

砂防に関する長期観測データを有効利用するためのデータベース作成・管理手法の開発

筑波大学生命環境科学研究科 ○今泉文寿 奈佐原顕郎
 京都大学防災研究所 堤 大三
 森林総合研究所水土保全研究領域 多田泰之

1. はじめに

近年気候変動が注目されているが、豪雨など極端気象の発生頻度の変化、あるいは凍結融解等の風化過程に影響する温度環境の変化により、流域内の土砂生産・流出状況が大きく変化する可能性が考えられる。これらについて実証するためには長期観測データが有用である。現在、各大学の研究実習施設（演習林等）や、官公庁の管轄流域などにおいて水文・流砂量・河床変動等の継続的な観測が行われており、中には観測期間が数十年に及ぶものもある。またこれらの観測サイトの多くは気象庁の気象観測網の展開が遅れている山地源流域に位置しているため、ここでのデータは貴重な研究資料である。しかしこれらの中には、観測後のデータ整理が充分でないためにデータの有効活用がなされていないケース、あるいは観測担当者の交代により以前のデータが散逸・紛失してしまうケースが多く見られる。今後、長期観測を維持してきた団塊の世代の退職が進み、さらには各大学・官公庁の人員・予算の削減が進展すると、さらなるデータの散逸が予想される。このような事態を防ぐためには、観測データを放置しておくのではなく、データベースとして保存・管理する必要がある。データベース化は気象・生態の分野では行われているが、砂防の分野ではほとんど行われておらず、その手法が確立されていない。そこで本研究では、過去の長期観測データのデータベース化を実践し、データベース化手法の開発を行う。

2. 長期観測データの管理状況

まず、これまでの長期観測データの管理状況について岐阜県高山市の京都大学穂高砂防観測所および静岡市の筑波大学井川演習林を例に検討し、データ管理上の問題点を議論する。

穂高砂防観測所はわが国を代表する砂防観測サイトであり、1966年の設置以来長期観測が続けられている。しかし1990年代以前のデータは主に記録紙の状態で保管されており、デジタル化のなされていないものが多い。また観測項目が多岐に渡るため、系統的なデータの保管が行われていない。このため、データをすぐに利用できるとは言い難い状態となっている。穂高砂防観測所に限らず、砂防分野では土砂移動量、流量、侵食量など観測項目が様々であるため、同様の問題は他のサイトにも共通してみられると考えられる。

南アルプスに位置する井川演習林では1968年から気象等の長期観測が行われている。起伏の激しい演習林内に7箇所の気象観測サイトを有するため、山岳気象に関する有用なデータが蓄積されている。また、土砂生産が活発なことを利用して、過去に河床変動の測量が行われた。データの多くは手書きで表にまとめられた「日表」、「月表」という形で保管されており、デジタル化されている部分はわずかである。また、過去に記録紙などのオリジナルデータの廃棄が行われた。



図-1 観測サイト位置図

3. 観測データのデジタル化

本研究では、実際に井川演習林の観測データをデータベース化することで、データベース化作業における問題点を検討した。データベース化の前段階として、記録紙など紙媒体に記録されているデータのPCへの入力（デジタル化）が必要である。記録紙をスキャナで読み込みデジタル化する試みもあるが（松野、2006）、井川演習林では記録紙が廃棄されているため、キーボードからの手入力を行った。約30年間・15項目にわたる1日間隔（一部1時間間隔）の観測値の入力および入力データのチェックに延べ約300時間を要した。データの入力にあたり以下の問題点が浮上した。

- ・信頼性が疑わしいデータがいくつかあったが、記録紙が既に廃棄されているためにデータの確認をとることができなかった。
- ・特に降水量は、観測値ゼロと機械の故障等による欠測の2つが区別されていない時期があった。

4. データベース化

次に、デジタル化されたデータのデータベース化する作業を行った。穂高砂防観測所や井川演習林でのデータの保管・利用上の問題点を踏まえ、データベースの構造には以下の点に留意した。

- ・ 様々な観測項目、観測インターバルのデータも統一的に扱えるようなデータのフォーマットとする（拡張性）。
- ・ 元データにたどり着けるよう、データソースやデータの整理者名などの情報を記録する。また、観測機器に関する情報も明記する（遡及可能性）。
- ・ Microsoft Excel 形式等の特定のアプリケーションに対応したファイルは、そのアプリケーションを使っていないユーザーにとって使い勝手が悪い。誰でもデータを利用しやすいデータ形式にする（汎用性）。

本研究では図－1 に示すようなデータ構造のファイルを各観測項目、観測場所、観測年ごとにテキスト形式で作成した。テキスト形式は自作プログラムでの処理が容易である上、PC に精通していない者でも Microsoft Excel によって表の形で閲覧が可能である。ファイル名は「年_観測サイト_観測点名_項目_観測インターバル」の規則に基づいて設定した（例：1990_IKW_SG1kadc_TM_h.txt, 1990 年に井川演習林内の総合 I という観測点で「カデック」というデータロガーによって観測された気温の 1 時間間隔のデータ）。また観測に関する情報は Read me ファイルとして別に作成した（図－2）。欠測は“n”，観測値ゼロは“0”と記入することで両者を区別した。ちなみに、データロガーから回収された過去約 15 年分、30 項目のデジタルデータおよび今回手入力したデータのデータベース化に延べ 100 時間以上要した。

#Please read 'Readme.txt'					#Please read 'Readme.txt'				
#year	month	day	hour	temperature (*C)	#year	month	day	rainfall (mm)	
2006	1	1	1	1.0	2006	1	1	1.0	
2006	1	1	2	0.9	2006	1	2	0.5	
2006	1	1	3	0.9	2006	1	3	0	
2006	1	1	4	0.8	2006	1	4	n	
:					:				

図－1 データファイル例（左：1 時間間隔で観測された気温データ，右：1 日間隔で観測された雨量データ）

・ データファイル名	・ 観測間隔	・ データソースの種類・名称	・ データファイル最終更新日
・ 整理者	・ データロガーのメーカー、型番、シリアル番号	・ 欠測期間	・ 備考

図－2 Read me ファイルへの記入事項

5. データの公開と管理

以上のようにして作成されたデータベースは以下のサイトで公開されている。

<http://www.sakura.cc.tsukuba.ac.jp/~enshurin/kishou/index.html>

このような長期観測データは様々な研究・教育に利用されるべきであるが、その一方で特に公的な機関ではデータが商業利用されないよう監視する必要がある。またデータを提供する側はデータを利用して得られた研究成果を把握する必要がある。そこで Web サイト上でユーザー名・パスワードを設け、利用希望者は一度管理者に問い合わせるユーザー名・パスワードの提供を受けた後に利用するという方法を採用した。データの管理においてはデータ構造や管理方法を徹底し、担当者が変わっても方針が変わらないようにすべきであり、またデータベースの元となる観測記録（記録紙、データファイル）の保管にも細心の注意を払う必要がある。

6. おわりに

今後の課題として、今回行わなかった測量結果や粒度分布などより複雑な構造のデータのデータベース化があげられる。本研究で行ったようなデータベース化は多くの手間と労力を要し、また過去の観測者からの観測に関する情報の提供も必要である。データベース化は今後の研究の進展や観測サイトの有効利用に大きく資するため、関係者の理解と協力のもと、各地の観測施設で実践していく必要がある。

本研究は平成 19 年度砂防学会若手研究助成を受けて行われた。また、データの整理にあたっては筑波大学関係者のご協力を得た。ここに付記して感謝いたします。

引用文献：松野潤(2006) 観測所における未公開長期観測資料の公開を目的としたデータベース化に関する研究、筑波大学修士論文