

データの高度利用に向けた同一基盤上でのリアルタイム多種データの収集; 降雨量データと河川水位データの場合

徳島大学大学院 賛助会員 ○尾崎光洋
徳島大学大学院 正会員 中田成智

1. はじめに

近年、地球温暖化に伴う未曾有の気象災害が日本各地で発生しており、多くの人が生活基盤や命を失っている。しかし、平成 27 年の鬼怒川の堤防決壊や平成 30 年 7 月豪雨による真備町での堤防決壊のようにハード面の対策だけでは限界がある。そこで、気象災害から命を守るためには、ハード面だけでなく、ソフト面の対策がより一層重要となっている。そのためには、工学的観点からさらなる気象データの高度分析が必要である。気象データの高度分析では、異種データを扱う必要がある。そこで本研究では、ある時刻の面的なデータである降雨情報と、ある場所の点的な時系列データである河川水位情報の 2 つの異なるフォーマットのデータを収集し、これらの情報をデータベースに蓄積して互いに関連付ける基盤となるシステムを構築する。また、降雨状況、および河川水位の時系列変化を可視化することで、データに関連付けの検証を行う。

2. 気象観測と河川観測のデータモデル

本研究では、異種データを同一基盤上で管理し、今後の関連性の解析を容易にするための環境を構築していく。そのためにはまず、それらのデータを収集するとともに、データベースで管理する環境を構築する。また、本研究で収集したデータに関連付けや検索、可視化などが有効に機能していることを検証するための情報システムの開発も行う。本研究で開発するシステムは、データを収集してデータベースに保存する機能と、データベースから必要なデータを検索する機能、検索したデータに関連付ける機能の 3 つから成る。

2.1 降雨量データのデータベースによる管理

本研究では京大生存圏研究所[1]が公開している 10 分ごとの GRIB2 形式の 1km メッシュ全国合成レーダーエコー強度 GPV を降雨量データとして利用する。GRIB2 形式は世界共通の気象情報のフォーマットの一つであり、膨大な気象データを圧縮するために特殊なバイナリ形式となっている。これを Python でデータ処理するために、NetCDF 形式に変換する。収集したファイルには各地点のレーダーエコー強度の数値と観測された地点の緯度経度が入っている。さらに、図 1 のように緯度方向、経度方向に分けられたメッシュに与えた番号（以下、メッシュ番号）も加えてデータ管理する。

2.2 河川水位データのデータベースによる管理

本研究では、国土交通省水質水文データベースで公開されている 10 分ごとの各観測所の河川水位をシステムにより収集する。そして、観測所番号と各観測所の時刻、水位をデータベースに保存する。

2.3 降雨量データと河川水位データに関連付け

本研究で収集した降雨量データと河川水位データは独立している。そこで、それらのデータに関連付けるために、観測所の位置データもデータベースで管理する。観測所の位置データには各観測所の観測所番号と緯度経度が含まれている。河川水位データと観測所の位置データの間には観測所番号が共通して存在する。また、降雨量データと観測所の位置データの間には、緯度経度が共通して存在する。このことから、観測所の位置データを通して、図 2 のように独立した 2 種類のデータに関連付けることができる。

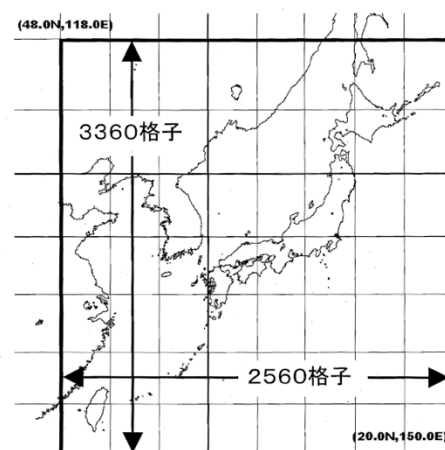


図 1. メッシュの仕組み[2]



図2. 降雨量データと河川水位データのつながり

3. データの関連付けの検証

本研究で収集した2種類のデータの関連付けができていないことを検証する。以下では、2020年9月台風10号の降雨状況と、その際の降雨量と河川水位の変化を表示したものを用いる。図3から降雨状況が表示されており、データが正常にデータベースに保管されていることがわかる。

一方図4から、降雨量データと河川水位データの2つの独立したデータが時系列で同時に表示されている。このことから、データ間の関連付けができていないことがわかる。また、他の条件の場合でも同様に機能することを確認した。

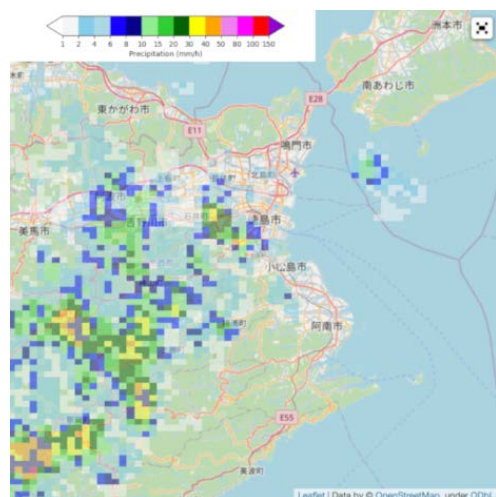


図3. 2020年9月6日21時40分の徳島市周辺の降雨状況

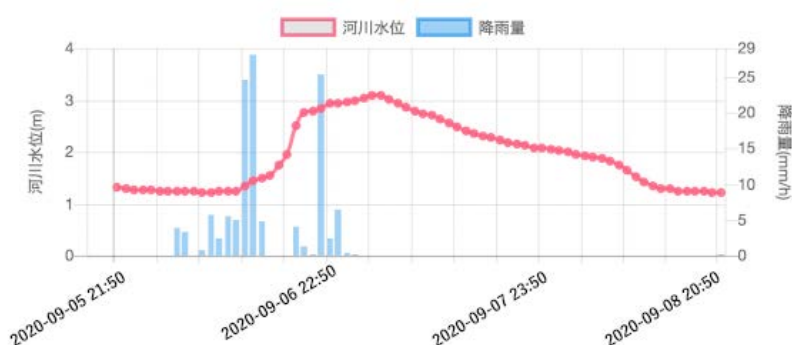


図4. 2020年9月5日21時40分~8日21時40分の古庄周辺の降雨量と河川水位の変化

4. まとめ

本研究で、多種データの同一基盤上での収集を行うことができた。また、各データを関連付けできることも確認した。今後は、本研究で構築した基盤を利用して、動的システムの理論を用いた解析モデル化を行い、河川水位の短期予測を可能にする基盤を構築していく。また、河川水位には降雨量だけでなく、地形やダムの放流などの複数の要因が考えられるため、それらのデータを用いた短期予測等も行い、データの違いによる精度の変化も確認していきたい。本研究では例として那賀川水系のデータを利用したが他の河川でも同様に行うことができるため、システムの一般化が容易である。多種のデータから関連性を見つけ、時系列分析を行い短期予測につなげる必要があるため、データサイエンスによる解析モデル化ができる環境を構築する。

謝辞

本研究では、京都大学生存圏研究所と国土交通省のデータを利用しました。ここに記して感謝いたします。

参考文献

- [1] 京都大学生存圏研究所, <http://www.rish.kyoto-u.ac.jp>
- [2] 気象庁観測部. 配信資料に関する技術情報（気象編）第162号~1kmメッシュ全国合成レーダーGPVの提供等について~, 2004年4月22日, https://www.data.jma.go.jp/add/suishin/catalogue/format/ObdObs001_format.pdf