

2019年度博士学位論文

定量的・定性的情報に基づく  
局地的降雨情報の推定に関する研究

お茶の水女子大学大学院  
人間文化創成科学研究科 理学専攻

佐野 仁美

2019年 9月



---

# 概要

日本は陸地の70%が山間地であり、地形的に台風や豪雨による影響を受けやすい。このため、国内には降雨に対し脆弱と考えられる地域が多数存在している。これまでの国内における降雨事象は均一的な降雨が広範囲に発生する例が多く、極端な大雨による水災害の事例は主に台風の影響によるものであった。しかし、近年は季節を問わず気象状況が短時間で急激に変化し局地的大雨となる事象が増え、これに伴う水災害も甚大化する傾向にある。この傾向は今後も長期的に続く見込みと考えられ、局地的大雨に伴う水災害への対策は重要な課題と言える。

こうした水災害から身を守るためには、降雨が小規模である段階から避難を開始することが有効である。しかし、局地的大雨は極めて狭域の地点に突発的に発生し、市町村単位で発表された直近の天気予報では当該地点の高いリスク状況を迅速に判断することが出来ない。従って、水災害による被害軽減には、自治体や住民らが現在地周辺の状況を主体的かつ早期に把握して今後の傾向を経験的に推定し、適切な行動判断を早期に開始することが最も有用な対策の一つと考えられる。そのためには、現在の局地的な降雨状況を誰もがリアルタイムで取得できる環境の整備が必要と考える。本研究では、「局地的な降雨情報の迅速な情報共有」を目的とし、次の3つのテーマで研究を行った。

まず初めに、気象観測レーダの一つである「XバンドMPレーダ」による観測情報を可視化し、リアルタイムで同レーダ情報を閲覧可能とする環境を構築した。XバンドMPレーダは国土交通省が全国に設置しているレーダであり、国内最狭レベルの高分解能および高頻度での配信を実現した。従って、同レーダによる観測情報は現時点においては国内で最も有用な局地的降雨の観測情報と言える。本研究では、局地的な降雨情報の取得および共有を目的とし、地球環境データ関連では国内で最大レベルのストレージを有する「データ統合・解析システム (DIAS)」上に同観測情報をリアルタイムで取得し画像公開する環境構築を行った。しかし、この開発にあたっては大別して3つの課題が生じた。具体的には、取得時の処理性能、重複データの混在、および表示性能の課題である。まず、取得時の処理性能の課題では、XバンドMPレーダデータの取得、画像作成処理、および公開環境へのデータ転送に関する全ての処理を同レーダ配信間隔である1分以内に終えなければならない制約があった。一方、データ量は従来のレーダと比べ1日当たり約1000倍に増加していた。このため本研究では、処理内容に応じて実行機器を振分けて物理的な負荷分散を図り、処理を並列化し、

---

---

処理状況の監視をしながら割込処理を組み合わせて遅延制御を行うことにより、総処理時間を平均 40 秒とする大幅な性能改善を実現した。次に、重複データ混在の課題では、国土交通省から配信される地点単位のデータに混在する重複データについて、DIAS 側で値を評価しながら解消する「高リスク優先手法」を提案し解決を図った。最後に、表示性能の課題では、携帯端末を利用した閲覧表示の性能改善に対処した。具体的には、低性能の要因であった大容量画像の解像度を表示時にユーザが指定する倍率に応じて変更する「高値採択型圧縮手法」を提案し、最終的に同性能を 3 秒以内に改善することができた。このように構築した X バンド MP レーダデータの共有環境は、2015 年から一般公開している。

次に、X バンド MP レーダデータに別の気象観測情報を合成して新たな降雨情報を生成する手法を提案した。X バンド MP レーダは短波ゆえに遮蔽物等により電波が減衰または消散する場合がある。このため、同レーダの設置場所は限定され、国内には観測範囲外の地点が多く存在する。従って、本研究ではこれらの地点の観測を補完することを目的として、主に広域的範囲において安定的な観測を行う C バンドレーダの観測データを合成し、国内全域の局地的な降雨情報を取得する新たな手法を提案した。

最後に、局地的な降雨情報を投稿情報内の言語表現から推定する手法を提案した。物理的な観測機器は設置場所に制約があり、これに伴って観測範囲も限定される。一方、近年では投稿情報の集合知からさまざまな情報を得る試みが広がっており、投稿情報は人を「動く観測センサ」とみなして現況を把握する情報源として有用視されつつある。気象情報に関する投稿情報の利用では、現状、晴れ・曇り・雨等の大きな区分での捉え方が中心であるが、これを降雨の程度（降雨強度）を取得する手法に発展できれば、物理情報の設置場所によらずに局地的な降雨情報を得る新たな手段を獲得できる。そこで本研究では、代表的な SNS 情報の一つである Twitter を利用し、発信情報に含まれる言語表現から局地的な降雨強度の推定を行う新たな手法を提案した。本提案手法では、まず、過去の膨大な Twitter の発信情報と X バンド MP レーダデータを解析して、言語表現と降雨強度の数値的な相関を捉えた。次に、特定の言語表現毎に月別の降雨強度のレベルを推定するモデルを構築した。そして、本モデルを実際の Twitter 情報に適用した結果、本手法により得られた推定レベル値は X バンド MP レーダデータの観測値（レベル換算）と極めて高精度で一致し、本提案手法が言語表現から物理情報相当値を推定する手法の一つとして有用であることを示した。本稿ではこれらの研究成果について順に述べる。

---

---

# Abstract

Because 70% of land in Japan is mountainous, typhoons and heavy rains have strong impact. Japan has many areas that are vulnerable to rain. Thus, uniform rain was common in the past few years aside from typhoons. However, in recent years, because of the phenomenon of sudden change in weather, localized heavy rain falls frequency, and water disasters are starting to become huge. Since this trend will probably continue for a long time, the measure for water disaster caused by local heavy rain is urgently needed.

In order to save lives from water disasters, it is must to start evacuation while the rainfall is not too heavy yet. Localized heavy rains occur randomly in quite small areas. However, weather forecasts are predicted in each unit of municipalities which are too wide to represent conditions in small areas. Therefore, it's difficult to grasp the high risk in small areas by current weather forecasts. One of the effective evacuation measures is that municipalities and residents should grasp the rainfall conditions around their current location proactively, make an appropriate action and judgment through an estimation and comparison of past experiences. Thus, the most important method is that "Grasping localized rainfall on real time". The main contents in this study are following 3:

First, the infrastructure sharing the data of X-band MP radar that realized the highest resolution in Japan and frequent data distribution was developed. The data volume of X-band MP radar is larger than the conventional radar, therefore, there are performance issues. These were solved and the infrastructure was published since 2015. Second, the method for data fusion of X-band MP radar and C-band radar was proposed. Since the waves of X-band MP radar are short, the electric waves of X-band MP radar may be attenuated or lost by obstacles. Therefore, there are many areas outside the observation range of X-band MP radar in Japan. In this study, as the purpose of complementing data at the areas, the new method for the data fusion with 2 radars was proposed for grasping localized rain of entire Japan. Finally, the method that estimates the localized rain using SNS information was proposed. Although the physical observation devices have limited installation location, I regarded people as "moving observers" and proposed a method to estimate rainfall information from the linguistic expression of Twitter which is one of the representatives of SNS. As a result, the estimated value by this method matched the observation value of X-band MP radar with high precision and proved the usefulness of this proposed methods.

---

---

# 目次

<b>第1章.</b>	<b>序論.....</b>	<b>9</b>
1.1	研究背景.....	9
1.2	研究目的.....	10
1.3	本論文の構成.....	11
<b>第2章.</b>	<b>局地的気象観測データの可視化.....</b>	<b>12</b>
2.1	研究背景.....	12
2.2	国内の降雨関連気象観測データの代表例.....	14
2.2.1	直接観測による気象観測データ.....	14
2.2.2	遠隔観測による気象観測データ.....	15
2.3	XバンドMPレーダの特徴.....	16
2.4	処理概要.....	19
2.4.1	全体処理概要.....	19
2.4.2	データ取得.....	20
2.4.3	画像生成および閲覧利用.....	22
2.5	研究課題および解決手法.....	23
2.5.1	並列化および割込処理による遅延解消.....	23
2.5.2	高リスク優先手法による重複データの排除.....	26
2.5.3	高値採択型圧縮手法による表示速度の高速化.....	30
2.6	DIASにおける画像閲覧機能.....	33
2.7	第2章のまとめ.....	35
<b>第3章.</b>	<b>気象観測データの合成による降雨情報の推定手法.....</b>	<b>36</b>
3.1	研究背景.....	36
3.2	研究課題.....	37
3.3	レーダ情報合成に関する提案手法.....	39
3.3.1	合成タイミングの同期.....	39
3.3.2	空間分解能の同期.....	41
3.3.3	同一地点におけるデータ値の採択手法.....	41

---

---

3.4	実装.....	43
3.4.1	全体処理概要.....	43
3.4.2	合成処理仕様.....	45
3.5	実装結果および考察.....	47
3.6	第 3 章のまとめ.....	49
<b>第4章.</b>	<b>言語表現を用いた降雨強度レベル推定手法.....</b>	<b>50</b>
4.1	研究概要.....	50
4.2	関連研究.....	52
4.3	対象 Tweet の特定.....	53
4.3.1	発信地点の特定.....	54
4.3.2	有効 Tweet 特定の処理概要.....	55
4.3.3	推定言語表現の特定方法.....	56
4.3.4	時点の絞り込み.....	58
4.3.5	有効 Tweet の精度に関する検証.....	59
4.4	降雨強度推定モデルの構築.....	61
4.4.1	Tweet が示す対象降雨の絞り込み.....	61
4.4.2	推定言語表現と X バンド MP レーダデータの照合.....	62
4.4.3	モデルの要素.....	66
4.4.4	推定値設定パターンの決定.....	67
4.4.5	Tweet 件数における人口密度の影響.....	68
4.4.6	降雨強度推定モデル構築手法のまとめ.....	70
4.5	適用結果および考察.....	70
4.5.1	2018 年の事例における適用結果.....	70
4.5.2	モデルの構成要素有無に関する考察.....	71
4.5.3	精度に関する考察.....	73
4.6	第 4 章のまとめ.....	74
<b>第5章.</b>	<b>本研究のまとめ.....</b>	<b>76</b>
	謝辞.....	77
	参考文献.....	78

---

---

## 図目次

図 1 : DIAS で扱われるデータ分野の概要.....	13
図 2 : X バンド MP レーダのアンテナ配置図.....	18
図 3 : X バンド MP レーダの観測地域.....	18
図 4 : DIAS における全体処理概要.....	19
図 5 : X バンド MP レーダデータのフォーマット.....	20
図 6 : ブロックデータのフォーマット.....	21
図 7 : 各サーバにおける処理イメージ.....	24
図 8 : データ取得および可視化処理の実行イメージ.....	25
図 9 : 富山・石川地域の地域画像および拡大画像(2015/09/09 03:57 JST).....	27
図 10 : 近畿地域の地域画像および拡大画像(2015/09/09 03:57 JST).....	27
図 11 : 豪雨周辺の主要レーダ設置イメージ.....	28
図 12 : 高リスク優先手法による全国合成(2015/09/09 03:57 JST).....	29
図 13 : 高値採択型圧縮手法のイメージ.....	31
図 14 : 全国合成画像の比較 (ズームレベル 5).....	32
図 15 : 「AMeNOW!」のサンプル画面.....	33
図 16 : 研究者向け閲覧画面.....	34
図 17 : Google Map へのオーバーレイ.....	34
図 18 : GrADS の利用による画像.....	34
図 19 : C バンドレーダによる観測情報 (2017/04/17 19:00 JST).....	37
図 20 : X バンド MP レーダによる観測情報 (2017/04/17 19:00 JST).....	37
図 21 : C バンドレーダの配信タイミングでの合成 (案 1).....	39
図 22 : X バンド MP レーダの配信タイミングでの合成 (案 2).....	39
図 23 : レーダ情報合成の処理タイミングイメージ.....	40
図 24 : 空間分解能の同期イメージ.....	41
図 25 : 案 A および案 B によるデータ合成イメージ.....	42
図 26 : 両レーダデータの合成イメージ.....	43
図 27 : レーダ情報の合成に関する全体処理概要.....	44
図 28 : C バンドレーダデータの配信イメージ.....	45
図 29 : X バンド MP レーダデータの配信イメージ.....	45

---

---

図 30 : 2次元配列への設定イメージ.....	46
図 31 : 実装結果(2017/04/17 19:00 JST).....	47
図 32 : 案 A 適用後の実装結果.....	48
図 33 : 案 B 適用後の実装結果.....	48
図 34 : 推定位置情報付加の概要.....	54
図 35 : 有効 Tweet 抽出の処理過程.....	55
図 36 : Tweet 発信時刻と降雨ピークの時間差.....	62
図 37 : 遡及時間内の X バンド MP レーダのレベル値.....	63
図 38 : 降雨強度レベルの推定言語表現別・月別集計(2015 年 8 月).....	64

---

## 表目次

表 1 : 観測手法で分類した降雨関連の気象観測データの代表例.....	14
表 2 : C バンドレーダと X バンド MP レーダの比較.....	16
表 3 : レーダ値の比較結果別領域数および割合.....	48
表 4 : 降雨強度レベル.....	51
表 5 : 有効 Tweet の位置情報精度 (2018 年 7 月) .....	59
表 6 : 有効 Tweet の内容に関する精度 (2018 年 7~9 月) .....	60
表 7 : 降雨強度レベル値の年間傾向 (2015~2017 年) .....	65
表 8 : 推定値設定パターン.....	66
表 9 : 推定言語表現の分類.....	67
表 10 : 一致率の年間平均 (2017 年) .....	68
表 11 : Category 毎の推定値設定パターン.....	68
表 12 : 人口密度に応じた按分件数.....	69
表 13 : 本モデルの月別適用結果 (2018 年) .....	71
表 14 : 一致率年間平均のパターン別比較 (2017 年) .....	72
表 15 : モデルの条件変更による一致率の比較 (2018 年) .....	72

## 第1章. 序論

### 1.1 研究背景

近年、気象状況が短時間で急激に変化し局地的大雨をもたらす事例が増加している。局地的大雨は季節問わず発生しており、特に「線状降水帯」<sup>(注1)</sup>を形成した場合には、地形的脆弱性が大きい地域に甚大な被害をもたらしている。2018年7月に発生した「平成30年7月豪雨」(通称「西日本豪雨」)では、西日本を中心とした広範囲の地域で記録的な大雨をもたらし、死者が220人を超える大災害となった。当事象に際し気象庁は大雨特別警報<sup>(注2)</sup>を11府県に発令している。しかし、発令前に河川が氾濫し土砂崩れが起きていた地域も存在していた。現在、警報は主に市町村単位で発表されているが、河川や山間部など降雨に脆弱な地域は各県に点在し、それぞれの特性によりリスクの度合いは異なる。このため、これまでの市町村単位による平均的なリスク評価では降雨に脆弱な地域への備えが遅れる可能性が高い。加えて、局地的な降雨事象は発生時刻も実際の降水量も狭い範囲毎に大きく異なり、必要な対策も千差万別である。従って、水災害の被害軽減対策としては、これまでの市町村レベルではなく、より狭域な範囲でのリスク判断を速やかに行うことが重要であると言える。

また、水災害における住民の意識改革も課題である。2015年9月に発生した「関東・東北豪雨」は栃木県を中心に長時間にわたる大雨をもたらし、栃木県内の各観測所では観測史上最多雨量を記録した。この結果、増水した鬼怒川の堤防が決壊し、浸水域では4000人を超える住民が孤立した。この例では鬼怒川中・下流域を管理する国土交通省下館河川事務所が水戸地方気象台・茨城県・栃木県と共同で鬼怒川の氾濫の危険性を段階的に伝える指定河川洪水予報を発表し、常総市からは避難指示や避難勧告が出されている。しかし、NHK放送文化研究所の世論調査<sup>(注3)</sup>によると、自宅などに留まる「屋内安全確保」を選択した住民が38%おり、そのうち約半数以上が孤立したという。また、上記で述べた西日本豪雨の事例でも自宅にとどまり逃げ遅れた人は多く、民間気象情報サービスのウェザーニューズのアンケート<sup>(注4)</sup>によれば、避難すべき状況にも拘わらず実際に避難しなかった人は84%にのぼっている。

---

(注1) 強い雨を降らせる積乱雲が帯状に連なる状況

(注2) 気象庁から「台風や集中豪雨により数十年に一度の降雨量となる大雨が予想される場合、若しくは、数十年に一度の強度の台風や同程度の温帯低気圧により大雨になると予想される場合」に発表される警報。

[https://www.jma.go.jp/jma/kishou/now/bosai/warning\\_kind.html](https://www.jma.go.jp/jma/kishou/now/bosai/warning_kind.html)

(注3) 避難指示・避難勧告の対象となった地域に居住していた20歳以上の男女1,000人を対象にした面接法による世論調査。

(注4) 減災調査2018 <https://weathernews.jp/s/topics/201808/300095/>.

これについて、自然災害リスク対策の行動の難しさに関する研究[1]では、国民による自然災害への意識は高まってきているが実際の出現行動は向上しにくいいためリスクに対し自力での対策が可能という感覚を抱かせることが重要、と述べる。つまり、住民自らが現時点の状況把握とパターン習得による経験を積み重ね、適切な行動判断が行える環境構築が必要であると考えられる。従って、降雨のリアルタイムでの状況を誰もが把握でき、地点毎の脆弱性や降雨の大きさに応じた被害の程度等を経験的に会得できることが望ましいと言える。特に、「線状降水帯」は同じ地点に強い雨を降らせる積乱雲が帯状に連なる特殊な形状で判別もつきやすい。線状降水帯は当該地域に最も深刻な被害をもたらす可能性が高いことから、リアルタイムで把握した降雨形状から避難準備等の対策を早期に講じることにより、被害を軽減化させることが可能と考える。局地的大雨は日本の各地に洪水、崖崩れ、土石流等の甚大な水災害をもたらす可能性が高く、その進行や状況は予測できない例が多い。しかし、現況を迅速に捉え共有することにより住民自らがその後の行動判断に繋げることが可能となり、被害軽減に極めて有用な対策の一つになると考えられる。

以上により、本研究では降雨に関する局地的なリスク評価および自治体や住民等の経験と学習による意識向上を企図し、「局地的な降雨情報のリアルタイムでの取得および情報共有を可能とする環境の構築」を目指すこととした。

## 1.2 研究目的

本研究では局地的な降雨情報の迅速な情報共有を目的とし、これを可能とする環境基盤の構築および局地的な降雨情報を推定する手法の提案を行う。このため、まずは局地的な降雨情報として現時点で最も有用な X バンド MP レーダによる観測情報を共有する基盤構築を行い、更に、当該情報を補完し得る新たな降雨情報（推定値）を他の気象観測情報や投稿情報の利用により生成する手法を提案した。本研究は以下の3つに大別して構成される。

第1は、国内における代表的な局地的な降雨観測データの共有を目的とした研究（「局地的気象観測データの可視化」）である。国土交通省が設置した X バンド MP レーダは従来のレーダに比べ 16 倍の高い空間分解能と高頻度（1 分）での配信を実現し、最も有用な局地的な降雨の観測情報である。しかし当該データ量は膨大で、全観測データを観測元の国土交通省で蓄積する環境構築は困難であった。

---

このため、国内で最大規模の容量を持つ地球環境情報プラットフォームである「データ統合・解析システム」(DIAS<sup>(注5)</sup>)上で取得・蓄積する開発が行われた。本研究は、これらの蓄積データを遅延なく可視化処理し、情報公開する環境構築を行った。

第2は、XバンドMPレーダデータの情報補完を目的とした研究(「気象観測データの合成による降雨情報推定手法の提案」)である。XバンドMPレーダは短波ゆえに設置場所が限られ、国内には同レーダで観測不可能な地域が多く存在する。このため、災害時に他の気象情報を合わせ見る必要があるなど機動的対処に課題があった。本研究ではこの課題解決として、XバンドMPレーダに従来のレーダであるCバンドレーダデータを合成することにより国内全域の局地的な降雨情報を推定する新たな手法を提案した。

第3は、局地的な降雨情報を物理情報以外から取得することを企図とした研究(「言語表現を用いた降雨強度<sup>(注6)</sup>レベル推定手法」)である。物理情報は観測値の正確性は高いが、設置場所が限定される。一方、投稿情報を利用する場合、人を「動く観測センサ」とみなすことで人の移動範囲(発信地点)を観測地点とし、迅速に情報を得ることが可能となる。従って、物理的な観測器で取得不可能な情報を人の発信情報で補完することが期待できる。このため、本研究では投稿情報とこれに対応する過去のXバンドMPレーダとの相関を詳細に検証し、言語表現から物理情報の実測値に近似する降雨情報を推定する新たな手法を提案した。

### 1.3 本論文の構成

本論文の構成は次のとおりである。まず、第2章では「局地的気象観測データの可視化」として、XバンドMPレーダのデータを取得、蓄積、可視化およびデータ公開を可能とする基盤構築の方法を述べる。第3章では「気象観測データの合成による降雨情報推定手法」として、XバンドMPレーダデータとCバンドレーダデータの合成に関する提案手法を述べる。第4章では「言語表現を用いた降雨強度レベル推定手法」として、投稿情報のうち特にSNSの代表であるTwitterから降雨関連情報を抽出し実際の降雨レベルを推定する提案手法を述べる。最後に第5章で本研究を総括する。

---

(注5)「データ統合・解析システム」(DIAS: Data Integration and Analysis System) 文部科学省から委託を受け東京大学地球観測データ統融合連携研究機構(EDITORIA)が開発・運用しているシステム (<https://www.diasjp.net/>)

(注6) 降雨の1時間当たりの雨量

## 第 2 章.

# 局地的気象観測データの可視化

本研究では、局地的な気象観測データの一つである X バンド MP レーダの全データをリアルタイムで可視化し、降雨の現況を把握可能とする情報提供環境を構築した。X バンド MP レーダによる観測データは高解像度・高頻度配信であるゆえにデータ量が大きく、処理時間や表示時間の最適化が必要であった。これに加えて、国土交通省の配信データには重複データが混在する課題も生じていた。このため、本研究では、これらの課題解決を図り、迅速かつ安定的に情報提供し得る環境構築を目指した。本章では、X バンド MP レーダの可視化に際して生じた課題と解決手法を中心に、同レーダの特徴、開発基盤である DIAS の概要、データ取得概要および実装結果を述べる。

## 2.1 研究背景

増加傾向にある局地的大雨の事象に対処するため、これまでの広域的な降雨情報以外に局地的な降雨情報を取得する必要があった。この手段として国土交通省が全国に新たに設置したのが X バンド MP レーダである。X バンド MP レーダは従来のレーダである C バンドレーダに比べ 16 倍の高分解能での観測と高頻度での配信 (1 分間隔) を実現したマルチパラメータレーダである。これにより、X バンド MP レーダによる観測データは国内全ての気象観測情報の中で局地的な降雨の観測に最も有用なデータの一つとなった。同レーダは大都市、火山、災害復旧現場など豪雨に対して脆弱性を持つ観測地域に優先的に配置され、現在は全ての政令指定都市に設置されている。しかし、同レーダのデータ量は、高解像度・高頻度の配信ゆえに大きく膨らみ、従来の主要レーダであった C バンドレーダに比べ、1 日当たり約 1000 倍のデータ量に増加した。このため、観測元である国土交通省では過去に亘る全てのデータを蓄積する環境構築が困難となっていた。このため、国土交通省で観測された全ての X バンド MP レーダデータは DIAS 上で取得および蓄積する開発が行われることとなった。

---

第2章  
局地的気象観測データの可視化

DIAS は地球環境関連データを対象とする大規模データベースシステムである。図1に示すように、「地球規模・各地域の観測で得られたデータを収集、永続的な蓄積、統合、解析するとともに、社会経済情報などとの融合を行い、地球規模の環境問題や大規模自然災害等の脅威に対する危機管理に有益な情報へ変換し、国内外に提供することにより、我が国の総合的な安全保障や国民の安全・安心の実現に資することを目的」(注7)として構築された。現時点の地球観測関連データのシステムでは国内最大規模のストレージを有し、特に降雨関連情報は国内の代表的な情報をほぼ網羅する。国土交通省で観測された全てのXバンドMPレーダデータはDIAS上で取得および蓄積する開発がなされ、2014年から運用が開始されている。

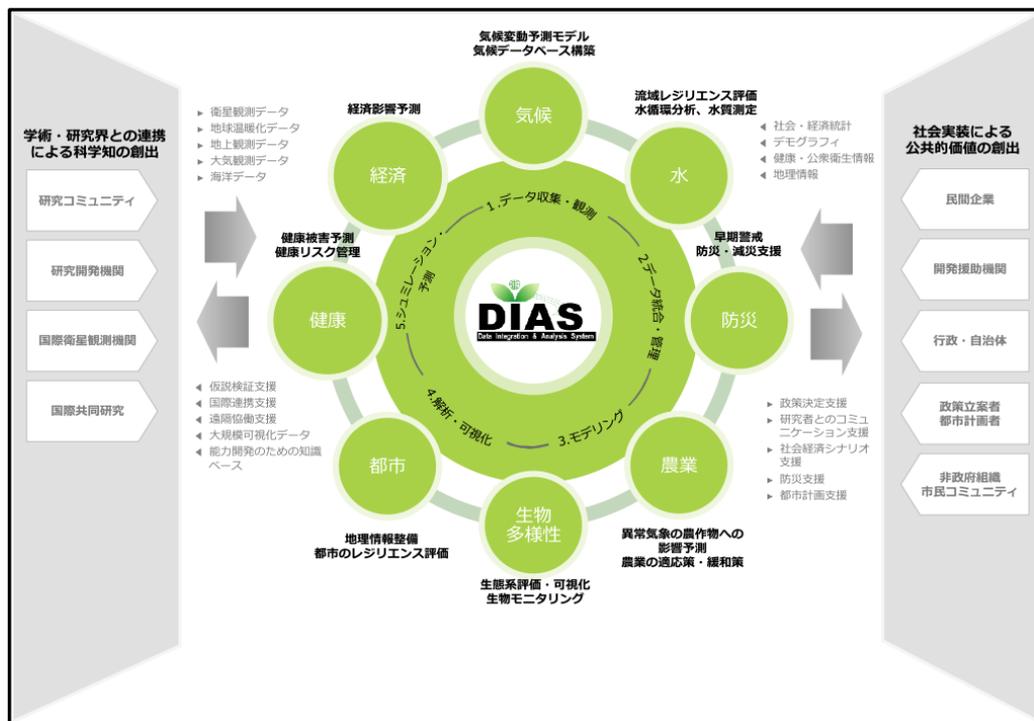


図1：DIASで扱われるデータ分野の概要(注8)

本研究では、XバンドMPレーダデータの迅速な情報共有を目的として、DIAS上に蓄積されたこれらのXバンドMPレーダデータを取得後即時に開示(閲覧)することを可能とする環境構築を行った。

(注7) データ統合・解析システム HP 「DIASとは」(<https://diasjp.net/about/>)の文中から引用。  
(注8) 同「DIASとは」(<https://diasjp.net/about/>)の図から引用。

## 2.2 国内の降雨関連気象観測データの代表例

本セクションでは国内の代表的な気象観測データ（降雨関連）を紹介する。

気象観測データは観測手法で大別すると、直接観測および遠隔観測（リモートセンシング）に分類できる。当該分類別に代表例を表1に示す。

表1：観測手法で分類した降雨関連の気象観測データの代表例

観測手法		代表例
直接観測		地域気象観測システム (通称：AMeDAS <sup>(注9)</sup> )
遠隔観測	地上リモートセンシング	XバンドMPレーダ Cバンドレーダ CバンドMPレーダ フェーズドアレイ気象レーダ
	衛星リモートセンシング	ひまわり

### 2.2.1 直接観測による気象観測データ

直接観測とは、ある場所の状態を実際にその場で測ることを言う。そして、直接観測を行う気象観測データの代表例としては気象庁の「地域気象観測システム」（通称「AMeDAS」）が挙げられる。AMeDASは国内約1300箇所に設置されている気象庁の無人観測所であり、地上における降雨の直接観測を行う。主に降水量、気温、日照時間、風向き・風速の気象要素を観測し、観測網の密度は降水量が約17kmのグリッド範囲、要素の観測点は約20kmのグリッド範囲で概ね20～200km規模（都府県レベル）の気象現象の把握に効果を発揮する。また、現時点のデータ配信は10分毎である。このように、直接観測による観測データは一般的に空間分解能が限定され配信頻度も高くないが、例えばCバンドレーダデータの補正に利用される等、高品質なデータとしての利用価値が高い。

(注9)：AMeDAS: Automated Meteorological Data Acquisition System

## 2.2.2 遠隔観測による気象観測データ

遠隔観測は、離れた場所から電磁波等により間接的に観測するものであり、地上に設置された測器による観測（地上リモートセンシング）と人工衛星搭載の測器による観測（衛星リモートセンシング）に大別される。現時点の地上リモートセンシングによる気象観測データの代表例では、XバンドMPレーダ、Cバンドレーダ、CバンドMPレーダおよびフェーズドアレイ気象レーダによる観測情報が現時点で最も利用価値が高い情報と考えられる（XバンドMPレーダおよびCバンドレーダの詳細は2.3に後述する）。

フェーズドアレイ気象レーダは、ゲリラ豪雨や竜巻などの観測を目的に、東芝、大阪大学および情報通信研究機構が開発した最短10秒間隔で隙間のない三次元降水分布を100mの分解能で観測可能な気象レーダとして開発された。2.3で後述するように、Cバンドレーダは日本全国を広域的に観測する一方、分解能が低く局地的な観測には向いていない。また、局地的な観測に向くXバンドMPレーダは地上付近の降雨を短時間で取得でき、降雨の現況取得に適するが、予兆を捉えるのは難しい。フェーズドアレイ気象レーダは予兆の把握を目的としており、上空の雲内における豪雨の前兆現象を捉え、3次元立体構造での観測を可能とする。パラボラアンテナを機械的に回転させるXバンドレーダやCバンドレーダと異なり、多数のアンテナ素子を配列し各素子における送信及び受信電波の位相を制御することで電子的なビーム方向を瞬間的かつ自由に変化させることが可能である。しかし、現時点では設置台数が極めて少なく、観測場所は大阪および埼玉等を中心とした地域に限定される。このため、フェーズドアレイ気象レーダは突発的な気象状況の監視や短時間予測への有用性が期待できるが、現状はXバンドMPレーダやCバンドレーダほど普及しておらず、広く実用化するには至っていない。

衛星リモートセンシングによる観測データの代表例では、ひまわり（現時点では8号および9号）が挙げられる。この衛星は世界最先端の観測能力を有する可視赤外放射計を搭載し、直接観測を行うことが困難な海洋や砂漠・山岳地帯を含む広い地域の雲や水蒸気の分布、海面の温度などを観測する。このため、本観測情報は主に雲の動きを捉えるものであり、降雨状況を直接表すものではない。しかし、他の降雨情報と合わせ見ることによって降雨全体の規模や状況把握が可能と考えられ、降雨予測等には有用なデータの一つであると言える。

以上のとおり、降雨関連の気象観測データは数多く存在する。しかし、国内全地点の局地的な降雨の状況を単独で把握可能な気象観測データは現時点では存在しない。従って、国内全地点の降雨状況を把握するには、複数の気象観測データを組み合わせ、各地点の状況を総合的に把握する必要がある。防災上の観点では局地的な降雨情報を捉える重要性が高い。このため、国内全地点の降雨状況を把握するには、まず X バンド MP レーダによる観測データを主軸とし、これに他の気象観測データを組み合わせることが現実的な方法であると考えられる。

## 2.3 X バンド MP レーダの特徴

本セクションでは、X バンド MP レーダの特徴について、従来のレーダである C バンドレーダを表 2 のとおり比較しながら述べる。

表 2 : C バンドレーダと X バンド MP レーダの比較

	X バンド MP レーダ	C バンドレーダ
観測間隔	1 分	5 分
空間分解能	250m メッシュ	1km メッシュ
定量観測領域	半径 60km	半径 120km
波長	3 cm	5 cm
配信単位	地域毎	全国
国内全域における 受信データ総計	約 20 MB / 1 分 (約 30 GB/日)	約 0.1 MB / 5 分 (約 0.03 GB/日)

まず、観測間隔について述べる。C バンドレーダは単偏波レーダであり、降雨強度を雨粒で反射した電波の強度から算出する。この算出に必要なパラメータは降雨の状態により異なるため、地上雨量データの AMeDAS を利用して値の補正を行う。従って、C バンドレーダの配信には当該補正時間が含まれ、現状の観測は 5 分間隔である。一方、X バンド MP レーダは 2 種類の偏波(水平・垂直)を送受信して雨粒の形状を把握し、雨滴の扁平度等から降雨強度を推定するマルチパラメータ方式であるた

---

め、高精度な観測を可能とし地上雨量計による補正を要しない。このため、観測間隔は約1分であり、Cバンドレーダの5倍の高頻度である。

次に、空間分解能を比較する。はじめに、空間分解能を表すのに用いられる「メッシュ」の用語について説明する。「メッシュ」は、緯度経度に基づき地域を隙間なく網の目の区域に分けたものをいい、昭和48年行政管理庁告示第143号の「標準地域メッシュ」を基にしている。総務省統計局を始め国の行政機関が作成する地域メッシュ統計と同様、気象観測情報も当該基準地域メッシュを利用している。基準地域メッシュは第1次地域区画を基に区画され、第1次地域区画は緯度を40分間隔、経度を1度間隔に区分した区画を示す。これを通称「1次メッシュ」と呼ぶ。これを縦横8等分した区画が第2次地域区画であり、通称「2次メッシュ」と呼ぶ。さらにこれを縦横に10等分した区画が基準地域メッシュ（第3次地域区画）となり、通称「標準基準メッシュ」または「3次メッシュ」と呼ぶ。Cバンドレーダの空間分解能はこの基準地域メッシュに該当し、その区画距離から1kmメッシュとも呼ばれる。一方、XバンドMPレーダの空間分解能は、基準地域メッシュ（第3次地域区画）を更に縦横4当分した単位（4分の1メッシュ）である。当該空間分解能はCバンドレーダの縦横4倍（16倍）の高分解能であり、これは現時点で国内最高レベルである。この点から、XバンドMPレーダは狭域での観測として国内で最も適した観測情報と言える。

観測領域については、XバンドMPレーダの場合が半径60kmであり、狭域な範囲の観測に適する。一方、Cバンドレーダの場合は半径120kmと広域であるため、台風の進行状況や降雨全体の状況把握に適していると考えられる。

次に、波長およびこれに伴うデータへの影響を述べる。各XバンドMPレーダは国土交通省において地形に合わせたレーダの仰角設定等が厳密に行われ、安定的な値の取得を確保している[2]。しかし、XバンドMPレーダは波長が短いことから、極めて強い降雨域が生じた場合に当該降雨がXバンド帯の電波の進行を阻み、電波が減衰・消散して観測が不能となる場合がある。この対策として国土交通省では各観測地域を複数のXバンドMPレーダで囲むように配置し、互いに取得したデータを補完しながら欠測リスクを低減している。つまり、ある地点の合成雨量を取得する際には、当該地点に近い複数のXバンドMPレーダの観測値を按分し求めることとなる。参考として、日本の関東地域におけるXバンドMPレーダの配置を図2に示す。一方、Cバンドレーダは、波長がXバンドMPレーダより長く、レーダ波の減衰によるデータ欠損は殆ど発生していない。

---

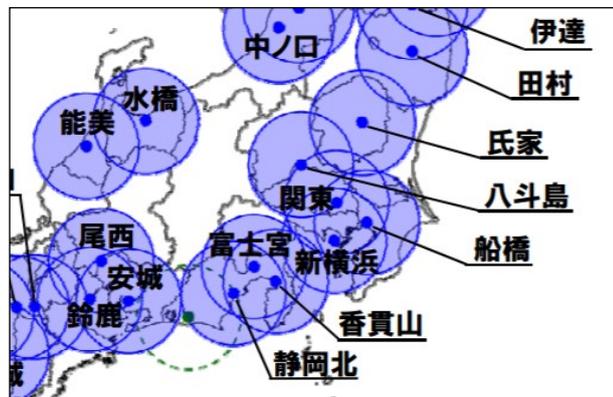


図2：XバンドMPレーダのアンテナ配置図(注10)

配信単位については、XバンドMPレーダの場合は全国を14分割した地域単位での配信である。しかし、当該地域は図3に示すように互いに重なり合い、同一地点に重複データが存在する課題がある。一方、Cバンドレーダの場合は全国一括で配信され、同一地点におけるデータ重複はない。

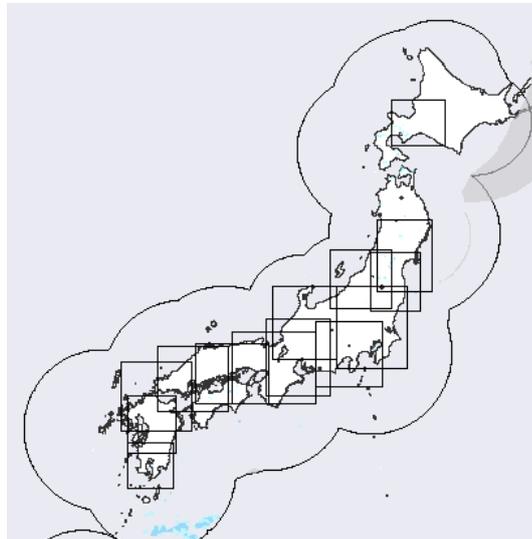


図3：XバンドMPレーダの観測地域[3]

データ量については、表2に示すとおり、XバンドMPレーダの1日当たりのデータ量はCバンドレーダデータのデータ量の約1000倍となっている。

(注10) 国土交通省資料 「XRAINの整備状況」より図を引用(2019年5月時点)：  
<http://www.mlit.go.jp/common/001046714.pdf>

## 2.4 処理概要

本セクションでは、国土交通省から X バンド MP レーダデータを取得し、DIAS 上で可視化された画像を生成するまでの全体処理概要を述べる。

### 2.4.1 全体処理概要

前述 2.3 のとおり、国土交通省が観測した X バンド MP レーダの全データは DIAS 上に蓄積される。当該蓄積データは①加工前の観測情報（生データ）、②一次処理後の加工値、③合成雨量の 3 種である。このうち、本研究で可視化処理に利用するデータは③合成雨量である。

国土交通省で観測、加工および合成処理が行われた X バンド MP レーダのデータ（①～③）は、DIAS に観測からほぼリアルタイムで配信される。その後、DIAS 上で当該データを利用して可視化画像を生成する。この処理概要を図 4 に示す。

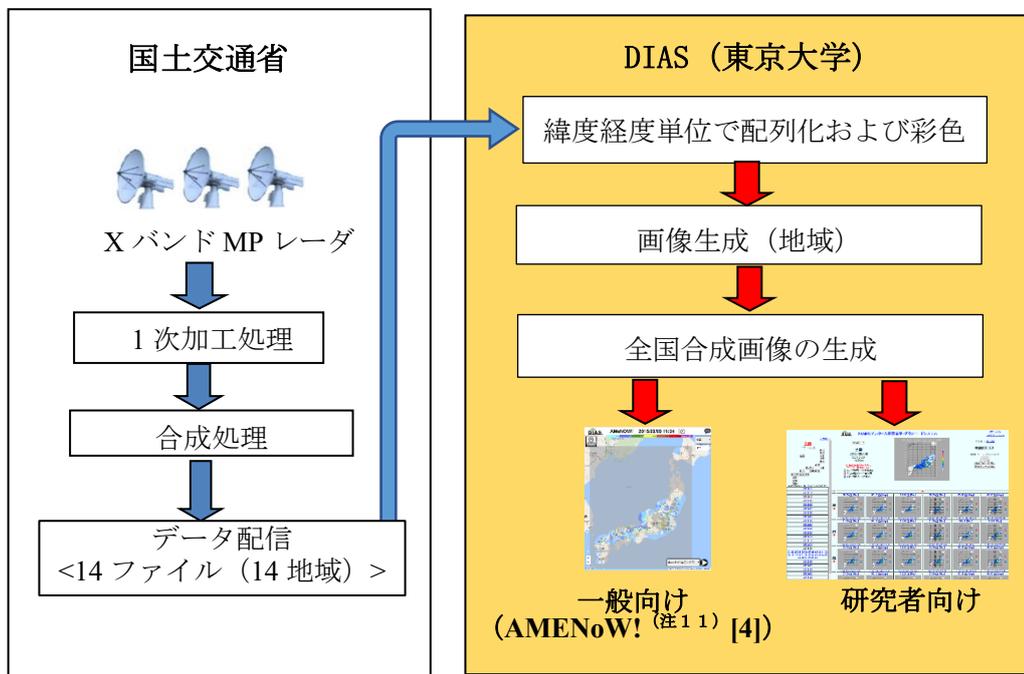


図 4 : DIAS における全体処理概要

(注 11) AMeNOW! (<http://rain.diasjp.net/>)

## 2.4.2 データ取得

XバンドMPレーダの全データ（前述①～③のデータ）は、国土交通省において観測、加工および合成処理後、リアルタイムでDIASに配信される。いずれのデータも14の地域単位（14ファイル）で毎分配信される。

DIASでは国土交通省から配信データを受信後ただちに処理を開始し、取得データのうち合成雨量を利用して、一般閲覧用および研究者閲覧用を含めた全ての画像生成処理（以下、これらを総称し「可視化処理」と呼ぶ。）を行う。

以降、本研究における可視化処理の概要を説明する。まず、各地域ファイル（14ファイル）を取得後、全地点の位置情報（緯度経度）および合成雨量を抽出する。XバンドMPレーダのデータは全国を14の地域に分割した上で、図5に示すフォーマットにより送信される。このうち、ブロックヘッダ4バイト内に当該データの位置を示す緯度、経度および2次メッシュコードが含まれている。

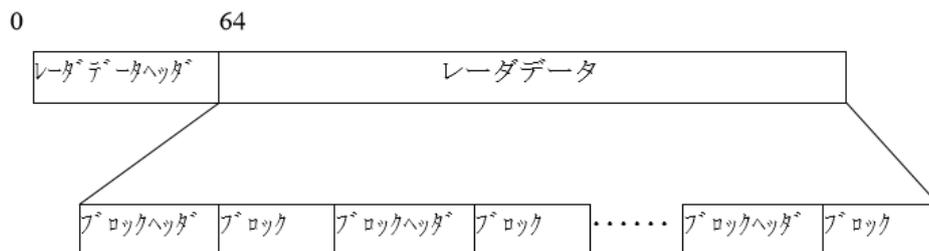


図5：XバンドMPレーダデータのフォーマット[3]

XバンドMPレーダのデータは標準地域メッシュを縦横4分割した4分の1地域メッシュである。ブロックデータ内の4分の1メッシュデータをブロックヘッダの緯度経度に基づくXY座標による2次元配列に展開する。各ブロックデータ内におけるXY座標の位置を図6に示す。

---

第2章  
局地的気象観測データの可視化

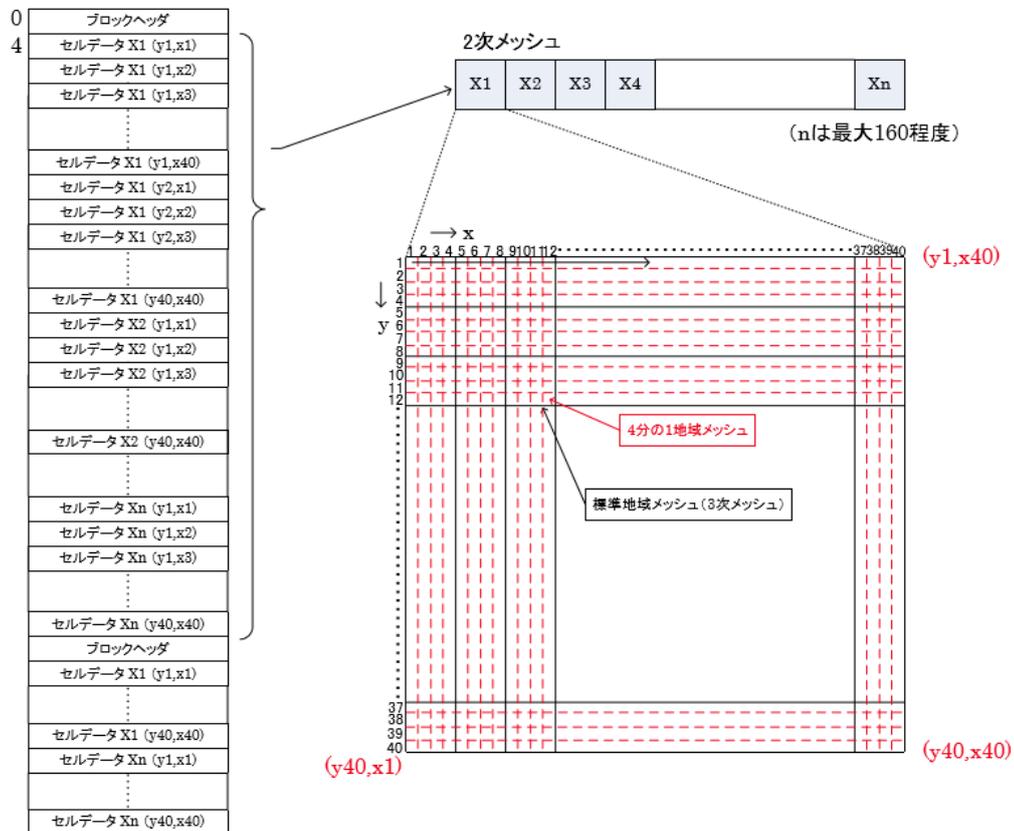


図6：ブロックデータのフォーマット[3]

各地域におけるブロックデータの X 座標および Y 座標が 0 である地点は、ブロックヘッダに含まれる 2 次メッシュコードの位置に該当する。従って、全国合成を行う場合には、国内全体の領域における各地域の XY 座標のスタート位置を各地域の 2 次メッシュコードから以下の式により求める。なお、1 つの 2 次メッシュ内における経度方向および緯度方向のデータ数はそれぞれ 320 (8 領域 \* 10 分割 \* 4 分割) である。

X 座標のスタート位置：(地域の最西端経度－全国領域の最西端経度) \* 320

Y 座標のスタート位置：(全国の最北端緯度－地域の最北端緯度) \* 320

### 2.4.3 画像生成および閲覧利用

2次元配列に展開した各地域のデータは合成雨量の値に基づき彩色し、当該彩色データ（3原色の構成値）をファイル出力する。そして、全ての地域データの処理終了後、全国画像の合成を行う。当該合成処理では日本を覆う全国の地点をXY座標による2次元配列に展開する。地域毎に重なり合う地点については後述2.5.2で述べる手法により値の評価を行い、最終的に全地点の値を決定する。そして、地域データと同様に値に応じた彩色設定を行い、彩色データ（3原色の構成値）をファイルに出力する。

続いて、これらの出力ファイルを利用し、14地域および全国合成の画像を生成する。画像に重ね合わせる地図はGoogle Map および緯線経線のマトリクスの2種とした。Google Map にオーバーレイする画像は彩色データ（3原色の構成値）をConvertして生成する。緯線経線のマトリクスとして生成する画像は地球科学関連のデータを処理・図化するための対話式ソフトであるGrADS<sup>(注12)</sup>を利用する。そして、これらの生成画像は次のXバンドMPレーダの配信時間（1分後）が到来する前に即時公開する。公開の方法は次の2通りとした。

- ① 一般用閲覧用
- ② 研究者向け閲覧用

一般用の閲覧機能は「リアルタイム降雨情報 (AMeNOW!)」として、認証を必要とせず誰もがリアルタイムでパーソナルコンピュータや携帯電話等を通して利用可能としている。ここで閲覧可能な画像はGoogle Map へのオーバーレイ画像である。本機能では地域と日付を指定することにより、指定日時における画像の閲覧や、指定時刻からの動画再生も可能としている。

研究者用の閲覧機能ではGoogle Map および緯線経線のマトリクス地図の両タイプの画像を閲覧できる。一般向けと同様に地域および日時指定により画像閲覧や動画再生が可能であるほか、データダウンロードも可能としている。

---

(注12) Grid Analysis and Display System (<http://cola.gmu.edu/grads/>)

## 2.5 研究課題および解決手法

本研究では X バンド MP レーダデータを可視化する際に生じた①処理遅延, ②重複データの混在および③表示性能の低さの3つの課題があった。このため, 当該課題に対する改善策として①並列処理と割込み処理の組み合わせた処理時間短縮, ②高リスク優先手法によるデータ重複解消および③表示画像の切り替えによる表示性能改善の解決手法を適用した。本セクションでは, これらの課題および解決手法を述べる。

### 2.5.1 並列化および割込処理による処理遅延解消

X バンド MP レーダの情報取得から可視化までの全処理を順次処理で実装すると, 総処理時間は3分を超過し大幅遅延となった。従って, 配信データの取得および後続の可視化処理は全て並列化して実装した。しかし, 並列化しても総処理時間が60秒を超え, 毎分配信される X バンド MP レーダの処理に遅延が生じた。本セクションではこれらの課題の要因および解決手法を述べる。

#### 2.5.1.1 処理遅延の要因

処理時間超過には2つの要因があった。まず1つ目の要因は, DIAS 上で頻繁に生じるユーザ処理要求によるシステム負荷であった。DIAS 上では, ユーザが X バンド MP レーダデータをリアルタイムでダウンロード可能とする環境を併せて構築した。このため, X バンド MP レーダデータの取得・蓄積・加工の処理中に, ユーザからのデータダウンロードおよび画像閲覧に関する処理要求が随時発生し, システム全体の負荷は常時高い状況にあった。特に, 過去の1ヵ月分の X バンド MP レーダデータのダウンロード要求が発生した際は, 対象データが大量であったため, 可視化処理が30分以上遅延した。

次に2つ目の要因として, 国土交通省によるデータの配信遅延やファイル欠損が生じる事例が判明した。国土交通省からの X バンド MP レーダデータ配信では, 観測範囲の降雨域の場所や降雨の強さ等により, 遅延やファイル欠損が生じる場合がある。加えて, 地域単位での配信時刻や配信順序も, 降雨の状況に左右され定型的ではない。このため, 次の配信(1分後)までに

全ての地域データファイルが揃わない事象はしばしば発生していた。XバンドMPレーダデータの可視化処理では、全ての地域ファイルが揃った後に全国画像の合成処理を実行することとしている。従って、一部の地域のデータファイル配信に遅延が生じた場合は、その時刻の全国画像の生成が実行できず、処理が全体的に遅延する。

### 2.5.1.2 2つの課題の解決

まず、DIAS上におけるXバンドMPレーダデータの処理負荷に関する課題では、処理毎に物理的環境を分割する解決を図った。具体的には、国土交通省からの配信データ取得や可視化処理を実行する「データ取得および可視化専用サーバ」（以下、「サーバA」と呼ぶ。）と、データのダウンロード要求やデータ閲覧等のユーザ要求処理を実行する「ユーザ関連処理サーバ」（以下、「サーバB」と呼ぶ。）の2系統で処理を行う（図7）。

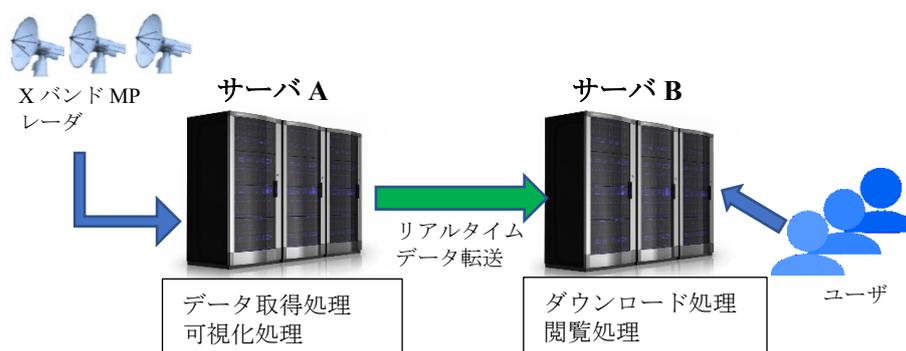


図7：各サーバにおける処理イメージ

処理対象サーバを分割したことで、ユーザによる処理要求の影響は解消された。しかし、この変更に伴い、サーバAからサーバBへのファイル転送時間が追加され総処理時間は増えたことになる。従って、サーバBでXバンドMPレーダのデータをリアルタイムで閲覧するには、サーバAにおけるデータ取得、可視化処理およびサーバAからサーバBへのファイル転送を次の配信時刻までに完了する必要があるが生じた。当初より更なる性能向上が必要となったが、これらは次に述べる配信遅延に関する課題の中で同時に解決を図ることができた。

国土交通省からの配信遅延およびファイルの欠損に関する課題では、以降に述べる処理方式に変更して解決を図った。

当初は「現時点の地域データの受信および可視化処理を並列処理し、全地域分処理後に全国合成を実行し次の時間帯の処理を行う」としていた。しかし、地域データに配信遅延や欠落が生じた場合、全地域分の処理終了後に全国合成を行うと当該処理の開始が遅延する可能性がある。このため、「現時点の処理を優先して間に合わない処理はスキップし、後続の時間帯以降の隙間時間に割込処理させる」方式に改めた。当該イメージを図8に示す。

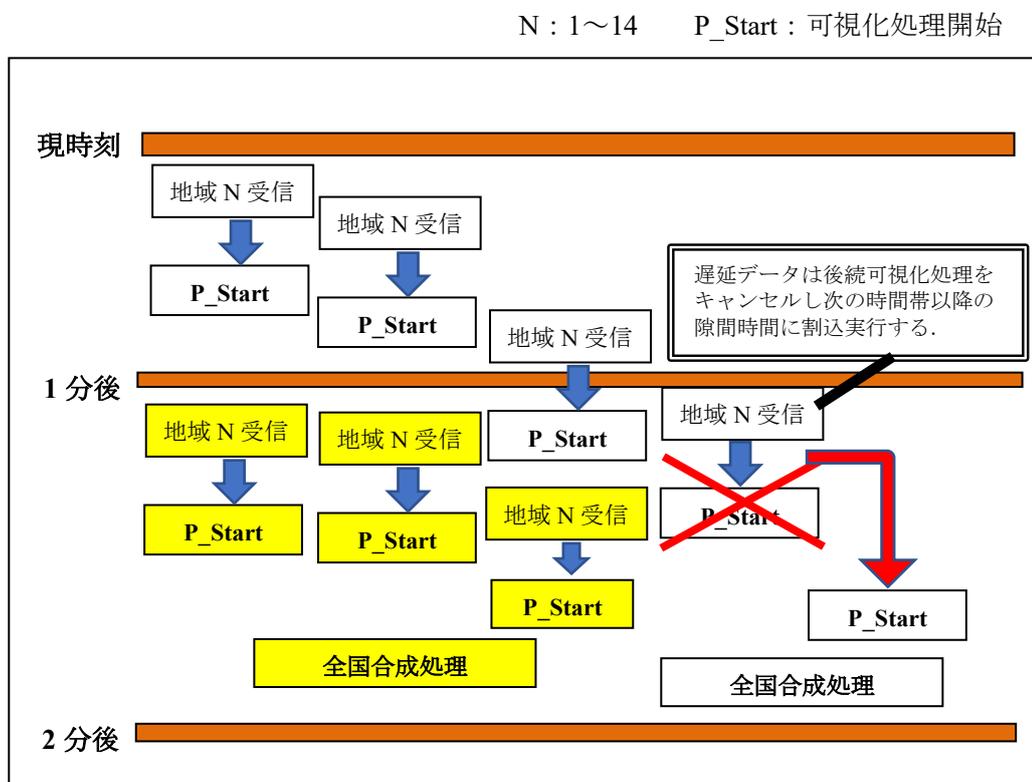


図8：データ取得および可視化処理の実行イメージ

このように、本来の時刻に遅延したデータは、次の時間帯以降の隙間時間に割り込ませる。遅延か否かは現在時刻とファイル名に付された時刻を比較して判断する。また、合成画像の生成は全地域画像が揃ったか否かをクロージングチェックし、揃った時点で開始させる仕様とした。

### 2.5.1.3 実装結果および考察

本提案手法による実装の結果，データ取得から可視化処理に関する全処理を平均 42 秒以内で完了することができた．次回配信までに約 20 秒程度の余力を持たせることが可能となり，処理時間の大幅な改善を図れた．併せて，ユーザ処理要求に関しても，データ取得または可視化処理の負荷の影響を受けず，現状は遅延なく稼働している．

## 2.5.2 高リスク優先手法による重複データの排除

国土交通省から配信される各地域ファイルには重複データの問題が生じていた．各地点の値は管轄する地域毎に観測および品質管理がなされ，どちらの地域の値が正であるかの決め手がない．しかし，本研究では配信された地域データを全国画像として合成するため，各地点の値を一意に決定する必要がある，解決手法の検討を行った．本セクションでは，当該重複データの要因および防災上の観点における解決手法を述べる．

### 2.5.2.1 重複データ混在の要因

前述 2.3 のとおり，国土交通省では地点を囲む複数のアンテナにより取得された値から合成雨量を算出することによりデータ欠損を防いでいる．しかし，レーダの観測，品質管理および運用は各地域で独自で行われている．従って，各地点の合成雨量は当該 14 地域毎に算出されるが，同一地点が複数地域に重複する場合でも地域間での連携はない．つまり，複数地域に重複する地点では各地域で合成雨量が個別に生成され，地域間で値が異なった場合も調整はされない．このため，配信データには同一地点における重複データが含まれる状況となっていた．

具体的な事象を「2015 年 9 月関東・東北豪雨」の画像（図 9 および図 10）に示す．富山・石川地域の画像（図 9）および近畿地域の画像（図 10）は同時刻の画像であり，それぞれの矢印は同一地点（琵琶湖南西部）を指す．双方とも左側が地域全体の画像，右側は対象地点の拡大画像である．矢印で示した箇所は富山・石川地域と近畿地域の両地域に重複して存在しており，同一地点にも関わらず両者で表示された降雨の強度が異なっていることがわかる．

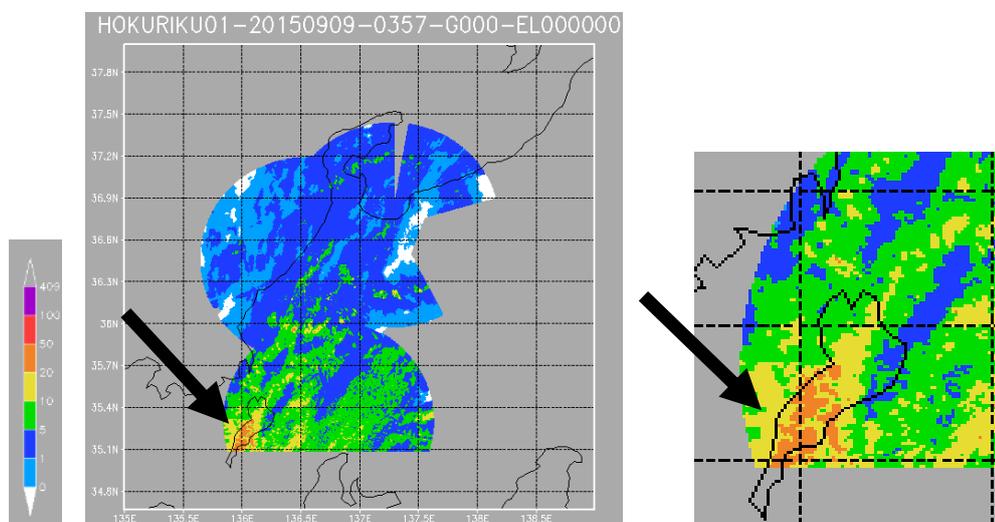


図9：富山・石川地域の地域画像および拡大画像(2015/09/09 03:57 JST)

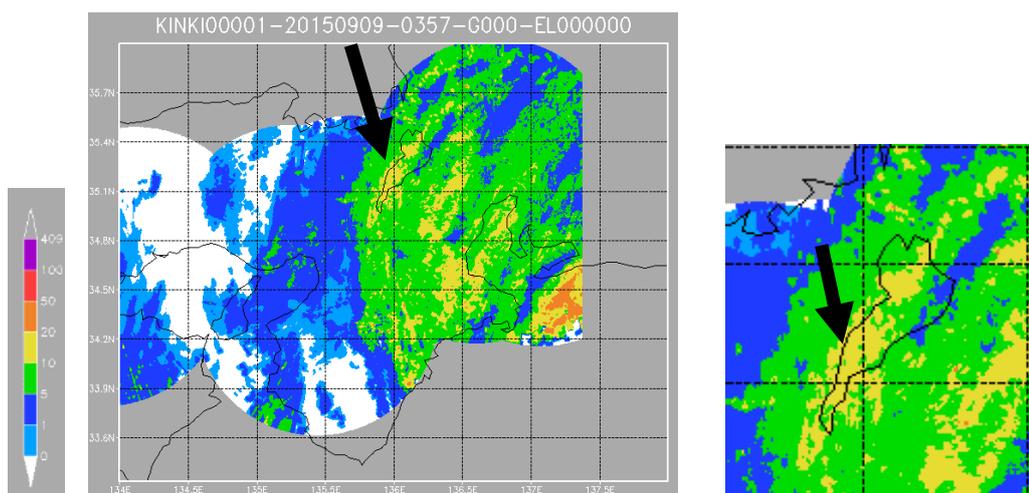


図10：近畿地域の地域画像および拡大画像(2015/09/09 03:57 JST)

図9の拡大図において赤い色で示された部分は強度の大きい降雨（豪雨等）を示し、リスク度が大きい降雨観測情報と言える。一方、当該情報は図10の拡大図では検出されていない。このように同一地点にも拘わらず地域毎の結果は異なっている。ここで、本事象が発生した地点周辺におけるXバンドMPレーダの設置状況を図11に示す。



図 11：豪雨周辺の主要レーダ設置イメージ

図 11 の赤丸は富山・石川地域に設置されていた X バンド MP レーダの観測範囲を示し、黒丸は近畿地域に設置されていた X バンド MP レーダの観測範囲を示す。近畿地域で強度の大きい降雨情報を取得できなかった決定的要因は特定できていない。しかし、両レーダの間に強度の大きい降雨が発生していたことがわかっており、状況的に、近畿地域に設置された X バンド MP レーダのレーダ波が当該降雨に阻まれ減衰したと推測される。この結果、両地域間で異なる合成雨量の算出結果が地域間で調整がされず、重複データになったと考えられる。

#### 2.5.2.2 高リスク優先手法の提案

各地域の算出結果は地域単位で正しく計算された結果であり、いずれの地域の値が正かの即時判断はできない。しかし、防災対策上、地点の潜在リスクの見逃しやこれに伴う行動判断の遅れを防ぐ必要性から、本研究ではリスクが高いと推測される値を優先して表示すべきと考えた。つまり、本研究では、「大きい値は高いリスクを持つ情報」と位置付け、同一地点に複数の異なる値が存在する場合は最も強度の大きい（リスクの高い）合成雨量の値をその地点の値として採択する手法を提案する。本研究ではこの提案手法を「高リスク優先手法」と呼ぶ。

全国合成処理では「高リスク優先手法」を以下の処理手順で適用する。

- ① 日本全国を覆う全地点のデータの初期値を「-9999」とする。
- ② 処理対象の値が同地点における設定済の値より大きい場合は、処理対象の値を当該地点に設定する。逆の場合は変更しない。

### 2.5.2.3 実装結果および考察

提案手法による実装結果を図12に示す。矢印で示した部分を参照すると、富山・石川地域の観測情報が近畿地域における同地点の観測情報より優先され、当該地点で最も強度の大きい降雨を示す情報として表示されている。このように、本提案手法はXバンドMPレーダデータの重複データを排除し、高リスク情報を優先表示する。これにより、本提案手法は防災上の観点から迅速かつ有用な情報提供が可能と考える。

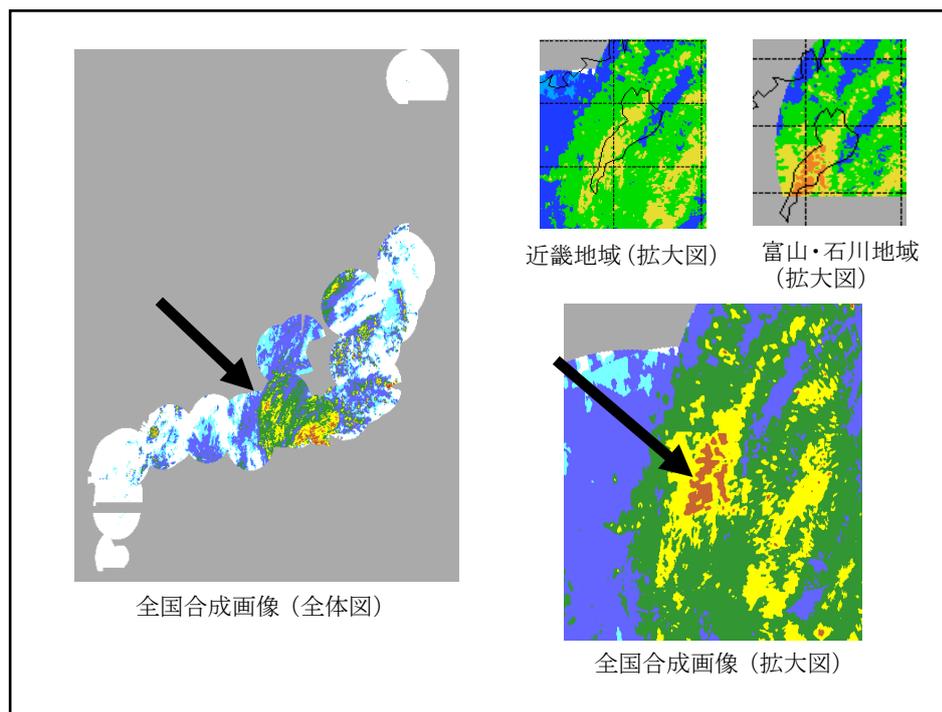


図12：高リスク優先手法による全国合成(2015/09/09 03:57 JST)

### 2.5.3 高値採択型圧縮手法による表示速度の高速化

2.4.3 で述べた X バンド MP レーダデータの画像閲覧方法のうち、一般向けの閲覧では、利便性を考慮して携帯端末を利用した閲覧を可能としている。この閲覧時に表示倍率（Google Map のズームレベル）を指定した際、全国合成画像の表示性能が低く、倍率指定から画像が完全に表示されるまで3分以上要する事例が多く見られた。従って、全国画像を携帯端末で表示する際の表示性能について改善が必要となった。本セクションでは、当該課題の詳細と解決手法を述べる。

#### 2.5.3.1 拡大表示時の性能に関する課題

一般向けの閲覧で表示倍率を指定した際、表示対象が地域画像の場合は性能の低下が確認されず、概ね3秒以内で全ての倍率での表示が可能であった。しかし、地域画像から全国画像へ表示切替を行った場合や、全国画像の表示において倍率を変更した場合には、切替指定から3分以上経過しても表示されない事例があった。

この要因として、全国合成画像の情報量の大きさの影響が挙げられた。全国合成画像の情報量は78MBで地域画像の0.7~2MBと比べ約39~111倍の情報量となっている。このため、本研究では地域画像と同等の3秒以内での表示性能を満たすことを企図し、全国合成画像の情報量を圧縮する手法を検討した。

#### 2.5.3.2 高値採択型圧縮手法の提案

一般的に、地点の状況をズームアップし拡大表示する場合は、高解像度での情報閲覧が求められる。しかし、ズームアウトして全体表示を行う場合には、解像度を低減させても見え方に影響を及ぼす度合いが低い。従って、本研究では、必要な解像度は指定倍率により異なると考え、ユーザが指定した表示倍率（Google Map のズームレベル）に応じて表示対象を圧縮画像に切替え、性能改善を図ることとした。

まず、画像を4分の1に圧縮する案と、Cバンドレーダの解像度と同等の16分の1に圧縮する手法で表示性能を比較した。この結果、4分の1に圧縮する案では、表示に平均1分20秒を要した。一方、16分の1とした場合は地域画像の表示と同等の3秒以内を実現できた。このため、圧縮の度合いをCバンドレーダの解像度と同等の16分の1とすることとした。

また、現状の携帯端末では、日本全体が液晶画面で閲覧可能なのは Google Maps API のズームレベルが 0~6 の間であった。従って、ユーザの指定倍率がズームレベル 0~6 までの場合には全国合成の圧縮画像を表示し、7~21 (最大) までの場合には各地域画像を加工せず表示する。

次に、圧縮の方法を述べる。ここで、「16 分の 1 に圧縮する」とは「16 の X バンド MP レーダデータを 1 データに集約すること」を示す。当該手法は次の 2 案で検討した。

案 A : 16 地点における最大値を抽出し、その値を 1 領域として設定。

案 B : 16 地点における平均値を抽出し、その値を 1 領域として設定。

C バンドレーダの観測値は 1 km 四方の平均的な降雨を示すため、案 B の結果は C バンドレーダの観測値と近似する可能性がある。しかし、X バンド MP レーダは局地的な情報であり、ある地点に豪雨が観測されても隣接地点の観測値は異なる強度を示す可能性が高い。従って、仮に案 B を採用した場合、強度の大きい観測情報（高リスク情報）は隣接する弱雨と平均化されて値が低くシフトされ、リスク情報が十分に表現されなくなる危険性がある。本研究は、防災上の観点から降雨に関する高リスク情報の迅速な提供を目的としているため、高リスク情報を十分に表現できない案 B は適切でないと判断した。従って、本研究では案 A を採択する。当該手法を「高値採択型圧縮手法」と呼び、全国合成画像の解像度低減に使用する。

具体的な処理内容としては、1 地点毎に降雨の強度を評価し、最も高い値を当該地点の「16 分の 1 情報」として設定する (図 13)。

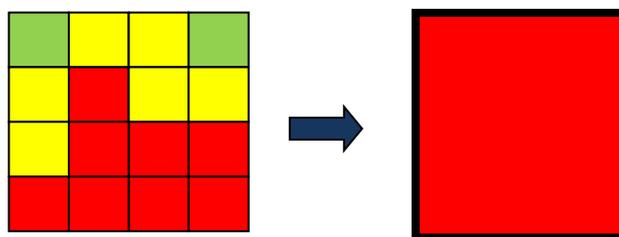


図 13 : 高値採択型圧縮手法のイメージ

処理手順は次のとおりとする。

- ① XバンドMPレーダのXYの2次元配列[ARRAY1] (250mメッシュ) についてXおよびYとも4分の1にした別の2次元配列[ARRAY2] (1kmメッシュ)の領域確保を行い、初期化する(初期値は“-999.9”とする)。
- ② [ARRAY1]における16のXバンドMPレーダデータを順に読み込み[ARRAY2]の値と比較する。ここで、[ARRAY1]>[ARRAY2]の場合のみ、当該[ARRAY1]のデータを[ARRAY2]に書き込む。これを[ARRAY1]の全データについて処理する。

### 2.5.3.3 実装結果および考察

実装結果として、携帯液晶画面に日本全国が見えるズームレベル(ズームレベル5)における高値採択型圧縮手法の適用前後の画像(図14のうち左側が適用前、右側が適用後)を比較して示す。

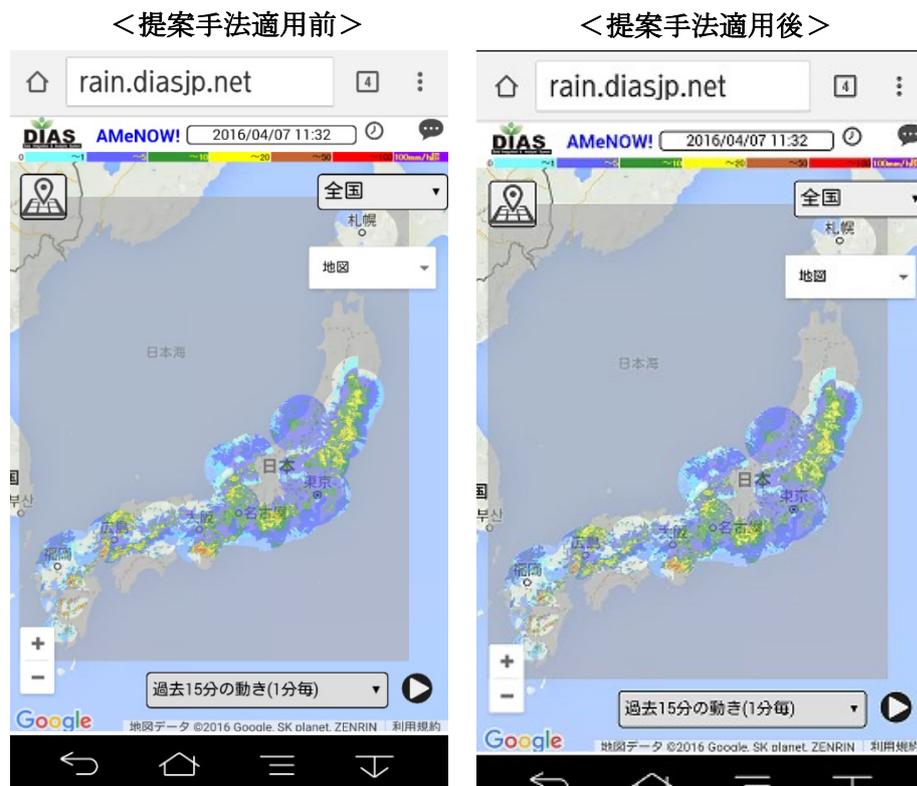


図14：全国合成画像の比較(ズームレベル5)

図14に示したとおり、両者において圧縮の適用による見た目の差はほぼない。一方、表示時間は当初の3分強から平均3秒以内まで大幅に改善し、全てのズームレベルにおいて高い表示性能を実現できた。本手法は強度の大きい降雨情報を優先表示するため、低解像度であっても高リスク情報を見逃すことはない。また、ズームレベルを7以上に拡大した場合は各地域の詳細画像に切り替え、本来の高解像度での局地的な降雨情報を表示する。このように、全体情報として潜在的な高リスク情報を優先表示し、詳細情報に切り替えれば本来の高解像度画像から局地的な降雨情報を把握可能であるため、防災上のリスク管理に適した実装ができたと考える。

## 2.6 DIAS における画像閲覧機能

XバンドMPレーダデータの閲覧環境は2015年9月より一般公開を開始した。閲覧方法は2.4.3で述べたとおり、一般向けおよび研究者向けの2通りで提供している。以降、各画面の表示例を示す。

まず、一般用の情報閲覧機能は「AMeNOW!」として公開している。ユーザは最新情報の閲覧のほか、日時指定による過去画像の閲覧や動画再生も可能である。ズームレベルはGoogle Maps APIの仕様の範囲で選択できる(図15)。

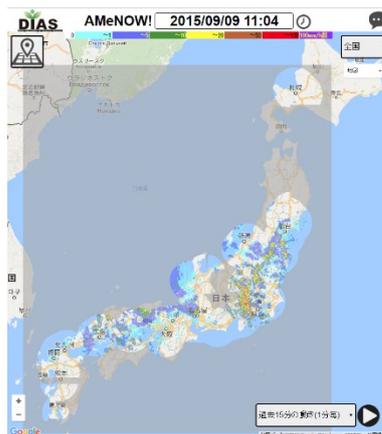


図15:「AMeNOW!」のサンプル画面

---

## 第2章 局地的気象観測データの可視化

次に、研究者向けの閲覧機能では、リアルタイムでの閲覧のほか、日時指定による過去画像の閲覧、動画再生およびデータダウンロードも可能としている（図16）。

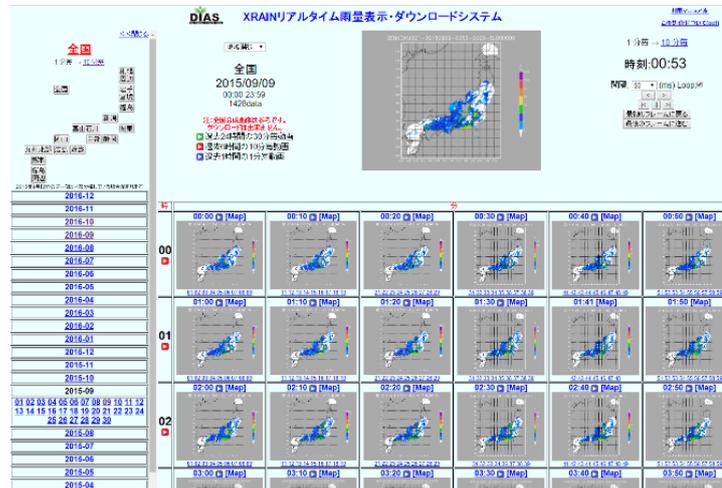


図16：研究者向け閲覧画面

なお、研究者向けに提供している画像は Google Map へのオーバーレイ（図17）と GrADS を利用して緯度経度の格子上に展開した画像（図18）の2種で表示可能であり、サムネイル上で選択して表示する。

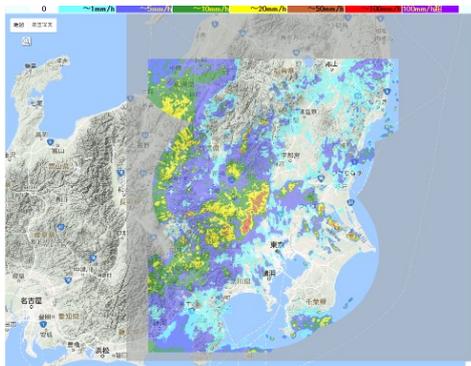


図17：Google Map へのオーバーレイ

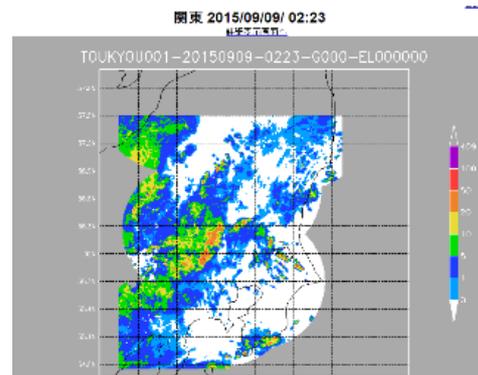


図18：GrADS の利用による画像

これらの画像はデータ取得後1分以内に閲覧可能であり、画像選択から表示までの時間は平均1秒（最大で3秒以内）である。これらの閲覧機能は2015年9月から公開を開始した。

## 2.7 第2章のまとめ

本章では、本研究で構築した「局地的な降雨情報をリアルタイムで閲覧可能とする環境」を紹介した。XバンドMPレーダデータはデータ量が膨大であるため、地球環境関連データでは国内最大規模のストレージを保有するDIAS上にデータの取得および蓄積を行う環境が開発された。これを受けて、本研究では当該蓄積データを迅速に閲覧可能とする環境を構築した。この際に生じた3つの課題（①処理遅延、②重複データの混在、③表示性能の改善）では、①は画像生成およびユーザ要求の処理環境を分割し、処理の並列化、および割込処理を組み合わせることで処理時間短縮を実現し、②は高リスク優先手法の提案、③は高値採択型圧縮手法の提案により、それぞれ解決を図った。

昨今の急激な気象の変化を勘案すると、局地的な降雨に関する気象観測情報は更に利用ニーズが高まると考えられることから、現在のXバンドMPレーダだけでなく、今後さまざまな高分解能かつ高頻度を実現する観測情報が整備されると予想される。その場合、膨大なデータを取得、加工および表示する際に処理性能面で本研究と同様の課題が生じる可能性が考えられる。

本研究は、局地的な降雨情報を迅速に情報共有することを目的とし、XバンドMPレーダデータを画像として迅速に閲覧できる環境を構築した。加えて、データ取得および表示に関しては実用的な性能を確保した。本研究における提案手法が同様の事例における参考となれば幸いである。また、本環境の利用により、局地的な降雨情報に関する経験学習が広く蓄積されることを期待したい。経験学習を重ねることで水災害に対する防災意識が高まり、被害軽減に繋げることができれば幸いである。

## 第3章.

# 気象観測データの合成による降雨情報の 推定手法

第2章ではXバンドMPレーダの情報共有基盤の構築について述べた。しかしXバンドMPレーダは短波ゆえに設置場所が限られ、国内には同レーダの観測情報を取得できない地点が多数存在する。当該地点におけるXバンドMPレーダの代替情報としては、XバンドMPレーダの次に解像度が高いCバンドレーダの観測情報が挙げられる。しかし、Cバンドレーダデータは空間分解能および配信頻度が異なり、単純に情報を重ね合わせるができない。

そこで本研究では、両者の長所を活かしながら国内全地点の「局地的かつリアルタイムの降雨情報」として推定する手法を提案する。併せて、両レーダの観測情報が重複する地点では、第2章で述べた「高リスク優先手法」により値を選別し、潜在的な高いリスク情報を優先して提供する仕組みとした。本章では当該手法の詳細を述べる。

### 3.1 研究背景

XバンドMPレーダは短波ゆえに、山間部等を中心とした遮蔽物が多い地域には設置が難しい。逆に、こうした地域は土砂災害や上流河川の氾濫の契機となり得ることが多く、最も迅速に情報を把握可能なXバンドMPレーダで観測できない事情は課題であると言える。これらの地域の降雨情報は国内全域を観測域とするCバンドレーダから得ることが出来る。しかし、両レーダを同時に合わせ見る手段は確立されておらず、機動的な対処が難しい。このため本研究では、局地的な降雨を観測するXバンドMPレーダと国内全域を観測するCバンドレーダの双方のデータを合成し、国内全地点の局地的な降雨の推定値をリアルタイムで生成する新たな手法を検討した。

## 3.2 研究課題

まず、同一日時における C バンドレーダおよび X バンド MP レーダの観測情報を図 19 および図 20 に示す。なお、いずれの画像も左が全体図および右が拡大図であり、グレーの箇所は各レーダの観測範囲外地点を示す。

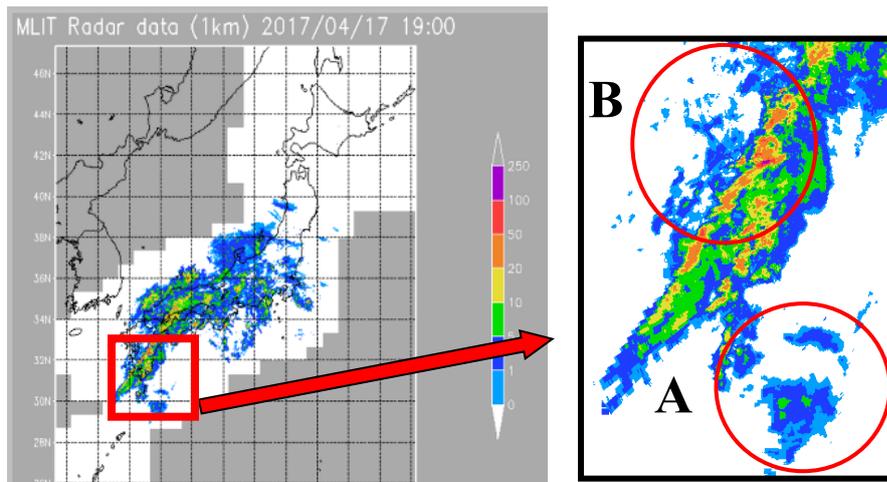


図 19 : C バンドレーダによる観測情報 (2017/04/17 19:00 JST)

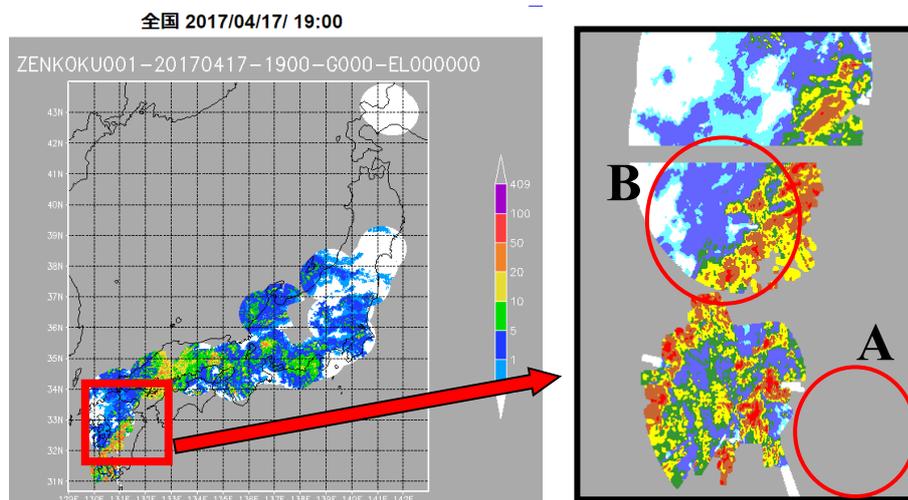


図 20 : X バンド MP レーダによる観測情報 (2017/04/17 19:00 JST)

図19に示すとおり、Cバンドレーダの観測範囲には日本全域が含まれる。一方、図20に示すとおり、XバンドMPレーダは観測範囲外となっている地域が多く、特に北海道地域や東北地方の北部はその殆どが観測範囲外である。従って、本研究ではこれらの地域の観測情報をCバンドレーダデータにより補完することを検討した。図19および図20を参照すると次の事項が確認できる。

- ① XバンドMPレーダの観測外地域における降雨情報をCバンドレーダで補完することが可能（図19のAおよび図20のA）。
- ② Cバンドレーダでは検出されていない強度の大きい降雨情報をXバンドMPレーダで補完することが可能（図19のBおよび図20のB）。

上述のとおり、両レーダ情報の合成によりそれぞれの長所を活かし、国内全域を対象とした高い空間分解能の降雨情報の生成が可能になると考えられる。しかし、両レーダ間で異なる条件や両レーダの観測地点が重なる場合にいずれの値を採択するかという点を含め、以下の課題解決が必要である。

1. 両レーダで異なる配信間隔を勘案した処理タイミング
2. 両レーダで異なる空間分解能を勘案したデータ分割方法
3. 同一地点における両レーダデータの採択基準

Cバンドレーダデータは1kmメッシュでの観測情報であり、これは、1領域が250mメッシュであるXバンドMPレーダの観測情報の16領域分の降雨を平均化した値とほぼ同値と考えられる。従って、同一地点における両レーダの観測値を比較した場合、XバンドMPレーダが示す値は当該地点を含むCバンドレーダの示す値の同程度以上の数値となる可能性が高い。つまり、リスクの高い局地的な降雨情報を取得する点では、CバンドレーダよりXバンドMPレーダの観測情報に優位性があると言える。このため、本研究では防災上の観点から局地的な高リスク情報を迅速に取得および共有を行うことが必要と考え、両レーダデータを合成する際はXバンドMPレーダの観測値を優先する。従って、両レーダで異なる配信間隔と空間分解能もXバンドMPレーダの条件を優先する。以降のセクションでは、両レーダデータの合成における提案手法の詳細を述べる。

### 3.3 レーダ情報合成に関する提案手法

本研究におけるCバンドレーダデータおよびXバンドMPレーダデータの合成は、防災上の観点から「高リスク情報の迅速な提供」を解決する提案とする。従って、両レーダで異なる観測間隔と空間分解能はXバンドMPレーダを優先する。また、同一地点に両レーダのデータが存在する場合は、高リスクである方（大きい値を示す方のレーダ観測値）を採択する。本セクションではこれらの提案手法を詳細に説明する。

#### 3.3.1 合成処理タイミングの同期

両レーダのデータ合成にあたり、まずは処理対象時刻を合わせる必要がある。この手法として、以下の2案を検討した。

案1：Cバンドレーダの配信タイミングでの合成（図21）

案2：XバンドMPレーダの配信タイミングでの合成（図22）

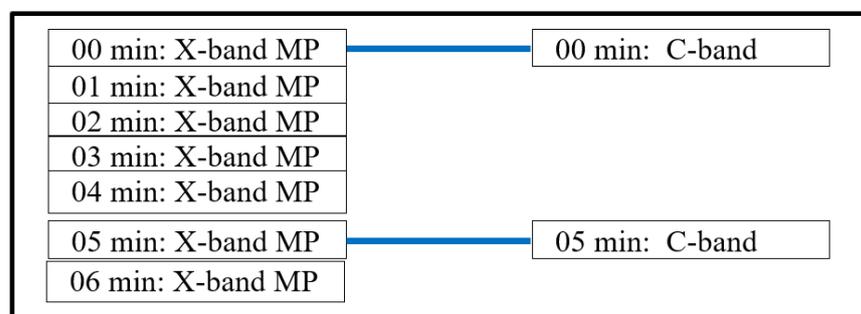


図21：Cバンドレーダの配信タイミングでの合成（案1）

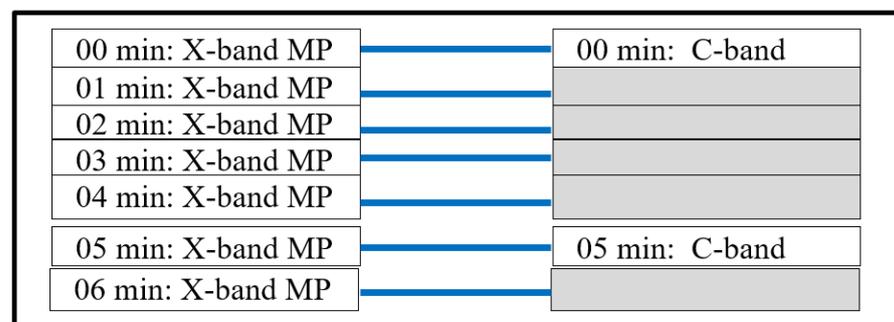


図22：XバンドMPレーダの配信タイミングでの合成（案2）

まず、案1は、合成処理が5分間隔で実行される。このため、当該間隔内（例えば1～4分、6～9分など）に観測されたXバンドMPレーダデータは処理対象外となる。従って、この時間帯に強度の大きいXバンドMPレーダの値が観測された場合は当該高リスク情報を見逃す可能性がある。ここで、Cバンドレーダデータの合成対象をCバンドレーダと同時刻のXバンドMPレーダではなく、Cバンドレーダと同時刻から直近5分間のXバンドMPレーダデータの平均値と合成する案を考える。しかし、この場合でも、降雨強度が大きいXバンドMPレーダの情報は他の時間帯の弱雨と平均化され、本来の高リスク情報を検出できない可能性がある。従って、案1は実際の観測値を利用するため信頼性が高い点がメリットである一方で、処理頻度が低く、強度の大きい情報をタイムリーに反映できないデメリットがある。

次に、案2は、合成処理が毎分実行される。このため、XバンドMPレーダの観測値をタイムリーに利用でき、局地的大雨の情報を逃さない。しかし、5分毎の時間帯以外（1～4分前等）は合成対象のCバンドレーダデータが存在しないため、当該時間帯のCバンドレーダデータを別途生成しなければならない。従って、案2はXバンドMPレーダの実測値を全て利用し局地的大雨情報の見逃さない点がメリットである一方で、推定値を別途生成することで値の信頼性が低下するデメリットがある。

これらを比較した結果、本研究では高リスク情報を漏れなく検出する案を優先し、案2を採択した。観測時間帯以外のCバンドレーダデータの推定値は、リアルタイム性を優先し「最新のCバンドレーダデータ」を利用することが適切と考えた。従って、当該時間のXバンドMPレーダデータの合成対象となるCバンドレーダは、直近の実測値を元として、図23のとおり対応させる。

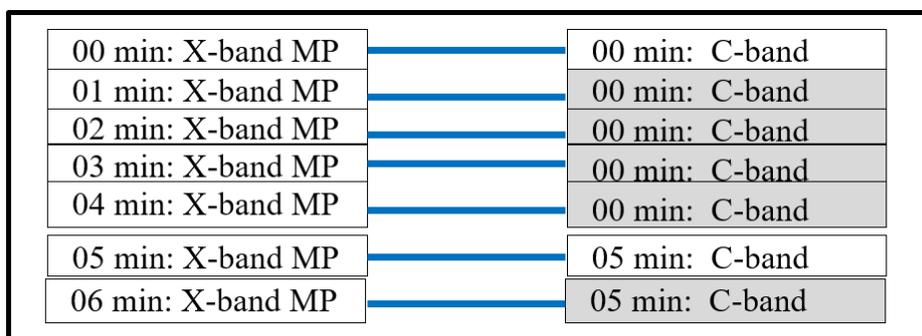


図23：レーダ情報合成の処理タイミングイメージ

### 3.3.2 空間分解能の同期

次に、両レーダデータにおける空間分解能の同期手法を述べる。まず、Cバンドレーダデータの空間分解能をXバンドMPレーダに合わせ16分割する。この分割および両レーダデータの比較イメージを図25に示す。

- ① Cバンドレーダデータの1つの観測領域を16分割する。
- ② 16分割後の各領域に当該Cバンドレーダデータの値を設定する。
- ③ 該当する地点のXバンドMPレーダの値と比較し、当該地点の値を決定する（値の採択手法は3.3.3に後述）。

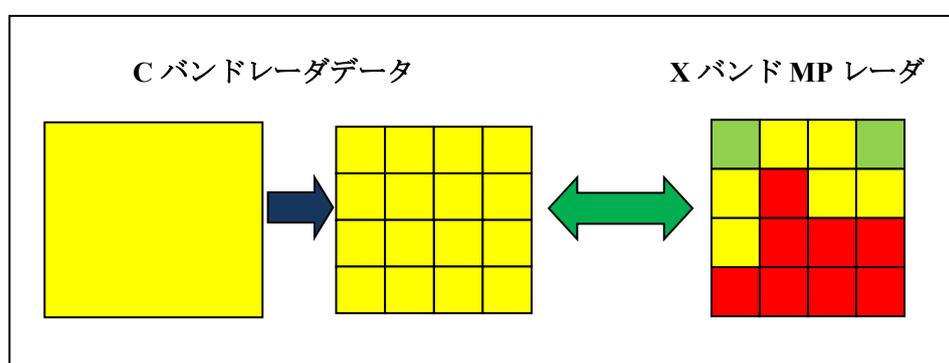


図24：空間分解能の同期イメージ

### 3.3.3 同一地点におけるデータ値の採択手法

同一地点における値の採択手法は以下の2案で検討した。

案A：地点の値は一律XバンドMPレーダの値を採択する。

案B：両レーダの値を比較し大きい方をその地点の値として採択する。

前述で示したCバンドレーダデータを16領域に分割設定した値は、平均化された1kmメッシュの観測値を便宜的に各領域に割り当てたものであり、実測値ではない。また、局地的な降雨情報の観測ではXバンドMPレーダに優位性があり、強度の大きい降雨が発生した場合には同地点を含むCバンドレーダデータの値よりXバンドMPレーダの値の方が大きくなる傾向がある。

案 A および案 B による両レーダデータの合成イメージを図 25 に示す。

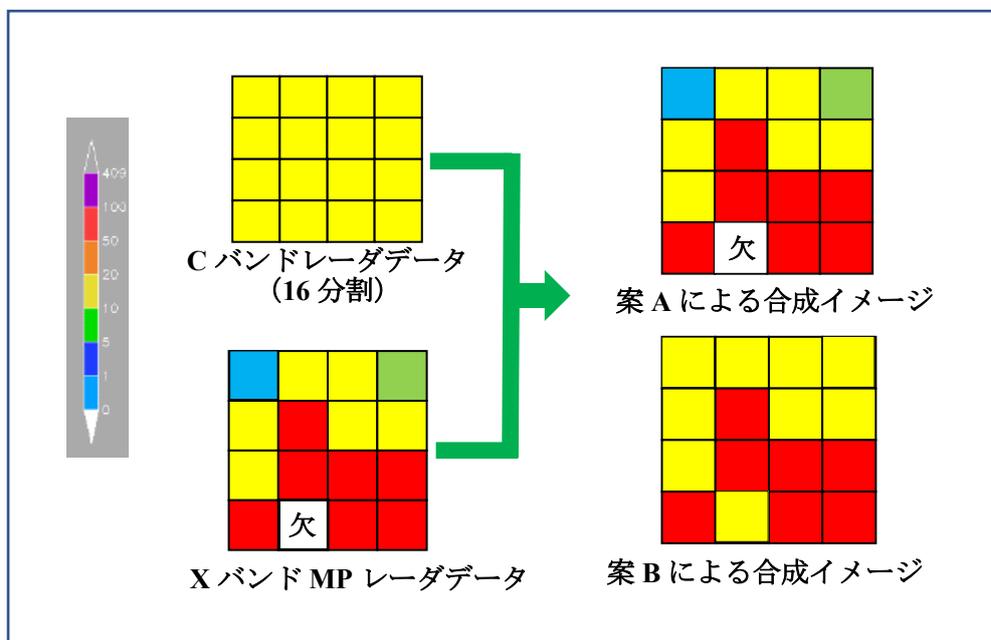


図 25 : 案 A および案 B によるデータ合成イメージ

まず、案 A は X バンド MP レーダデータの値を優先させるため、大雨等の高リスク情報を逃さない点で効果的と言える。しかし、X バンド MP レーダのレーダ波が減衰した場合には本来より低い値が設定される可能性があるほか、欠損の場合は補完できる情報が存在しないデメリットがある。

次に、案 B は 2 章で述べた「高リスク優先手法」の考え方に基づく手法である。本手法を適用すると、強度の大きい降雨は殆どの場合で X バンド MP レーダの値が採択される。また、X バンド MP レーダのレーダ波減衰またはデータ欠損の場合でも当該値を C バンドレーダデータで補完できるため、案 B は事象に応じて常に最大リスク値を表示できるメリットがある。しかし、C バンドレーダは 1 km メッシュの平均値のため、X バンド MP レーダの観測値が弱雨の場合は C バンドレーダの値の方が高くなる。従って、案 B の手法は全体的に実際より高い値として表示される傾向となる可能性がある。

以上を検討した結果、本研究では値を確実に取得し常に高リスク情報を表す点を優先し、案 B を採択することとした。

## 3.4 実装

### 3.4.1 全体処理概要

図 26 に示すとおり、C バンドレーダの観測領域は日本全国を完全に覆う。一方、X バンド MP レーダによる観測領域は C バンドレーダの観測領域内に包含され、観測範囲外地点も存在する。

従って、合成手法としては、まず C バンドレーダデータの全領域を X バンド MP レーダデータの解像度に合わせて 16 倍に拡大し、当該領域に X バンド MP レーダデータ（データが存在する地点のみ）の全国合成データを重ね合わせる。

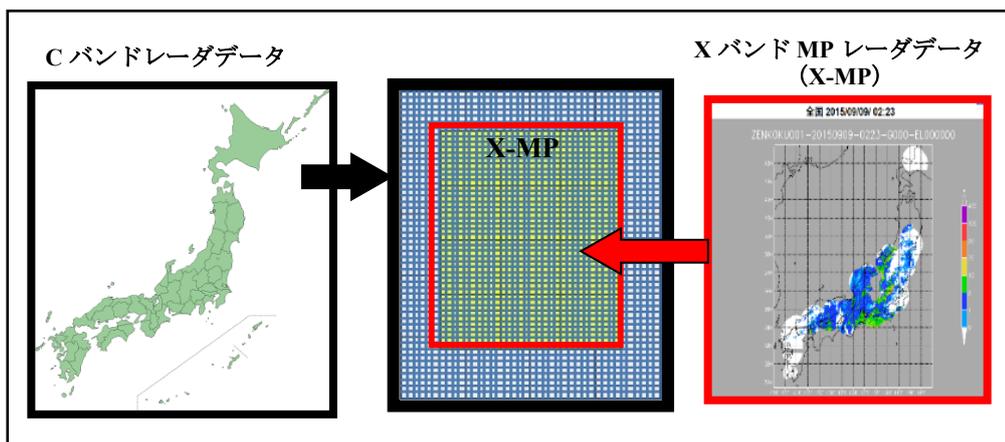


図 26 : 両レーダデータの合成イメージ

本研究で利用するデータは、国土交通省から配信され、DIAS 上に蓄積された X バンド MP レーダおよび C バンドレーダの観測データとする。DIAS 上での各レーダのデータ取得から両レーダデータの合成処理までの処理概要を図 27 に示す。

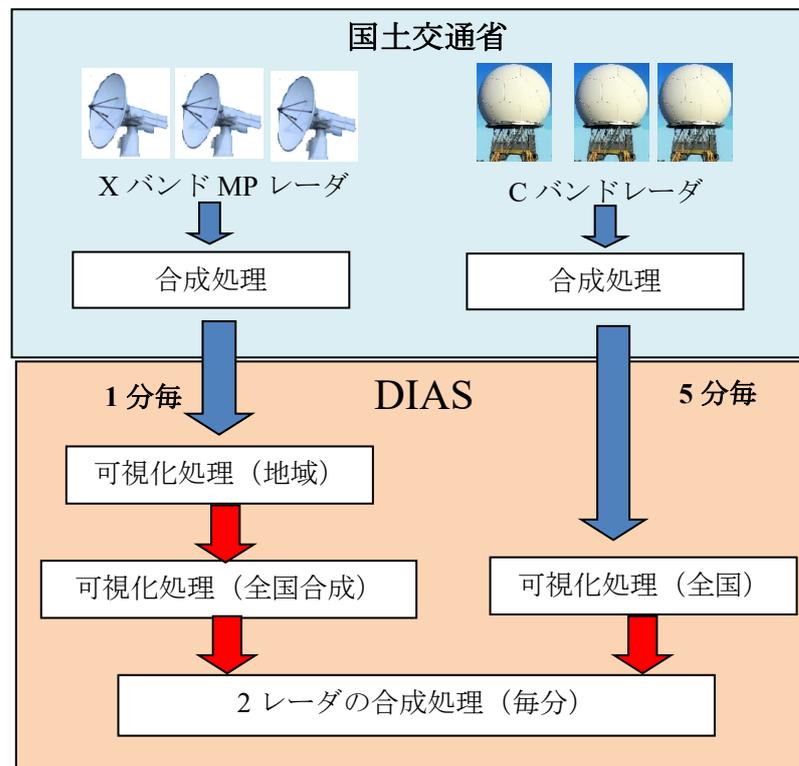


図 27：レーダ情報の合成に関する全体処理概要

処理は以下の順で実行する。

- ① Cバンドレーダの観測情報は国土交通省から DIAS に 5 分毎に配信される。一方、XバンドMPレーダのデータは国土交通省から 14 の地域毎に 1 分間隔で配信される。
- ② Cバンドレーダデータ（直近 5 分毎）と XバンドMPレーダ（毎分）を受信次第可視化する。
- ③ XバンドMPレーダの全国合成データ（毎分）を可視化する。
- ④ Cバンドレーダデータ（直近 5 分）と XバンドMPレーダの全国合成データを 1 分毎に合成し可視化する。

### 3.4.2 合成処理仕様

Cバンドレーダデータは各地域ブロックのセルデータに図28に示す3次メッシュ（1kmメッシュ）として格納されている。従って、ブロックヘッダ内の緯度経度情報からスタート地点Pを特定し、読み出したセルデータ（3次メッシュデータ）を2次元配列として展開した配列Aに格納する。同様に、XバンドMPレーダデータについても図29に示す1/16の3次メッシュ（250mメッシュ）として格納されているセルデータを読み出し、配列Bに格納する。

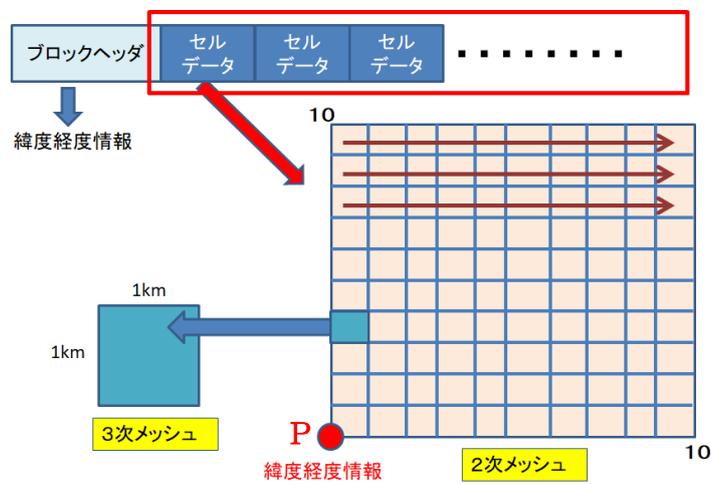


図28：Cバンドレーダデータの配信イメージ

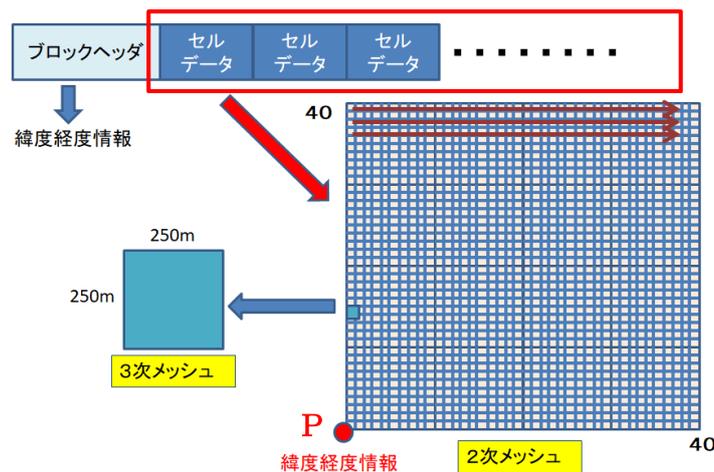


図29：XバンドMPレーダデータの配信イメージ

次に、設定済の配列 A および配列 B を利用して両データを合成する。具体的には、図 30 に示す C バンドレーダデータの 1 次メッシュ領域 (X 方向 23 および Y 方向 35) を 16 倍にした領域と X バンド MP レーダの 1 次メッシュ領域 (X 方向 14 および Y 方向 20) を、各スタート地点 S(C) および S(X) を起点として重ね合わせる。各配列数は、次のとおりとなる。

- ① C バンドレーダデータ (X 方向)  
23 (1 次メッシュ) \* 8 (2 次メッシュ) \* 10 (3 次メッシュ) \* 4 (4 分割)
- ② C バンドレーダデータ (Y 方向)  
35 (1 次メッシュ) \* 8 (2 次メッシュ) \* 10 (3 次メッシュ) \* 4 (4 分割)
- ③ X バンド MP レーダ (X 方向)  
14 (1 次メッシュ) \* 8 (2 次メッシュ) \* 10 (3 次メッシュ) \* 4 (4 分割)
- ④ X バンド MP レーダ (Y 方向)  
20 (1 次メッシュ) \* 8 (2 次メッシュ) \* 10 (3 次メッシュ) \* 4 (4 分割)

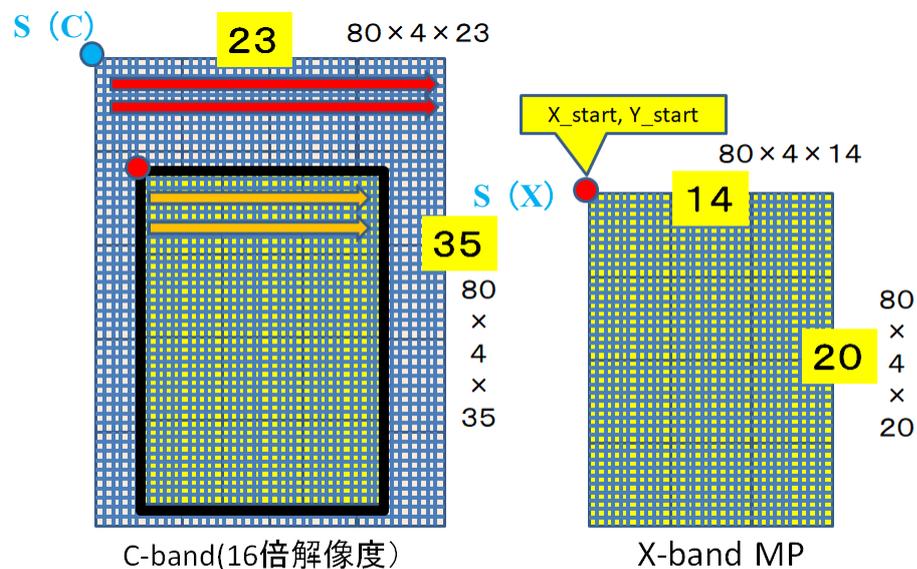


図 30 : 2 次元配列への設定イメージ

各配列 A(X,Y) と B(X,Y) 内の値を比較し、大きい方の値を当該地点の値として配列 C(X,Y) に設定する。これを全配列について処理して合成データとする。

### 3.5 実装結果および考察

提案手法を適用し2つのレーダデータを合成した結果を図31に示す。

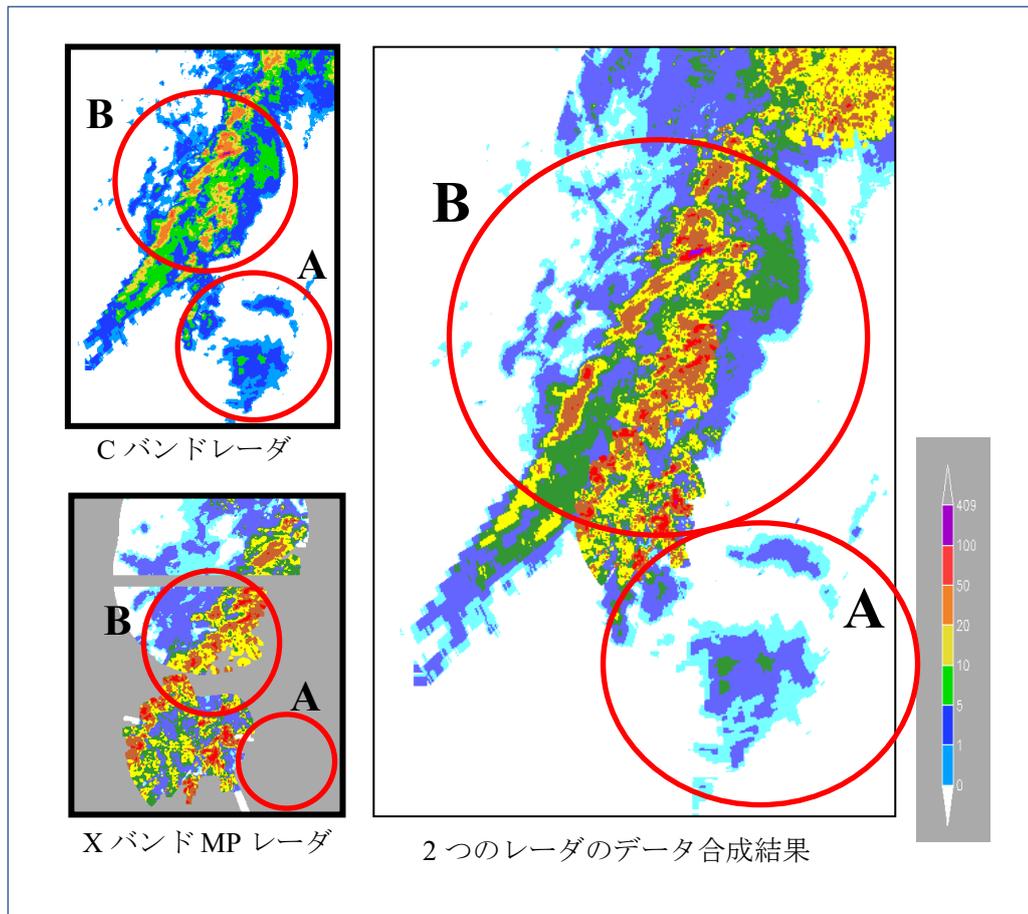


図31 実装結果 (2017/04/17 19:00 JST)

図31における合成画像のAを参照すると、XバンドMPレーダの観測範囲外地域における降雨情報がCバンドレーダにより補完されている。また、同合成画像のBを参照すると、Cバンドレーダでは検出できなかった局地的大雨情報がXバンドMPレーダの情報により補完されている。これらから、両レーダのデータを合成した結果は両者の長所を活かし補完し合っていることが示された。

ここで、両レーダが重なり合う地点において、どちらのレーダ値が採択されたかを当該領域数と全体に占める割合として表3に示す（表中ではCバンドレーダをC-band）、XバンドMPレーダをX-band MPと表す。

表3 レーダ値の比較結果別領域数および割合

2レーダの比較結果	領域数	全体に占める割合 (%)
C-band > X-band MP	1,813,352	28.9
C-band < X-band MP	2,424,375	38.6
C-band = X-band MP	2,037,506	32.5

表3を参照すると、まず、Cバンドレーダデータの値以上の数値を示したXバンドMPレーダの領域数は全体の71%超であり、そのうち同値となった割合は約半数であった。これにより、降雨強度の値として、XバンドMPレーダデータはCバンドレーダデータに優位性があると言える。逆に、XバンドMPレーダよりCバンドレーダデータの方が大きい値を示した領域は約30%存在した。これにより、Cバンドレーダデータは、波長の減衰等によりXバンドMPレーダが取得困難となった地点のデータを補完していることが確認できる。

次に、前述3.3.3における案Aと案Bを適用後の実装結果をそれぞれ拡大図として図32および図33に示す。

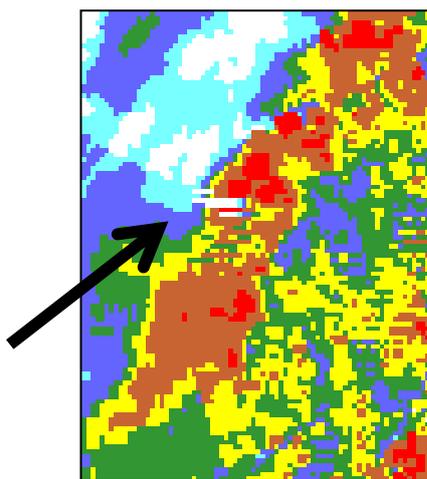


図32：案A適用後の実装結果

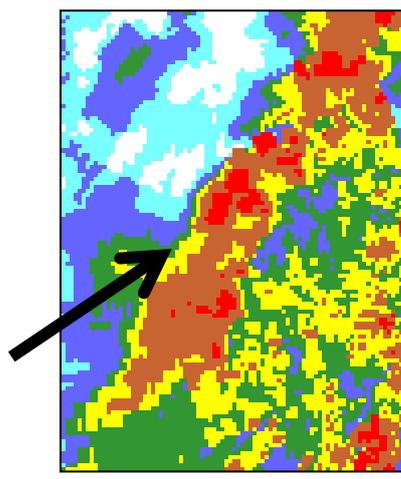


図33：案B適用後の実装結果

図を参照すると、常に X バンド MP レーダデータを優先する案 A の適用による実装結果では、矢印で示した部分の一部が X バンド MP レーダの値が取得できず情報が途切れている。しかし、案 B の適用による実装結果では、C バンドレーダの情報により欠損データが補完され、周囲の降雨とデータの連続性が保たれている。このように、X バンド MP レーダデータには取得時に不安定要因を内包することから、案 B が適切と結論付けることができた。以上により、本提案手法は観測範囲や降雨強度を補完しながら複数のレーダ情報を合成し、高リスク情報の提供が可能であることを示した。

### 3.6 第3章のまとめ

本章では、X バンド MP レーダデータに他の気象観測レーダを合成して、国内全域の降雨情報の推計値を得る手法を提案した。本手法の適用により、各レーダの長所を活かした合成結果を得ることが可能となり、国内全地点における局地的な降雨情報かつ高リスク情報をリアルタイムで取得できる。

現在の物理的な気象観測情報では、国内全域の局地的な降雨情報を全て取得できる観測方法（およびその情報）が存在しない。このため、非常時には複数の観測情報を合わせ見る必要があり、機動的な対処に課題があった。本研究における提案手法では、広域的な気象情報と局所的な気象情報を組み合わせ、現時点で把握可能な国内全域の局地的な高リスク情報を同時に提供することが可能である。国内全ての地点における危険な豪雨の兆候を、誰もがリアルタイムで検出することが可能になった点は、防災上の観点で有用性が高いと考える。本基盤が局地的な降雨情報のリアルタイム表示だけでなく、降雨データの研究材料として、今後発生し得る水災害への備えとして、または地域および住民の方々における降雨情報の経験学習として、広く貢献できれば幸いである。

## 第4章.

# 言語表現を用いた降雨強度レベル推定手法

これまで、既存の気象観測情報を組み合わせながら国内の局地的な降雨情報を提供する手法について紹介してきた。しかし最近では、これまで述べたような物理的観測情報のほか、投稿情報の集合知から気象情報を得る試みもある。投稿情報による表示は、現状、民間の情報サービス等において「晴れ・くもり・雨」等の天気区分で表すのが主流であり、弱雨か大雨か等の降雨の程度（降雨強度）までは把握できないものが大半である。しかし、投稿情報から降雨強度が把握できれば、防災上の観点からも有用な情報源となり、物理情報を補完する可能性も期待できる。

そこで本研究では、言語表現から降雨強度の推定を行う手法を提案する。本研究では、代表的な SNS の一つである Twitter の発信情報と国内で最も高解像度な観測情報である X バンド MP レーダとの数値的相関を解析し、言語表現と降雨強度の相関および降雨強度の季節的変動の特徴を捉え、言語表現パターンと季節（月）を組み合わせた推定モデルを構築した。本章では、本研究で提案する「言語表現から降雨強度の推定を行う手法」について述べる。

## 4.1 研究概要

本研究では言語表現の解析に、代表的な SNS の一つである Twitter を利用した。ソーシャルメディアに関する研究機関<sup>(注13)</sup>の最新動向調査<sup>(注14)</sup>によると、調査時点の国内の月間アクティブユーザー数では Twitter が約 4,500 万人と最も多く、続いて Facebook が 2,800 万人、Instagram が約 2,000 万人である。人による発信情報は地点や時点等に曖昧性があり、解析には統計的観点から多くのデータを必要とする。従って本研究では、より多くの情報を取得する観点から Twitter を用いることとした。

---

注13：ソーシャルメディアラボ <https://gaiax-socialmedialab.jp/>

注14：「2018年6月更新！11のソーシャルメディア最新動向データまとめ」  
<https://gaiax-socialmedialab.jp/post-30833/>

## 第4章

### 言語表現を用いた降雨強度レベル推定手法

本研究では、Twitter を利用した発信情報（以下、「Tweet」と呼ぶ。）に含まれる言語表現から降雨情報のレベルを推定する手法を提案する。

まず、本研究の対象地域は関東地域に限定した。そして、降雨強度の推定値は気象庁による雨の強さと降り方の関係<sup>(注15)</sup>を基準として、表4に示す5段階のレベルで表す。

表4 降雨強度レベル

強度レベル	降雨の強度
1	弱い雨： 10 mm/h 未満（0を除く）
2	やや強い雨： 10 mm/h 以上 20 mm/h 未満
3	強い雨：20 mm/h 以上 30 mm/h 未満
4	激しい雨：30 mm/h 以上 50 mm/h 未満
5	非常に激しい雨，猛烈な雨： 50 mm/h 以上

次に、本提案手法の概要を説明する。本研究は大別して以下の2部で構成している。

- ① 降雨強度レベルの推定を可能とする Tweet の特定
- ② 言語表現から物理情報を推定するモデルの提案

①の対象 Tweet は、当該 Tweet 内に「小雨」「豪雨」等、降雨強度レベルの推定を可能とする言語表現（以下、「推定言語表現」と呼ぶ）が含まれ、かつ位置情報および降雨情報の発生時刻が推定可能である Tweet とする。位置情報は先行研究を利用してジオタグ付きでない Tweet に推定位置情報を付加した。また、Tweet の示す降雨情報の発生時刻の推定には、独自の抽出手法を開発した。

②では、まず、推定言語表現を含む Tweet とこれに対応する X バンド MP レーダデータを照合した。この結果、「小雨」「霧雨」のように時間や季節によらず一定の降雨強度を示す表現や、「豪雨」「大雨」のように時間または季節により降雨強度が変動する表現など、推定言語表現と降雨強度に相関を検出できた。本研究ではこれらの推定言語表現を X バンド MP レーダの実績値と照合した際の一致度に応じて Category に分類し、当該 Category 毎に推定方法を割り当てる方法とした。

注15：気象庁「雨の強さと降り方」

[https://www.jma.go.jp/jma/kishou/known/yougo\\_hp/amehyo.html](https://www.jma.go.jp/jma/kishou/known/yougo_hp/amehyo.html)

---

降雨には弱雨から強雨に推移する時間的な変動と季節変動がある。このため、本研究では双方の変動の特徴を過去の Tweet および X バンド MP レーダの実績を用いて解析し、Category 毎に X バンド MP レーダから最適解を得る方法を定め、当該 Category 内の「推定言語表現」と「月」を要素として降雨強度のレベルを推定するモデルを構築した。

そして、本提案手法を 2018 年の事例に適用した結果、推定値と X バンド MP レーダの実績は年間平均 93%以上の精度で一致した。従って、本提案手法は言語表現から物理情報相当の値を得る有用な手法の一つと考える。

## 4.2 関連研究

SNS 情報と実データを関連付ける研究[5]では、Tweet の集合知から特定事象の把握や検出を試みるものが多い。代表的な例としては、Calton Pu らによる物理情報（地震と降雨）と SNS 情報（Twitter, Instagram, YouTube）を統合したシステムの開発[6]、地震発生を示す Tweet から高い確率で地震発生を検出する研究[7]、鉄道の遅延等に関する Tweet から当該事象を検知する研究[8]、自然災害の発生を検知しネットワーク切替えを試みる研究[9]、Tweet 内の言語表現から体感温度を取得する研究[10]、事故や渋滞に関する Tweet から交通障害の早期発見を行う研究[11]等、ジオタグ付 Tweet の内容や移動速度および距離から交通路を抽出する研究[12]、および Tweet 内の食に関する言語表現から言語圏毎に食の文化比較を行った研究[13]等がある。

降雨と Tweet の関連研究[114-214]では、「浸水・冠水」の関連 Tweet から災害検出を行う研究[14][15][16]、降雨の関連 Tweet から当該降雨の持続時間検出を行う研究[17]、Twitter から天気（晴れ、曇りまたは雨）を把握する研究[18]、局地的な降雨のバースト現象を捉える研究[19][20]、土砂災害の前兆を捉える研究[21]、気象庁の降雨データと「雨」「地滑り」の単語の関連性を述べた研究[22]、災害支援が必要な地点の検出を行う研究[23]等がある。

また、ジオタグ付きの Tweet から当該時点の状況を抽出する研究[24]では、Tweet から降雨有無を把握可能であり、位置情報を正確にする必要性を述べている。本研究ではこれを参考にし、ジオタグ付き Tweet に加えて先行研究により位置情報を付加した Tweet を対象とした..

投稿情報から天気を表示する試みでは、ヤフー株式会社が運営するポータルサイト Yahoo! JAPAN 上の「みんなで実況・今の天気」<sup>(注16)</sup>、および株式会社ウェザーニューズが運営する「ウェザーリポート」<sup>(注17)</sup>がある。「みんなで実況・今の天気」は、登録ユーザが投稿した天気区分（晴れ、くもり、雨、雪）の投稿数を直近1時間で集計し、10分毎に公開する。「ウェザーリポート」では、ユーザが投稿した画像や天気区分の内容をリアルタイムで共有する。両サービスとも投稿および集計は市町村区の比較的広域な単位であり、天気情報は設定区分（晴れ・くもり・雨等）に従うため降雨強度までは表示されない。

## 4.3 対象 Tweet の特定

本研究では、降雨情報を正確に得るため、地点と時点が特定可能な Tweet を対象とする。しかし、位置情報を把握可能なジオタグ付き Tweet は非常に数が少ない。また、「小降りだった」、「猛烈な雨が降っている」または「雨が降りそう」など、Tweet 内で示される降雨の時点はさまざまである。従って本研究では、先行研究を利用して位置情報付き Tweet の対象数を拡張するとともに、独自の抽出手法を開発し、Tweet の示す降雨情報の時点を特定する。本章では当該特定方法について述べる。

### 4.3.1 発信地点の特定

発信地点に正確性を求める場合、一般的にはジオタグ付き Tweet の利用が適する。しかし、これらは Tweet 全体の1%未満であり、非常に数が少ない[25]。このため、本研究では、伊藤・吉永らによる先行研究[26][27]を利用してジオタグがない Tweet に「推定位置情報（緯度経度）」を付加し、対象 Tweet を拡大した。

伊藤・吉永らの手法はマイクロブログ (Tweet) ストリーム中の投稿中に現れる単語を元に、Tweet に推定位置情報を付加する。本手法は、まず、Tweet がジオタグ付か否かで位置情報の付加方法を大別し、ジオタグ付の場合は当該位置情報を推定位置情報に設定する。一方、ジオタグがない場合は、地名辞書に基づき位置情報を推定する。本手法で利用する地名辞書は、Wikipedia<sup>(注18)</sup>に含まれる地名

---

(注 16): [https://www.yahoo-help.jp/app/answers/detail/p/616/a\\_id/44565/faq/pc-detail](https://www.yahoo-help.jp/app/answers/detail/p/616/a_id/44565/faq/pc-detail)  
(2019年9月30日時点)

(注 17): <http://weathernews.jp/report/> (2019年9月30日時点)

(注 18): <https://www.wikipedia.org>

と位置情報に加え、Foursquare<sup>(注19)</sup>のような位置情報サービスを介して投稿された Tweet から得た施設名と位置を組み合わせ構築しており、当該場所（施設）の中心地点が推定位置情報として割り当てられる。加えて、特定の位置と強く関連付けられる名詞も都度追加する。例えば、あるイベントが特定の場所で開催された場合、Tweet 内に当該イベント名が含まれば、場所情報を含むか否かに関わらず、イベント開催地点を推定位置情報に設定する。ただし、本研究では照合対象である X バンド MP レーダが極めて高解像度であるため、Tweet の位置情報誤差を勘案し、縣市町村名に基づき推定位置情報が付加された Tweet は除外した。この手法の全体概要を図 1 に示す。

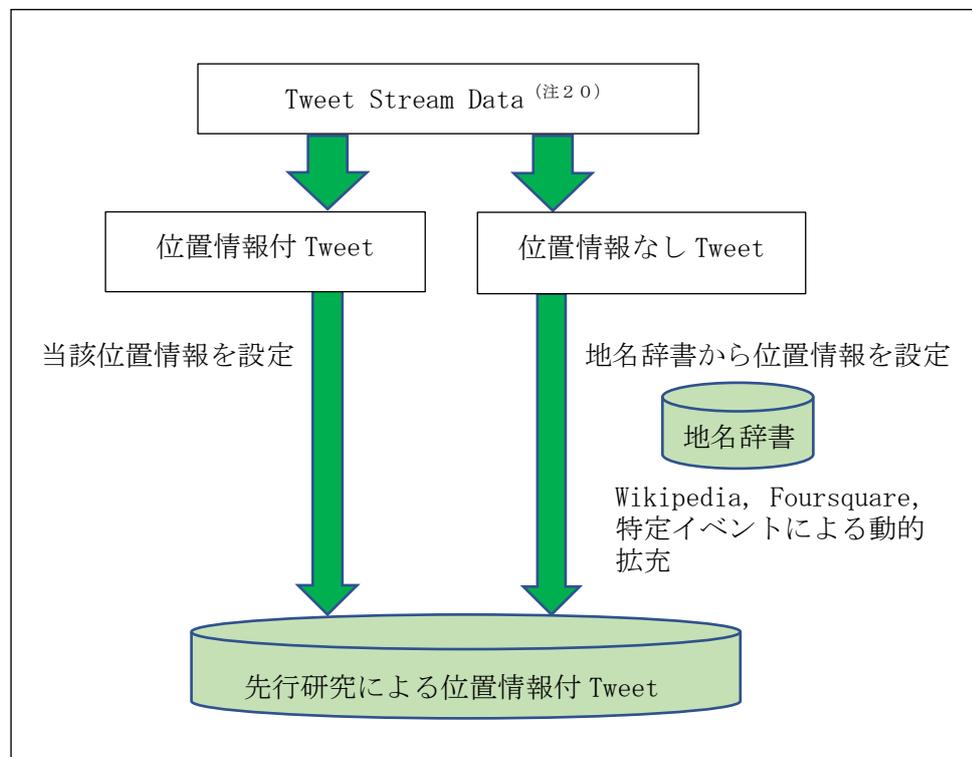


図 34 推定位置情報付加の概要

(注 19): <https://ja.foursquare.com/>

(注 20): 東京大学、喜連川・豊田研究室で 2011 年 3 月 11 日から Twitter API を用いて収集したツイートデータを用いている。ツイートの収集は 30 人の著名 Twitter ユーザを初期ユーザセットとして過去の投稿を取得するところから開始している。その後、彼らのタイムライン上でリツイートやメンションが行われたユーザを新たな対象としてユーザセットに加え、同様に過去の投稿を取得する。このように、ユーザセットの拡大とタイムラインのトラッキングを繰り返し、2018 年 11 月 18 日時点で約 574 億ツイートを蓄積している[27]。

### 4.3.2 有効 Tweet 特定の処理概要

本研究で使用する Tweet は、「推定言語表現を含み、Tweet 時に発生していた降雨を示す Tweet」とし、以下、これらを「有効 Tweet」と呼ぶ。

有効 Tweet の特定方法について説明する。まず、先行研究により推定位置情報が付加された位置情報付 Tweet から次の降雨関連用語を含む Tweet（以下、「降雨関連 Tweet」と呼ぶ。）を抽出する。

#### 降雨関連用語

雨（豪雨、大雨等の雨を含む単語）、土砂降り、どしゃ降り、夕立、スコール、小降り、本降り、ザーザー降り、台風、雹（ひょう）、霰（みぞれ）

次に、これらの降雨関連 Tweet から有効 Tweet を特定する。具体的には、降雨関連 Tweet から本研究で扱う推定言語表現が含まれる Tweet を抽出し、その中から更に、Tweet が示す降雨情報が Tweet 時点に発生していたものと推定できる Tweet を絞り込む。これらは本研究で構築した独自の抽出手法を使用している。この処理過程を図 35 に示す。

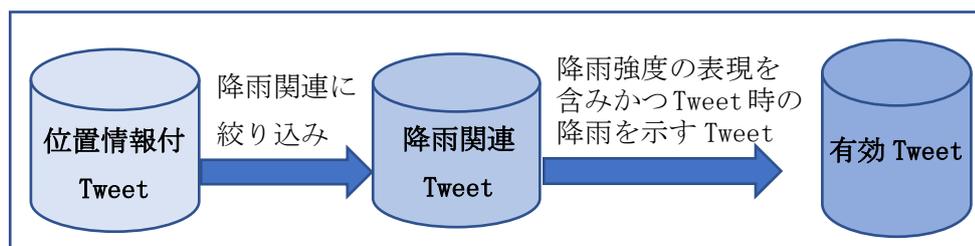


図 35 有効 Tweet 抽出の処理過程

降雨関連 Tweet から有効 Tweet を抽出する手法は、2015～2017 年の降雨関連 Tweet を利用して開発した。以降、推定言語表現の特定方法、開発した抽出手法の詳細および本手法により抽出された有効 Tweet の精度に関する検証結果について説明する。

### 4.3.3 推定言語表現の特定方法

本研究のモデルで使用する推定言語表現の特定方法を述べる。まず、本研究で開発した分類器を使用して 2015～2017 年の降雨関連 Tweet から推定言語表現の候補を抽出する。次に、その中から「高精度で降雨強度を示すと考えられる言語表現」を絞り込み、最終的に本研究で使用する推定言語表現を決定する。

分類器の仕様について概要を述べる。分類器には、まず、4.3.2 節で挙げた降雨関連用語を初期登録しておく。その上で、降雨関連 Tweet を入力データとし、Tweet を 1 件ずつ形態素解析して分類器内に登録済の推定言語表現と照合する。そして、照合結果 (Tweet 内における推定言語表現の有無に関する結果) を出力すると共に、本研究で使用する推定言語表現として追加登録の候補となり得る言語表現を併せて出力する。

分類器を使用した推定言語表現の特定における具体的な手順を説明する。

1. 4.3.2 節で挙げた降雨関連用語を分類器に初期登録する。
  2. 分類器を使用して対象年の降雨関連 Tweet から推定言語表現の登録候補を抽出する。分類器の仕様は以下のとおり。
    - (1) 形態素解析された降雨関連 Tweet 内の各言語表現を配列  $\text{array}[x]$  に割り当てる。ここで  $x$  は形態素解析された要素の先頭からの連続番号とし、 $0 \leq x < n$  とする ( $0$  を先頭、 $n$  は形態素解析された要素数)。
    - (2)  $\text{array}[x]$  について、条件①～④ (次頁参照。以下、同様。) と比較する。この結果、いずれかの条件を満たす場合は、当該降雨関連 Tweet をファイル (File1) に出力する。
    - (3) 条件①～④のいずれも満たさない場合は、当該 Tweet をファイル (File2) に出力し、併せて、当該 Tweet 中の全ての名詞、形容詞、副詞を「追加登録の候補である推定言語表現」として File3 に出力する。
-

条件①：以下のいずれも満たす場合

例：「豪雨」「大雨」「本降り」「スコール」)

- array[x]の品詞＝名詞
- array[x]が推定言語表現に登録済

条件②：以下の3条件をいずれも満たす場合

例：「すごい豪雨」「雨が弱い」

- array[x]の品詞＝形容詞
- 降雨関連用語の状態を示すと判定できる

(判定条件の一例)

array[x+1] = 降雨関連用語

array[x-2] = 降雨関連用語かつ array[x-1]の品詞＝助詞

- array[x]が推定言語表現に登録済

条件③：以下の3条件をいずれも満たす場合

例：「ポツポツ降り出した」「雨がザーザー」

- array[x]の品詞＝副詞
- 指定の表現(\*. 以降同様.) または降雨関連用語の降雨強度を示す

(\* 指定の表現：降る，降り出す，降り続く，降っている等

(判定条件の一例)

array[x+1] = 指定の表現(\*)

array[x-2] = 降雨関連用語 かつ array[x-1]の品詞＝助詞

- array[x]が推定言語表現に登録済

条件④：以下を満たす場合

(例：「滝のような」「バケツをひっくり返す」)

- array[x]～array[x+3]の連結＝登録済の推定言語表現

3. File3 に出力された登録候補の登録要否を検証し，追加登録が必要の場合は分類器に手動追加する（例えば，「冷たい」「暖かい」「優しい」等は降雨強度を示さないため追加しない），この2. および3. を繰り返し，追加分がこれ以上ないと判断できるまでデータ対象全てについて継続実施する。

#### 4.3.4 時点の絞り込み

2015～2017年の関東地域の降雨関連 Tweet を利用して、Tweet 時に発生していたと考えられる降雨を示す Tweet を抽出する手法を開発した。具体的には、次のいずれかの条件を満たす Tweet を有効 Tweet とし、いずれも満たさない Tweet は除外する。

- ・ 指定の表現（前述の条件③参照）+ 記号（「」以外）  
例：大雨が降っている。（「降っている」+「。」）等  
※形態素解析された各要素（「降る」（動詞）、「て」（助詞）および「いる」（動詞））を結合して条件判断に使用。以降、同様。
- ・ 降雨関連用語+記号（「」以外）  
—— 豪雨！（「豪雨」+「！」）等
- ・ 降雨関連用語と助動詞，助詞および記号（「」以外）の組み合わせ。  
—— どしゃ降りです。（「どしゃ降り」+「です」+「。」），  
また雨が！（「雨」+「が」+「！」）等
- ・ 降雨関連用語，助詞，形容詞または副詞，記号の組み合わせ。  
—— 雨がすごい。（「雨」+「が」+「すごい」+「。」）等。  
雨がポツポツ。（「雨」+「が」+「ポツポツ」+「。」）等。

本手法による抽出結果からは、過去（「降った」「降っていた」）、推量（「降りそう」「降るかも」「～らしい」「～みたい」等）の表現は除外される。

この効果として、例えば、東京で九州地区の豪雨を呟く等の場合（「九州では猛烈な雨が降っているらしい。」等）は推量の表現のため除外される。このため、結果的に、Tweet が示す降雨の場所が発信地点を示さない Tweet を除外することが可能となっている。

### 4.3.5 有効 Tweet の精度に関する検証

有効 Tweet の抽出手法について、2018 年のデータを利用して精度検証を行った。本検証の対象は、先行研究を利用して付加した推定位置情報および抽出された Tweet の内容の正確性である。

まず、推定位置情報の正確性は、2018 年の月別の有効 Tweet 数が最も多い 2018 年 9 月における有効 Tweet (計 507 件) を基に検証した。検証方法としては、先行研究により付加された位置情報 (緯度経度) を Google Map 上に展開し、当該位置が Tweet の示す場所と一致しているか (ジオタグの場合は Tweet の示す場所の一部、場所名称による推定の場合は当該場所の中心地であること) を確認した。この結果を先行研究における推定位置情報付加条件の区分単位で表 5 に示す。

表 5 有効 Tweet の位置情報精度 (2018 年 9 月)

区分	件数計	正答件数	誤件数	Precision
ジオタグ	64	63	1	0.984
場所名称による推定	443	415	28	0.937
バーストワードによる推定	0	0	0	-

表 5 のとおり、先行研究により付加された推定位置情報は 0.93 以上の高精度を得る結果となった。ジオタグにより付加された推定位置情報は高精度の結果が得られた。場所名称により付加された推定位置情報も高精度と言えるが、Tweet 内に複数の場所の名称が含まれる場合に、Tweet が示す場所と推定位置情報が一致しない例が少数ながら存在した。これは、最初に出現した場所名称で位置を推定する先行研究の仕様による。また、バーストワードによる推定は先行研究では極少数事例であり、本結果でも抽出されなかった。

以上の結果から、不一致事例は存在するが、全体精度に与える影響は低いと考えられるため、先行研究による推定位置情報は高精度であると言える。

また、今回検証した 2018 年 9 月における事例を参照すると、対象となる有効 Tweet 数はジオタグ付きの Tweet に限定する場合に比べ約 8 倍に拡大された。以上により、先行研究の利用による推定位置情報は高い正確性と対象 Tweet 数の拡大に有効であると結論付けられる。

---

次に、有効 Tweet における発信内容の正確性について、Tweet 件数の多い夏期（2018年7～9月）のデータにより検証を行った。当該データ期間は、7月には西日本豪雨（「平成30年7月豪雨」）が発生し、8月には台風発生数が1951年からの統計史上2番目となり、9月には近畿地方に甚大な被害をもたらした台風21号の上陸等により、2018年の降雨関連 Tweet 数が上位3位となった月である。この結果を表6に示す。

表6 有効 Tweet の内容に関する精度(2018年7～9月)

月	降雨関連 Tweet (件)	TP	FP	FN	TN	Precision	Recall	F 値
7月	14605	340	24	1667	12574	0.934	0.169	0.29
8月	12144	349	19	1043	10733	0.948	0.251	0.4
9月	19612	461	46	2336	16769	0.909	0.165	0.28

表6の Precision（適合率）を参照すると、平均して0.93を超えている。しかし、FN 値を参照すると、誤って除外した事例は相当数存在し、その割合は3ヵ月間の平均で約9%だった。この要因を検証すると、Tweet 内に省略や個別の言い回し等がなされ、本来の言語表現とは乖離するものが中心であった。このような表現には汎用的な条件設定が困難であるため、現状は「表現が不適切な（内容を機械的に判読不能な）Tweet」として一律除外せざるを得ない。

このように除外された Tweet 数の影響により、Recall（再現率）およびこれと連動する F 値は低い。しかし、本研究における降雨強度レベルの推定では物理情報相当の正確性を目指しているため、信頼性に疑念が生じ得る Tweet は極限まで排除する必要がある。つまり、本研究では有効 Tweet の正確性を高水準に引き上げること（Precision の高さ）を最も重要視し、抽出条件を極めて厳格に設定した。この結果、有効 Tweet の Precision は対象の3ヵ月間で高精度な値を示し、高い正確性を実証できた。従って、これらの Tweet は降雨強度レベルの推定を精緻に行うために有効であると結論付けられる。

有効 Tweet の抽出条件の工夫により対象範囲の拡大（Recall 値の向上）は可能である。Precision とのバランスは容易でないが、将来的な課題として、機械的判定の改善を試行しながら全体的な精度向上を目指す扱いとする。

---

## 4.4 降雨強度推定モデルの構築

本セクションでは、推定言語表現から降雨強度を取得するモデルの構築方法を説明する。まず、Tweet 内に含まれる推定言語表現が示す降雨強度のレベル値について、過去の Tweet と X バンド MP レーダの値（降雨強度レベルに換算）を照合して調査した。この結果、当該言語表現毎に各月に出現するレベル値を観察すると、一定のレベル値のみが検出される推定言語表現と、1~5 まで出現するレベル値に幅がある推定言語表現等、推定言語表現毎に示されるレベル値に一定の傾向があることがわかった。このため、推定言語表現を当該言語表現の示す統計的な精度に応じて 3 つの Category に分類し、より正確な推定手法を試行することとした（後述 4.4.3）。

降雨強度には大別して 2 種類の変動パターンが存在する。具体的には、弱雨からピークを迎え収束する「時間的な変動」と、冬に小さく夏に大きい傾向がある「季節的な変動」である。従って、推定言語表現から降雨強度のレベル値を推定する場合、両変動を考慮した仕組が必要と考えられる。本研究では両変動の傾向を捉えるため、Tweet の発信時刻と当該時刻前後に発生していた降雨強度を分析した。そして、過去の Tweet と X バンド MP レーダを照合しながら推定言語表現の示す降雨強度のレベルおよび変動傾向を学習し、最終的に「Category」、「推定言語表現」および「月」を要素とする降雨強度推定モデルを構築した。本セクションでは当該構築方法を説明する。

### 4.4.1 Tweet が示す対象降雨の絞り込み

一般的に、Tweet 発信時には入力等による時間差が生じる。このため、Tweet と照合する X バンド MP レーダは発信時刻時点のデータのみでなく、時間差を考慮した範囲の X バンド MP レーダのデータを含める必要がある。

本研究では、Tweet の発信は降雨状況の変化に誘発されると考え、Tweet の発信時刻と降雨群のピーク時刻との差を検証して当該時間差の傾向を把握することとした。検証方法は、2015 年の関東地域の降雨関連 Tweet (133,254 件) から有効 Tweet を抽出し、当該 Tweet に対応する X バンド MP レーダ（発信時刻から 30 分遡及）を抽出してピーク値の時刻を取得し、Tweet 発信時刻との差を 5 分単位で集計した。この集計結果が全体に占める割合(%)を図 36 に示す。

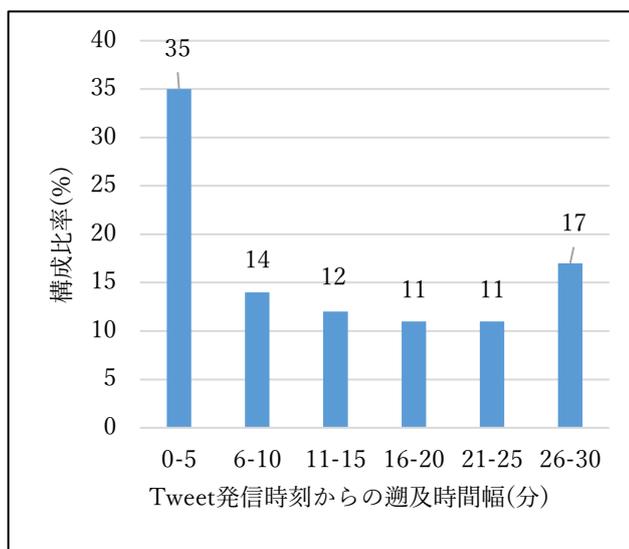


図36 Tweet 発信時刻と降雨ピークの時間差

図36によると、Tweet全体の約3分の1は降雨ピークから5分以内の発信であった。5分以上遡ると急激に割合が減少し、以降は同様の傾向を示す。また、30分まで遡ると再び発信件数に上昇が見られるが、これらは別の降雨群が発生している可能性が考えられる。この結果から、Tweetの示す降雨の多くは当該Tweet発信時刻から遡及5分間以内に集中する傾向があると言える。しかし、全体に占める割合は3割強程度と半数に届かないため、約半数となる「発信時刻から遡及10分間」に発生した降雨データを調査対象とすることが適切と考えた。このため、本研究においてTweetと照合するXバンドMPレーダデータは「Tweet発信時刻から遡及10分間に発生したXバンドMPレーダデータ」とする。

#### 4.4.2 推定言語表現とXバンドMPレーダデータと照合

本研究では、推定言語表現毎に示される実際の降雨強度の傾向はXバンドMPレーダの過去の実績を学習することで把握可能と考えた。従って、本研究では、推定言語表現を含む過去のTweetと当該Tweetに対応するXバンドMPレーダデータ（降雨強度レベルに換算）を照合した。以降、この実施手順および照合結果を説明する。

まず、有効 Tweet を抽出し、当該 Tweet 毎に照合対象とする X バンド MP レーダデータを遡及時間分（X バンド MP レーダは毎分データのため遡及 10 分では 10 件となる）取得する。

次に、それらの X バンド MP レーダデータから当該 Tweet に含まれる推定言語表現の示す降雨強度レベルを選定する。この選定基準は 3 パターンを検討した。具体的には、遡及時間範囲の X バンド MP レーダデータの選定基準を①最大値、②平均値、③最頻出値から検討する。参考として、「小雨」および「豪雨」の事例を図 37 に示す。

Tweet 時刻: 10:30 「小雨」				Tweet 時刻: 11:10 「豪雨」			
XRAIN				XRAIN			
時刻	雨量	レベル		時刻	雨量	レベル	
10:30	1mm/h	1		11:10	25mm/h	3	
10:29	3mm/h	1		11:09	29mm/h	3	
10:28	2mm/h	1		11:08	25mm/h	3	
10:27	1mm/h	1		11:07	37mm/h	4	
10:26	4mm/h	1		11:06	34mm/h	4	
	:				:		
	:				:		
(以下, 省略)				(以下, 省略)			

図 37 遡及時間内の X バンド MP レーダのレベル値

ここで図 37 を参照すると、「小雨」を含む Tweet に対応する X バンド MP レーダの降雨強度レベルは遡及時間内で全て「レベル 1」であり、Tweet 毎の降雨強度レベルは選定基準①～③のいずれも「レベル 1」である。従って、推定言語表現「小雨」の示す降雨強度レベルは一意に定まる。一方「豪雨」は、降雨強度レベルの結果が分散し、選定基準でも①レベル 4、②レベル 3、③レベル 3 となり一意に特定できない。

このため、推定言語表現毎の降雨強度レベルの傾向を月間で統計的に検証する。一例として、「小雨」「豪雨」の事例において 2015 年 8 月の有効 Tweet を基に選定基準①で Tweet 毎に降雨強度レベルを抽出した。そして、抽出した降

雨強度レベルを月間で集計し、全体における割合を月間構成比率として表した。  
この結果および Tweet 毎の降雨強度レベルのサンプルを図 38 に示す。

「小雨」			「豪雨」		
日付	時刻	レベル	日付	時刻	レベル
2015/08/22	14:16	1	2015/08/14	18:40	4
2015/08/27	09:49	1	2015/08/14	18:50	1
2015/08/27	11:00	1	2015/08/14	19:03	3
2015/08/27	11:07	1	2015/08/14	19:06	3
2015/08/27	12:33	1	2015/08/14	19:19	5
(以下, 省略)			(以下, 省略)		
月間構成比率:			月間構成比率:		
レベル 1: 100 %			レベル 1: 14.7%		
			レベル 2: 29.4%		
			レベル 3: 32.4%		
			レベル 4: 14.7%		
			レベル 5: 8.8%		

図 38 降雨強度レベルの推定言語表現別・月別集計(2015年8月)

2015年8月のデータでは「小雨」に該当した降雨強度レベルは月内全て「レベル1」であった。一方、「豪雨」では、図38のサンプルのとおり降雨強度レベルが分散している。2015年8月で最大件数となったレベル値は「レベル3」であったが、当該月での構成比率は約3割と低く、「豪雨」の示す降雨強度レベルが「レベル3」と特定することは難しい。

次に、前述の推定言語表現「小雨」および「豪雨」について、月間での構成比率の年間傾向を調査した。対象データは2015~2017年の3年間の有効 Tweet とし、各月で最も多く出現した降雨強度レベルと当該レベルが全体に占める割合を月間構成比率として表し、結果を表7に示す。

---

表7 降雨強度レベル値の年間傾向(2015~2017年)

	小雨		豪雨	
	レベル	月間構成比率 (%)	レベル	月間構成比率 (%)
1月	レベル1	100	レベル1	50
2月	レベル1	98	レベル1	67
3月	レベル1	100	レベル1	100
4月	レベル1	100	レベル1	71
5月	レベル1	100	レベル2	54
6月	レベル1	100	レベル2	62
7月	レベル1	100	レベル3	52
8月	レベル1	100	レベル3	48
9月	レベル1	100	レベル3	62
10月	レベル1	100	レベル2	56
11月	レベル1	100	レベル1	100
12月	レベル1	100	レベル1	85

表7では、「小雨」の降雨強度レベルは全て「レベル1」を示し、月間構成比率も96%以上であった。一方、「豪雨」は降雨レベル値に1~3の幅があり、夏に大きい値を示すなどの季節的な変動が確認できる。この結果から、本研究では、言語表現から推定値を一意に求めるより、季節の要素を含めた条件で細分化し決定する方がより高精度の結果を得られると考えた。この細分化の粒度は「月」や「週」が考えられるが、「週」は細かく繁雑になるほか各週のデータが少なく、季節的な降雨変動の特徴を捉えにくい。このため、本研究では降雨強度の推定単位を「月」とし、降雨強度レベルを推定するモデルを「推定言語表現」と「月」を要素として構築することとする。

また、表7の「豪雨」は、特に8月の月間構成比率が低い。本研究では、このような月間構成比率が低い推定言語表現は人による感覚差が大きい表現であり、一定の基準を下回る場合は推定材料として適さないと判断した。

従って、本研究で推定に利用する推定言語表現は、各月の月間構成比率が50%以上得られた推定言語表現に限定する。例えば上記事例の場合、8月の推定対象の言語表現から「豪雨」は除外される。

また、今回の推定言語表現は関東地域の有効 Tweet により構築し関東特有の特徴を含むため、他地域の降雨強度レベル推定には用いない。地域により気象の差や固有の表現（方言）等が考えられることから、推定言語表現は地域単位で抽出し特定する扱いとする。

#### 4.4.3 モデルの要素

本研究における降雨強度レベルの推定モデルは「推定言語表現」と「月」を要素として構築する。

ここで、月内の分散したレベル値から月毎のレベル値を選定するには次の2種の方法が考えられる。

- (i) 同月内で最大件数のレベル値
- (ii) 同月内で中央値を示すレベル値

従って、推定言語表現毎・月毎に降雨強度レベル値を決定する方法（以下、「推定値設定パターン」と呼ぶ。）は、Tweet 単位での X バンド MP レーダの選定方法（3 パターン）と月間の推定言語表現単位のレベル値選定方法（2 パターン）を組み合わせた以下 A～F の 6 パターンを考える（表 8 参照）。

表 8 推定値設定パターン

XRAIN の選別	月内抽出方法	
	最大件数	中央値
最大値	A	B
平均値	C	D
頻出値	E	F

加えて、推定言語表現の月間構成比率により最適パターンは異なる可能性が考えられる。このため本研究では、本研究において扱う全ての推定言語表現を2015～2017年の月間構成比率の年間平均値を基準として以下の3つのCategoryに分割した（表9参照）。

表9 推定言語表現の分類

Category	月間構成比率 (年間平均)	対象推定言語表現数
Category 1	95%以上	63語（小雨，霧雨，弱い，本降り等）
Category 2	85%以上 95%未満	12語（暴風雨，横殴り，大粒等）
Category 3	50%以上 85%未満	10語（豪雨，土砂降り，雷雨等）

表9のCategory 1には年間の平均降雨強度がレベル1の弱雨を示す推定言語表現が全て含まれた。しかし、同強度がレベル2以上の強い降雨を示すことが多い推定言語表現はCategory 2～3に分散され、特にCategory 3ではレベル3以上を示す推定言語表現が多い。この結果から、降雨強度が小さい推定言語表現は降雨強度レベルの月間構成比率が高く、降雨強度レベルをほぼ一意に推定可能と考えられる。一方、降雨強度が大きくなると人間の感覚によるバラつきが多くなる。従って、本研究のモデルではCategory毎に推定値設定パターンを決定し、推定言語表現別・月別の設定値（推定値）を設定した。

#### 4.4.4 推定値設定パターンの決定

本セクションでは、最終的に決定したCategory毎の推定値設定パターンと当該特定手順を述べる。

まず、2015～2017年の降雨関連Tweet（計386,357件）から抽出した有効Tweetとこれに対応するXバンドMPレーダデータを利用し、表8に基づく全6パターンのモデルを構築した。次に、2017年の有効Tweet（78,988件）に全てのモデルを適用し、得られた推定値とXバンドMPレーダの実績値を照合した。この照合結果を表10に示す。また、Category毎に最も高精度となったパターンを表11に示す。

---

表 10 一致率の年間平均 (2017 年)

パターン	推定言語表現の分類		
	Category 1	Category 2	Category 3
A	0.892	0.765	0.642
B	0.812	0.701	0.572
C	0.941	0.804	<b>0.785</b>
D	0.915	0.791	0.658
E	<b>0.988</b>	<b>0.829</b>	0.731
F	0.974	0.804	0.677

※ 一致率= 月内一致件数 / (月内一致件数+月内不一致件数)

表 11 Category 毎の推定値設定パターン

Category 1	Category 2	Category 3
パターン E	パターン E	パターン C

#### 4.4.5 Tweet 件数における人口密度の影響

本手法におけるモデルは関東地域を対象とし、地域全体を Tweet の集計単位とした。しかし一般的に、集計単位を地域全体とする場合、人口が少ない県より都市部の Tweet 数が優先され、都市部の特徴を強く表すレベル設定となりやすい。この場合、当該レベルは人口の少ない県の特徴と合わず全体の精度を下げる可能性がある。従って、レベル設定の際は、地域全体での集計結果と各県の人口密度を考慮して重みづけをした集計結果の比較を行うことが必要と考える。このため、本研究の事例で同比較を実施してみたが、関東地域では両者の比較に違いが現れず、結果として人口密度による重みづけの考慮は不要と判断した。しかし、例えば、中部地域や東北地域等、太平洋側と日本海側とで気候が異なる地域では、該当地域全域での集計と人口密度の調整を行った集計とで結果が異なり、県毎の特徴が顕現化する可能性がある。この場合は、集計単位を地域別ではなく、県別あるいは緯度経度を利用した一定の区分等に細分化したレベル設定とする等の考慮が別途必要になると考えられる。

## 第4章

### 言語表現を用いた降雨強度レベル推定手法

今回は人口密度の考慮を不要としたが、今後の他地域におけるレベル設定を鑑み、参考として関東地域のデータに対して実施した人口密度による影響確認の手順および結果を本節に示す。

まず、関東地域の人口密度を表12<sup>(注21)</sup>に示す。この表を参照すると、東京と栃木県等、都県間で約20倍の差が生じる事例がある。

表12 人口密度に応じた按分件数

県名	人口密度(人/m <sup>2</sup> )	JX：調整値
東京	6168.7	1
神奈川	3777.7	1.63
埼玉	1913.4	3.22
千葉	1206.5	5.11
茨城	478.4	12.89
群馬	310.1	19.89
栃木	308.1	20.02

そこで、統計上の信頼性の観点から、レベル毎のTweet数の集計時に都県の人口密度に応じて按分した場合の影響を調査した。以下に表す式のとおり、按分件数(RX)は、Tweetの位置情報(緯度経度)に応じた都県Xの件数KXに人口密度に応じた調整値(表12：JX)を乗じて算出する。

$$RX = KX \times JX ※ \quad ※表12参照$$

按分した件数に基づきモデルを再構築した結果、再構築前の設定値からの変更率は全体の約6%に留まった。従って、年間平均一致率も人口密度を考慮せず計算した場合とほぼ変わらない。これは関東地域に気象差および人の感覚差が少ないことが理由の一つと考えられる。この結果から、関東地域では人口密度に伴う統計的影響が少ないため、今回の研究では考慮しない。しかし、今後別地域でモデル構築する際には改めて同調査を行い、適切な集計単位(県別等)を検討する。

(注21)：総務省統計局「平成27年国勢調査(人口等基本集計結果)」

<http://www.stat.go.jp/data/kokusei/2015/kekka/kihon1/pdf/gaiyou1.pdf>

#### 4.4.6 降雨強度推定モデル構築手法のまとめ

本モデルの概要を総括する。本モデルは2015~2017年の3年間の降雨関連 Tweet (計 386,357 件) および X バンド MP レーダを利用し、適用結果の年間平均が最も高精度となるモデルとして構築した。構築手順は次の①~④のとおりとなる。なお、本研究では関東地域を対象としたが、本構築手順は全地域共通であり、対象の地域単位で行う扱いとする。

- ① 分類器を用いて推定言語表現を特定する。
- ② 過去の Tweet と X バンド MP レーダの実績を照合し、①の推定言語表現の月間構成比率を調査する。なお、X バンド MP レーダの抽出範囲は Tweet 発信時刻から遡及 10 分間とする。
- ③ ②の結果を基に推定言語表現を 3Category に分類し、各月の月間構成比率が 50%以上のものを利用する。
- ④ 過去の Tweet と X バンド MP レーダの実績を利用して推定値設定の全パターンを試行し、当該分類毎に最適なパターンを決定する（本研究における設定パターンは表 11 のとおり）。

### 4.5. 適用結果および考察

#### 4.5.1 2018 年の事例における適用結果

4.4. で構築したモデルを 2018 年の事例に適用した。手順は、2018 年の降雨関連 Tweet (139,694 件) から有効 Tweet を特定し、これらの有効 Tweet に本モデルを適用して得られた推定値（降雨強度レベル）と X バンド MP レーダの実績値（降雨強度レベルに換算）を照合する。照合した結果、一致した Tweet 数を月単位で集計し、当該月の有効 Tweet 数との割合（月別平均一致率）で表した（表 13）。

表 13 の各 Category の全体に占める割合を参照すると、Category 3 の推定言語表現の割合が少ないほど一致率が高く、夏期の一致率は低下していることがわかる。しかし、一致率の月間平均は 12 ヶ月中 9 ヶ月で 0.9 超となり、年間平均でも 0.93 を超え、全体的に高精度の結果が得られたと言える。

表 13 本モデルの月別適用結果 (2018 年)

対象月	一致率の 月間平均	Category 毎の全体に占める件数の割合(%)		
		Category 1	Category 2	Category 3
1 月	0.952	95	4	1
2 月	1.000	98	1	1
3 月	0.974	97	2	1
4 月	0.860	88	10	2
5 月	0.910	82	14	4
6 月	0.988	80	18	2
7 月	0.908	52	22	26
8 月	0.843	33	34	33
9 月	0.877	49	31	28
10 月	0.900	67	28	5
11 月	0.977	96	2	2
12 月	1.000	99	1	0
平均	0.932	---	---	---

## 4.5.2 モデルの構成要素有無に関する考察

### 4.5.2.1 実験内容

本モデルの構成要素である「月」および「Category 毎の推定値設定パターン」を含まない場合の手法を考察し、構成要素の必要性を評価する。具体的には、次の実験 1 および 2 による一致率を本モデルによる結果と比較して評価する。

実験 1：月別とせず推定言語表現毎に一意のレベル値を設定する方法

実験 2：Category 別にせず共通の推定値設定パターンでレベル値を設定する方法。

実験 1 では、2015～2017 年の各 Tweet から推定言語表現を取り出し、10 分間中の X バンド MP レーダの頻出値から推定言語毎の降雨強度レベルを決定し、3 年間分集約した中で最大件数となるレベル値を当該推計言語表現のモデル値とする。

次に、実験2の手順を述べる。表10の結果をCategory別にせずに算出した年間一致率の結果を表14に示す。

表14 一致率年間平均のパターン別比較 (2017年)

A	B	C	D	E	F
0.866	0.843	0.926	0.893	0.933	0.928

表14を参照すると、パターンEの一致率が最も高い。従って、全ての推定言語表現のレベル値をパターンEにより設定して実験した。

#### 4.5.2.1 実験結果および考察

2018年の有効Tweetに各実験におけるモデルを適用し、得られた推定値をXバンドMPレーダの実績値と照合した。この結果を表15に示す。

表15 モデルの条件変更による一致率の比較 (2018年)

対象月	実験1	実験2	本モデル
1月	0.902	0.948	0.952
2月	0.988	1.000	1.000
3月	0.912	0.952	0.974
4月	0.688	0.841	0.860
5月	0.782	0.880	0.910
6月	0.891	0.961	0.988
7月	0.640	0.766	0.908
8月	0.582	0.712	0.843
9月	0.613	0.743	0.877
10月	0.722	0.877	0.900
11月	0.862	0.965	0.977
12月	0.971	1.000	1.000
年間平均	0.796	0.887	0.932

実験1では、Category毎に異なる結果となった。まず、Category1の推定言語表現における設定値は年間を通じてほぼ同一のため、条件変更による一致率への影響はなかった。しかし、Category2およびCategory3ではレベル値を年間で一つに集約すると季節的な降雨変動が反映されず、冬期または夏期を中心に実際のXバンドMPレーダと不一致となる事例が大幅に増えた。従って、「月」を要素としてモデルを構築する手法は有効と考えられる。

実験2では、推定値設定パターンがCからEに変更となったCategory3の推定言語表現のみ設定値に変更が生じ、Category3の推定言語表現が使われる割合が多い月（7～9月）を中心に一致率が低くなった。Category3の推定言語表現は、降雨の時間的な変動幅が大きいですが、この変動を捉える手法では頻出値を採用する手法（パターンE）より平均値を採用する手法（パターンC）の精度の方が高い。従って、Category毎に推定値設定パターンを割り当てる手法は有効と考えられる。

以上の結果から、本モデルの構成要素の「月」および「Category」の必要性を実証できた。

### 4.5.3 精度に関する考察

これまでの検証結果のとおり、本推定値はXバンドMPレーダと高精度で一致する結果を得た。この精度を他の指標と比較する。

まず、一般的な公開気象情報と比較する。本研究の局地的かつリアルタイム配信の条件に最も近い情報としては、気象庁が公開する「高解像度降水ナウキャスト」(250mメッシュ・5分毎1時間までの予測)がある。当該精度は閾値が大きいほど予測精度が低くなるが、閾値1mm h<sup>-1</sup>における5分後のスレットスコア<sup>(注22)</sup>の値は最も高い場合で約0.8弱とされる[28]。このため、本研究における推定値の時間的優位性を差し引いても、同予測精度とほぼ同等以上の精度と言える。従って、本研究でのSNSの言語表現に基づく推定値は一般的な気象情報として有意性があると考えられる。

---

(注22):スレットスコア:予報,または実況で「現象あり」の場合の予報適中事例数に着目して予報精度を評価する指標。

気象庁「数値予報研修テキストで用いた表記と統計的な指標」

[https://www.jma.go.jp/jma/kishou/books/nwptext/42/Appendix\\_B.pdf](https://www.jma.go.jp/jma/kishou/books/nwptext/42/Appendix_B.pdf)

次に、防災上の数値指標として比較する。防災上の数値は一般的に大雨の予測に関する適中率により示されるが、気象庁では、局地モデル (LFM) (注23) を中心に解析される。しかし、当該モデルの精度は閾値が  $1 \text{ mm h}^{-1}$  以上かつ検証格子  $10 \text{ km}$  の条件下でのスレットスコア値として約  $0.4$  前後とされ[29]、現状では高精度の予測は難しい。当該モデルは予測値であり本研究とは時間幅および対象面積も異なるため単純比較はできない。しかし、現状の物理的観測情報以外に本研究による推定値を補助的な情報として組み合わせることは、総合的な精度向上のための選択肢の一つと考えられる。

## 4.6. 第4章のまとめ

本研究では、Tweet 内の降雨関連の言語表現が示す実際の降雨強度について、過去の膨大なデータの学習により、言語表現毎の特徴や同じ言語表現でも季節毎に示される強度の違いを明らかにし、言語表現から降雨強度レベルを推定する手法を提案した。推定元となる Tweet は、先行研究の利用により推定位置情報を付加して空間的精度を高め、独自の抽出手法により時間的精度を高めた Tweet とした。また、推定モデルは、推定言語表現を月間構成比率に応じて分類し、当該分類毎に推定値の算出方法を設定して、推定言語表現および月を要素とした手法として構築した。そして、適用例では実際の X バンド MP レーダの実績値と高精度で一致する結果を得た。

本提案手法は地域毎に推定言語表現を抽出し、当該地域の X バンド MP レーダとの照合結果から推定値を得る。地域毎としたのは、各地域で言語表現 (方言等)、気候差、地形、気候に対する感受性および人口密度の分布等により、示される降雨強度が異なる可能性を勘案したことによる。本手法により適用地域のデータを利用してモデルを構築すれば、どの地域でも降雨強度の推定が可能となる。従って、本稿では関東地域の事例としたが、本手法を利用して他地域における降雨情報の推定を実装することにより、地域間での気候や表現の違いを検出可能と考える。

(注23): 局地モデル (LFM) : 防災気象情報作成支援、航空予報支援を主な目的とする水平  $2 \text{ km}$  メッシュでの高解像度な数値予報モデル (毎時データ)。発達した積乱雲を直接表現することができ、集中豪雨や突風など局地的な現象を主な予測対象とする。

気象庁予報部数値予報課 「局地モデル LFM のモデルの特性と利用上の留意点について」  
<https://www.jma.go.jp/jma/kishou/minkan/koushu131120/shiryoul.pdf>

## 第4章

### 言語表現を用いた降雨強度レベル推定手法

---

また、本研究では、モデル構築のための学習として多くの推定言語表現と X バンド MP レーダとの比較検証を行った。この過程で、特定の推定言語表現において示される降雨強度に、今後の応用的利用方法が期待できる例も検出された。一例として、「雹（ひょう）」および「霰（みぞれ）」の推定言語表現は、特に夏において、強度の大きい降雨として観測される傾向が多いことを検出した。X バンド MP レーダは雨粒の大きさを観測して降雨の強度を推定するため、粒度が大きい雹や霰は強度が大きいと見なされると考えられるが、これらは一般的な降雨と比べ物理的破損（屋根、車等）や怪我の危険性が高い。このため、これらの推定言語表現の場合は水害よりむしろ物理的破損の危険性を知らせる等、推定言語表現により通知内容を振り分ける方式も有効と考えられる。

また、本研究では現時点の降雨を示す言語表現を利用し、リアルタイムでの降雨強度レベル推定を行う手法として提案した。しかし、例えば、「もうすぐ雨が降りそう。」等の未来の事象を差す表現と X バンド MP レーダの実績値との学習により、予測推定に応用可能とも考えられる。このほか、投稿情報の利用による天気情報の表示の詳細化、物理情報の観測対象外地域における代替値としての利用等も期待できる。

このように、本研究の成果は言語表現から物理情報を推定する手法の一つとして有効であり、応用的な利用も期待できると考える。

## 第5章. 本研究のまとめ

本研究では、「局地的な降雨情報の迅速な情報共有」を目的として3つのテーマにより研究を行った。

第一に、XバンドMPレーダデータを可視化し、誰もがリアルタイムで局地的な降雨情報を閲覧可能な環境を構築し公開した。本研究では膨大な量のXバンドMPレーダデータを毎分遅延なく処理および表示する対策を講じ、高リスク値を優先する手法を提案した。これにより、防災上最も必要となる高リスク情報を迅速に表示しながらXバンドMPレーダの情報共有を行う基盤を構築することができた。第二に、国内に存在するXバンドMPレーダの観測範囲外地点におけるデータの補完手法として、Cバンドレーダデータを利用して推定値を生成し、国内全地点における局地的な降雨情報の取得する手法を提案した。そして第三に、物理的な観測器によらず、人を「動く観測センサ」とみなしてTwitterの発信情報から局地的な降雨情報を抽出する手法を提案した。本提案手法による推定値はXバンドMPレーダの実績値と高精度で一致し、結果として、本提案手法の有用性のみでなく人による投稿情報自体の有用性をも示すことができたと考える。

昨今の気象現象は極めて局地的に発生し、かつ変化のスピードも速く、既存の物理的気象観測情報では把握できない場合も少なくない。迅速な状況判断にはより多くの情報が必要であり、本研究における単独の観測データ、複数の観測情報の合成データおよび人による投稿情報からの推定データは、いずれも互いの領域を補い合い利用価値が高いと考える。特に投稿情報からの推定データは、人を「動く観測センサ」として国内全域に観測地点を広げられ、貴重な情報源として今後の利用が期待できる。

本研究では局地的な降雨情報を迅速に得る手法として、さまざまなバリエーションでの提案を行った。しかし、情報共有環境の整備は実際に利用され活かされることに意味があり、特に、現状の課題である自治体や住民の水災害における避難時の行動判断には、主体的に情報を得ることが極めて重要であると考え。本研究の成果が自治体や住民の方々の気象変化に関する意識の高まりや、今後の気象予測等における情報源として寄与できれば幸いである。

## 謝辞

本論文を作成するにあたり、大変多くの方々からご指導やご支援を賜りました。指導教員のお茶の水女子大学理学部の小口正人教授には、最後まで親身で丁寧なご指導を頂きました。社会人博士として柔軟なご対応を頂き、多くの貴重な経験を通して大変お世話になりました。心より感謝とお礼を申し上げます。

副指導教員のお茶の水女子大学理学部の小林一郎教授には、本研究に関する有益かつ丁寧なご指導を頂きました。深く感謝申し上げます。

お茶の水女子大学理学部の吉田裕亮教授、浅本紀子教授、戸次大介准教授には論文審査委員として大変有益なご指導・ご助言を頂きました。深く感謝申し上げます。

国立情報学研究所所長・東京大学の喜連川優教授には、研究者として従事する場を与えて頂きました。女性活躍推進のご方針の下で、ワークライフバランスを取りやすい勤務環境、博士課程進学へのご理解とご推薦、研究上でのご指導・ご助言、および研究者としての心構えに至るまで、大変多くのご支援・ご厚情を賜りました。心より深い感謝とお礼を申し上げます。

東京大学の生駒栄司特任准教授には、日々の業務や論文執筆において多大なご指導とご支援を頂戴いたしました。厚く御礼申し上げます。

北海道情報大学の伊藤正彦准教授には、データのご提供や論文執筆において、多くのご指導とご支援を頂戴いたしました。深く感謝申し上げます。

東京大学の豊田正史教授、吉永直樹准教授には、データおよび関連プロダクトの使用を快諾下さり、御礼申し上げます。

お茶の水女子大学および東京大学喜連川研究室の先輩である絹谷弘子元特任助教には、多くのご支援・ご助言を頂きました。深く感謝申し上げます。

お茶の水女子大学小口研究室の皆様および東京大学喜連川研究室の皆様には、研究生活において多くのご支援を頂きました。ありがとうございました。

そして、生活面や精神面で支え応援してくれた主人と子供達、そして家族や友人の皆に心から感謝します。

最後に、母校で再び学ぶ喜びを与えて下さり、本論文に関わりご支援を下さいましたすべての皆様に心より御礼申し上げます。

## 参考文献

- [1] 海上智昭, 幸田重雄, 岡村信也, 堀田哲郎, 「自然災害リスク対策行動の難しさに関する態度研究に基づく論考」, 愛知工業大学研究報告, Vol.47, Mar. 2012.
  - [2] 国総研資料 第 909 号 <http://www.nilim.go.jp/lab/bcg/siryou/tnn/tnn0909.htm>
  - [3] 佐野仁美, 生駒栄司, 喜連川優, 「高リスク優先手法を用いた XRAIN 合成雨量の可視化」, 情報科学技術フォーラム講演論文集 15<sup>th</sup>, pp39-42, Aug. 2016.
  - [4] 生駒栄司, 佐野仁美, 松村浩道, 根本利弘, 小池俊雄, 喜連川優, 「XRAIN 雨量データ視覚化・ダウンロードシステム「AMeNOW!」の構築」, 情報処理学会全国大会講演論文集 78th(1), pp475-476, Mar. 2016.
  - [5] 榊剛史, 松尾豊, 「ソーシャルセンサとしての Twitter : ソーシャルセンサは物理センサを凌駕するか?」, 人工知能学会誌, 27 巻, 1 号, pp.67-74, 2012.
  - [6] A. Musaeu, C Pu, “Landslide Information Service based on Composition of Physical and Social Sensors,” 2017 IEEE 33rd International Conference on Data Engineering (ICDE), pp.1415-1416, Apr. 2017.
  - [7] T. Sakaki, M. Okazaki, and Y. Matsuo, “Earthquake shakes Twitter users: Real-time event detection by social sensors,” 19th International Conference on World Wide Web, pp.851-860, 2010.
  - [8] 長野伸一, 上野晃嗣, 長健太, 「ソーシャルセンサからの鉄道運行 情報検出システムの開発」, 電子情報通信学会論文誌, Vol.J96- D, No.10, pp.2262-2273, 2013.
  - [9] C. Maru, M. Enoki, A. Nakao, S. Yamamoto, S. Yamaguchi, and M. Oguchi, “Development of Failure Detection System for Network Control using Collective Intelligence of Social Networking Service in Large-Scale Disaster,” the 27th ACM Conference on Hypertext and Social Media (HT2016), pp.267-272, July 2016.
-

- [10] 馬賀嵩士, 三川健太, 後藤正幸, 吉開朋弘, 「気象情報と Tweet データの統合的分析による体感気温の定量化とその需要予測への応用」, 電子情報通信学会論文誌, Vol.J101- D, No.7, pp.1037-1051, 2018.
  - [11] 藤本雄紀, 今井龍一, 中村健二, 田中成典, 有馬伸広, 荒川, 貴之, 「交通現象に係わる SNS の情報収集及び活用技術に関する研究」, 日本知能情報ファジィ学会 ファジィ システム シンポジウム 講演論文集, 34, pp.403-404. 2018.
  - [12] 谷直樹, 風間一洋, 榊剛史, 吉田光男, 斉藤和巳, 「ジオタグ付きツイートを用いた交通路の抽出法」, 情報処理学会論文誌データベース (TOD) , 10(2), pp. 31-41. 2017.
  - [13] 中元雪絵, 大知正直, 榊剛史, 森純, 坂田一郎, 「食の描写表現の相違を用いた言語圏ごとの料理の感じ方の比較手法」, 人工知能学会全国大会論文集, 2018, pp. 2020S24a02. 2018.
  - [14] 国土交通省, 「土砂災害の前兆現象把握に「つぶやき情報」活用～官民共同研究 (平成 26～27 年度) の開始～」, July, 2014  
<http://www.nilim.go.jp/lab/bcg/kisya/journal/kisya20140714.pdf>.
  - [15] 森田直志, 早川誠, 高尾典佑, 「ソーシャルメディア情報を活用した防災システム」, 富士通, vol.68, no.6, pp.62-67, Nov. 2017.
  - [16] 武田邦敬, 滝口茂隆, 高橋哲朗, 山影譲, 渡部勇, 「豪雨時の Twitter データを活用した災害事象の検知」, 砂防学会研究発表会概要集, 2013, B, pp.218-219, May. 2013.
  - [17] Y. Tzeng, J. Jiang, P. Cheng, “Event Duration Detection on Microblogging,” 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Volume:1, PP.16-23, 2012.
  - [18] L. Butgereit, “Crowdsourced weather reports: An implementation of the  $\mu$  model for spotting weather information in Twitter,” 2014 IST-Africa Conference Proceedings, pp.1-9, 2014.
-

- [19] S. Kotozaki, K. Tamura, H. Kitakami, "Identifying burstiness of local topic using location-based burst detection with a classifier technique," 2014 IEEE 7th International Workshop on Computational Intelligence and Applications (IWCIA), pp.225- 230, 2014
  - [20] T. Sakai, K. Tamura, "Identifying bursty areas of emergency topics in geotagged tweets using density-based spatiotemporal clustering algorithm," 2014 IEEE 7th International Workshop on Computational Intelligence and Applications (IWCIA), pp.95-100, 2014.
  - [21] 国友優, 神山嬢子, 「Twitter を活用した土砂災害の前兆・発生状況把握の可能性」, 土木技術の総合情報誌, vol.57, no.9, pp. 18-21, Sep.2015.
  - [22] K. K. win, K. Zetsu, and K. Sugiura, "Geovisualization and correlation analysis between geotagged Twitter and JMA rainfall data: Case of heavy rain disaster in Hiroshima," 2015 2nd IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services (ICSDM), pp. 71-76, July. 2015.
  - [23] T. Funayama, Y. Yamamoto, M.Tomita, Y. Kajita, S. Tajima, K. Utsu, and O. Uchida, "Disaster mitigation support system using web services and SNS information," 2015 13th International Conference on ICT and Knowledge Engineering (ICT & Knowledge Engineering 2015), pp.42-45, 2015.
  - [24] 影澤秀明, 廣井慧, 奥矢淳, 香取啓志, 加藤朗, 砂原秀樹, 「Twitter を用いたセンシングシステムの提案と考察」, マルチメディア, 分散協調とモバイルシンポジウム 2014 論文集, 2014, pp.725-732, July. 2014.
  - [25] S.E. Middleton, L. Middleton, and S. Modafferi, "Real-time crisis mapping of natural disasters using social media," IEEE Intelligent Systems, 29, 2, Mar-Apr, 2014.
  - [26] M. Itoh, N. Yoshinaga, and M. Toyoda, "Word-Clouds in the Sky: Multi-layer Spatio-temporal Event Visualization from a Geo-parsed Microblog Stream," 20th International Conference on Information Visualization (IV2016), pp. 282-289, July. 2016.
  - [27] 伊藤正彦, 吉永直樹, 豊田正史, 「マイクロブログストリーム中の位置参照表現に着目した実世界イベントの時空間可視化」, 第8回データ工学と情報マネジメントに関するフォーラム (DEIM2016) , Feb-Mar, 2016.
-

- [28] 加藤亮平, 清水慎吾, 下瀬健一, 前坂剛, 岩波越, 中垣壽, 「局地的短時間強雨に対する気象庁高解像度降水ナウキャストの予測精度検証」, 第17回非静力学モデルに関するワークショップ, pp.55-56, Dec, 2015.
- [29] 気象庁: 数値予報解説資料(数値予報研修テキスト)「第51巻(平成30年度)第10世代数値解析予報システムと数値予報の基礎知識」第1部 最新の数値予報 第3章: 検証・事例調査  
[https://www.jma.go.jp/jma/kishou/books/nwptext/51/1\\_1\\_chapter3.pdf](https://www.jma.go.jp/jma/kishou/books/nwptext/51/1_1_chapter3.pdf)