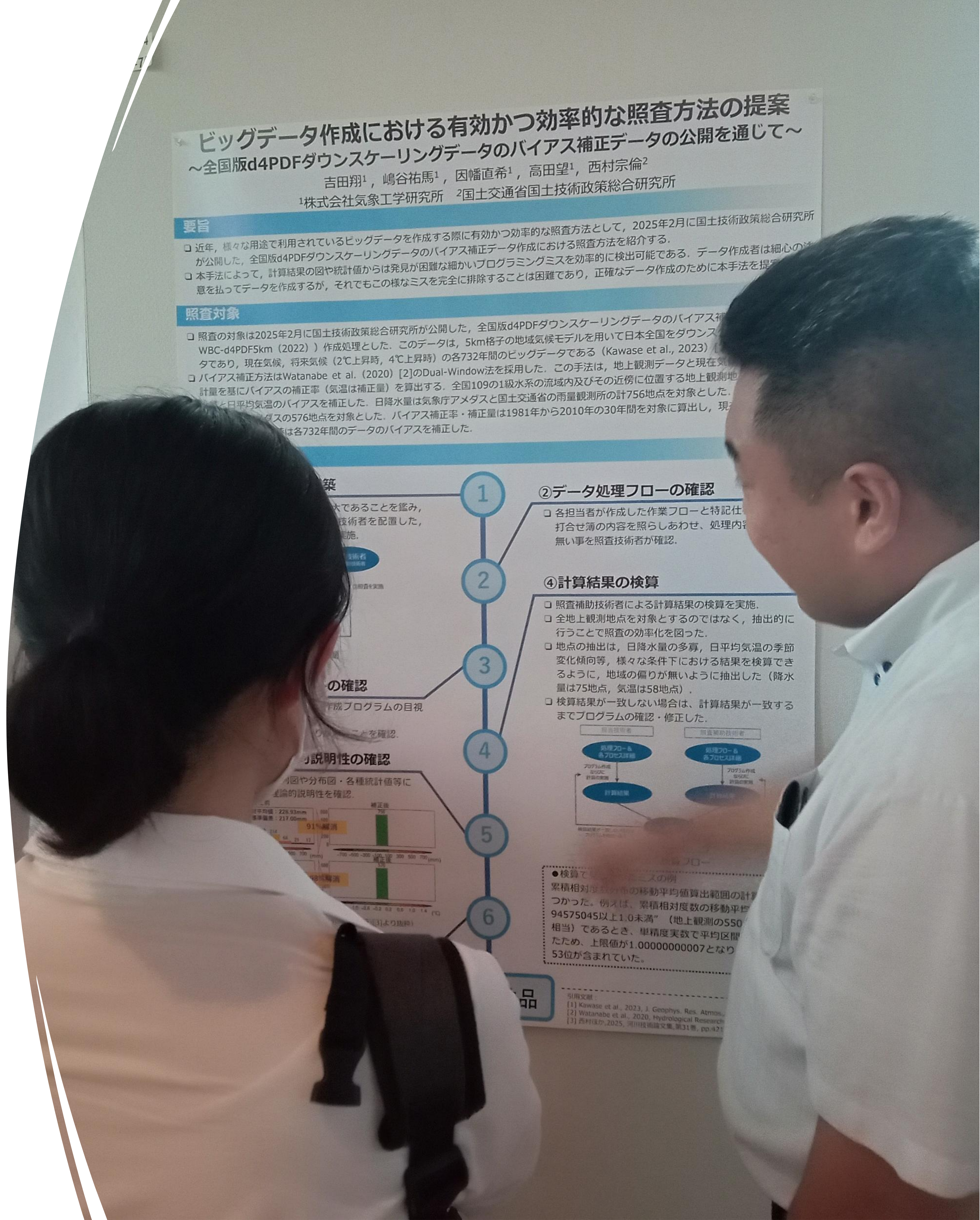


(当社社員 発表風景)



ビッグデータ作成における有効かつ効率的な照査方法の提案 ～全国版d4PDFダウンスケーリングデータのバイアス補正データの公開を通じて～

吉田翔¹，嶋谷祐馬¹，因幡直希¹，高田望¹，西村宗倫²
¹株式会社気象工学研究所 ²国土交通省国土技術政策総合研究所

要旨

- 近年，様々な用途で利用されているビッグデータを作成する際に有効かつ効率的な照査方法として，2025年2月に国土技術政策総合研究所が公開した，全国版d4PDFダウンスケーリングデータのバイアス補正データ作成における照査方法を紹介する．
- 本手法によって，計算結果の図や統計値からは発見が困難な細かいプログラミングミスを効率的に検出可能である．データ作成者は細心の注意を払ってデータを作成するが，それでもこの様なミスを完全に排除することは困難であり，正確なデータ作成のために本手法を提案する．

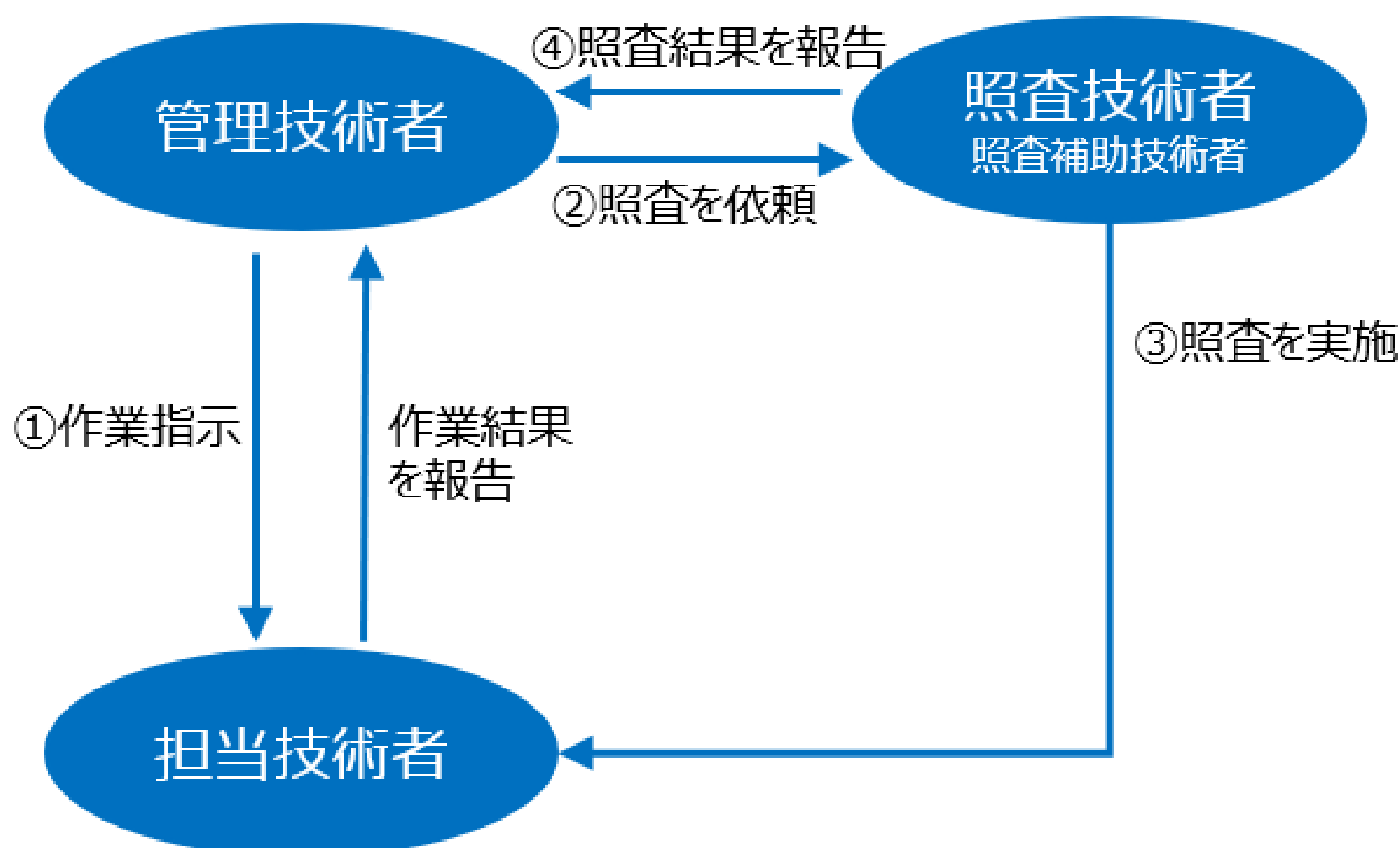
照査対象

- 照査の対象は2025年2月に国土技術政策総合研究所が公開した，全国版d4PDFダウンスケーリングデータのバイアス補正データ（略称：WBC-d4PDF5km（2022））作成処理とした．このデータは，5km格子の地域気候モデルを用いて日本全国をダウンスケーリングしたデータであり，現在気候，将来気候（2℃上昇時，4℃上昇時）の各732年間のビッグデータである（Kawase et al., 2023）[1]．
- バイアス補正方法はWatanabe et al.（2020）[2]のDual-Window法を採用した．この手法は，地上観測データと現在気候データの順序統計量を基にバイアスの補正率（気温は補正量）を算出する．全国109の1級水系の流域内及びその近傍に位置する地上観測地点における日降水量と日平均気温のバイアスを補正した．日降水量は気象庁アメダスと国土交通省の雨量観測所の計756地点を対象とした．一方，日平均気温は気象庁アメダスの576地点を対象とした．バイアス補正率・補正量は1981年から2010年の30年間を対象に算出し，現在気候は360年，2℃上昇，4℃上昇時は各732年間のデータのバイアスを補正した．

照査方法

①品質管理体制の構築

- 照査対象となるデータ量が膨大であることを鑑み，照査技術者に加え，照査補助技術者を配置した，充実した品質管理体制で照査を実施．



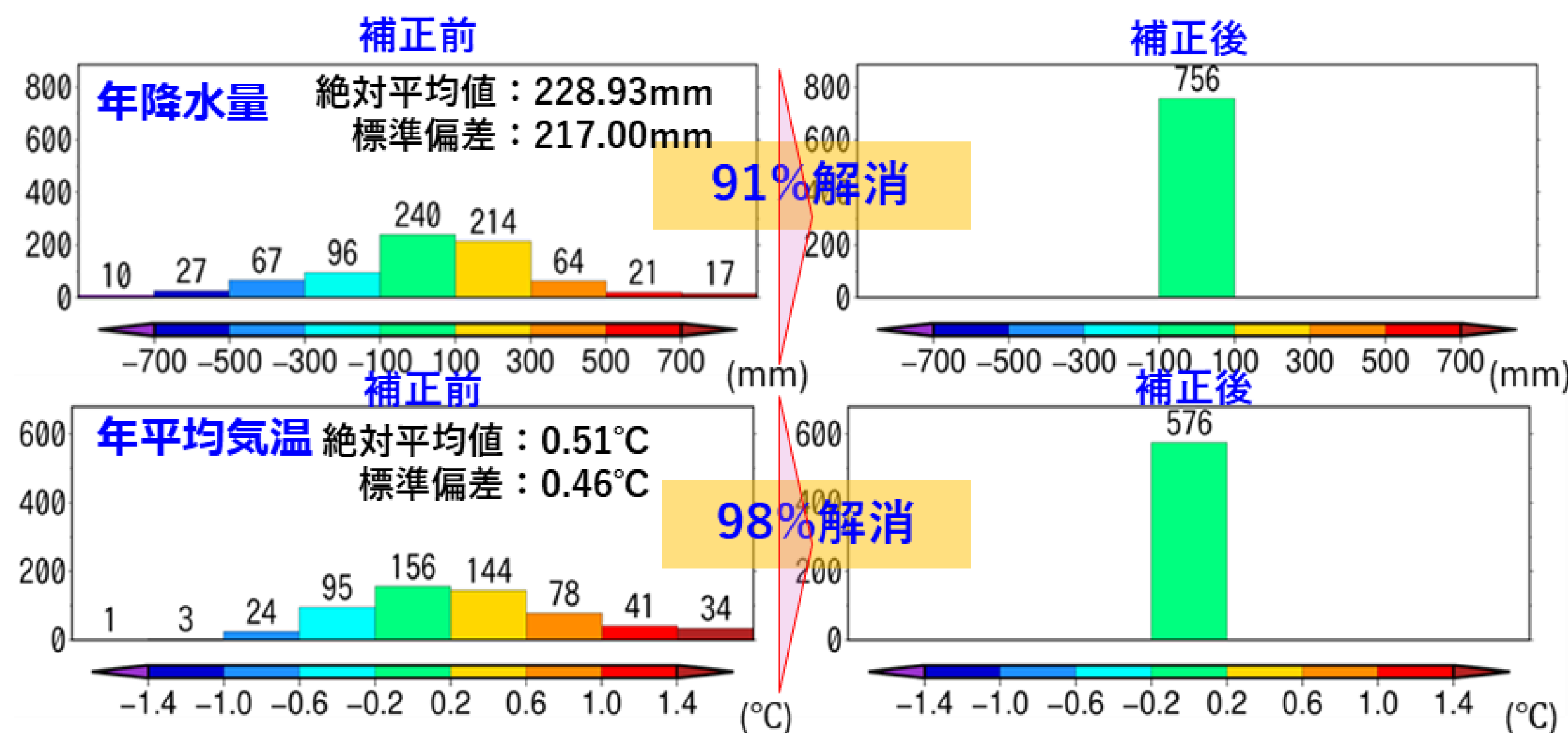
図．品質管理体制

③プログラムコードの確認

- 照査技術者によるデータ作成プログラムの目視確認を実施．
- 計算方法や計算条件に誤りが無いことを確認．

⑤異常値&理論的説明性の確認

- 計算結果の時系列図や分布図・各種統計値等による異常値や理論的説明性を確認．



図．理論的説明性の確認（西村ほか(2025)[3]より抜粋）

⑥作成データの数量の確認

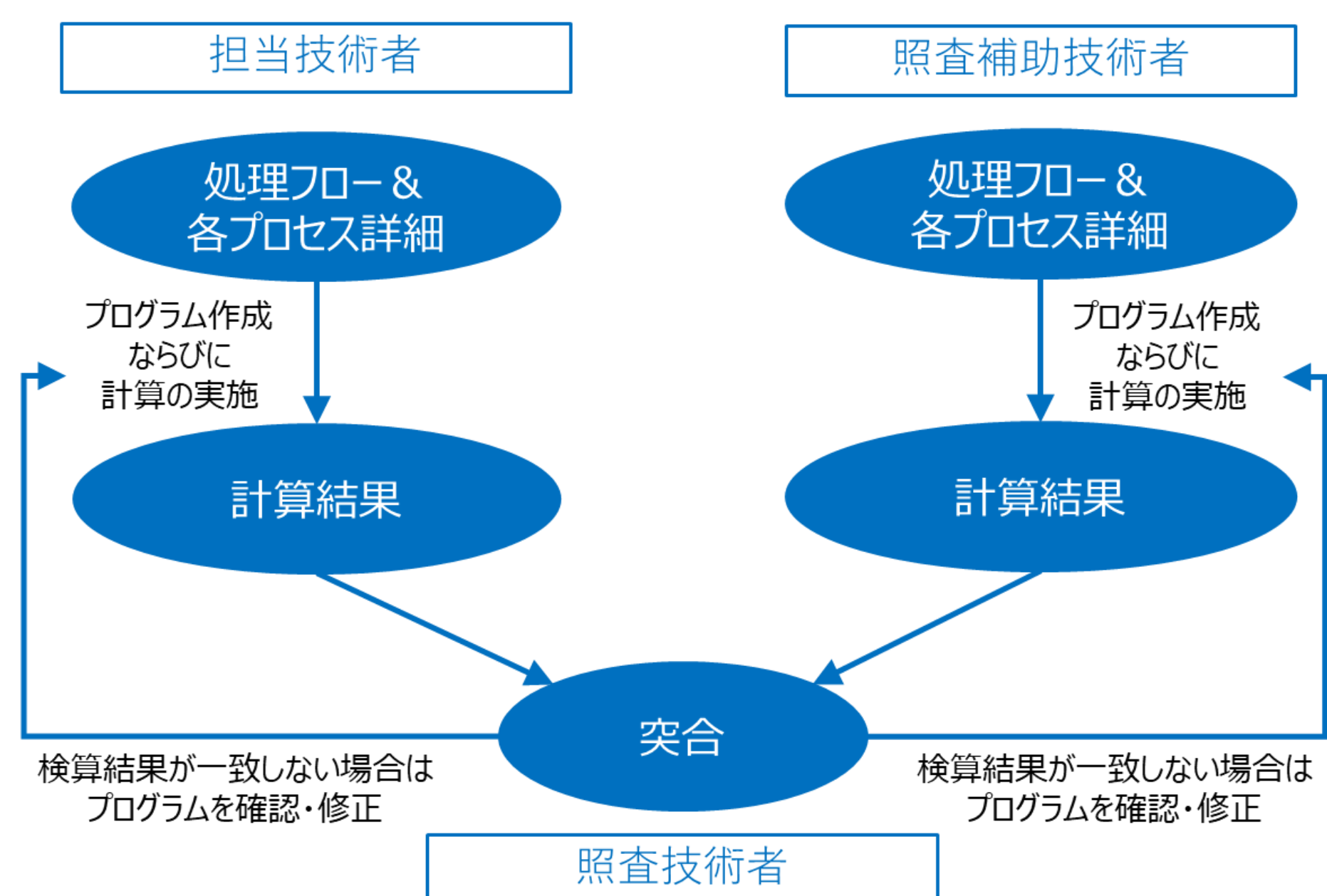
- 出力データの数量が膨大であるため，数量を確認するためのチェックプログラムを作成し，所定のデータが作成されていることを確認．

②データ処理フローの確認

- 各担当者が作成した作業フローと特記仕様書および打合せ簿の内容を照らしあわせ、処理内容に漏れが無い事を照査技術者が確認．

④計算結果の検算

- 照査補助技術者による計算結果の検算を実施．
- 全地上観測地点を対象とするのではなく，抽出的に行うことで照査の効率化を図った．
- 地点の抽出は，日降水量の多寡，日平均気温の季節変化傾向等，様々な条件下における結果を検算できるように，地域の偏りが無いように抽出した（降水量は75地点，気温は58地点）．
- 検算結果が一致しない場合は，計算結果が一致するまでプログラムの確認・修正した．



図．検算フロー

●検算で見つかったミスの例

累積相対度数分布の移動平均値算出範囲の計算ミスが見つかった。例えば、累積相対度数の移動平均区間が“0.9 94575045以上1.0未満”（地上観測の550位～552位に相当）であるとき、単精度実数で平均区間を計算していたため、上限値が1.00000000007となり、地上観測の553位が含まれていた。

データ納品

引用文献：

- [1] Kawase et al., 2023, J. Geophys. Res. Atmos., 128.
- [2] Watanabe et al., 2020, Hydrological Research Letters, 14(3), p117-122.
- [3] 西村ほか, 2025, 河川技術論文集, 第31巻, pp.421-426