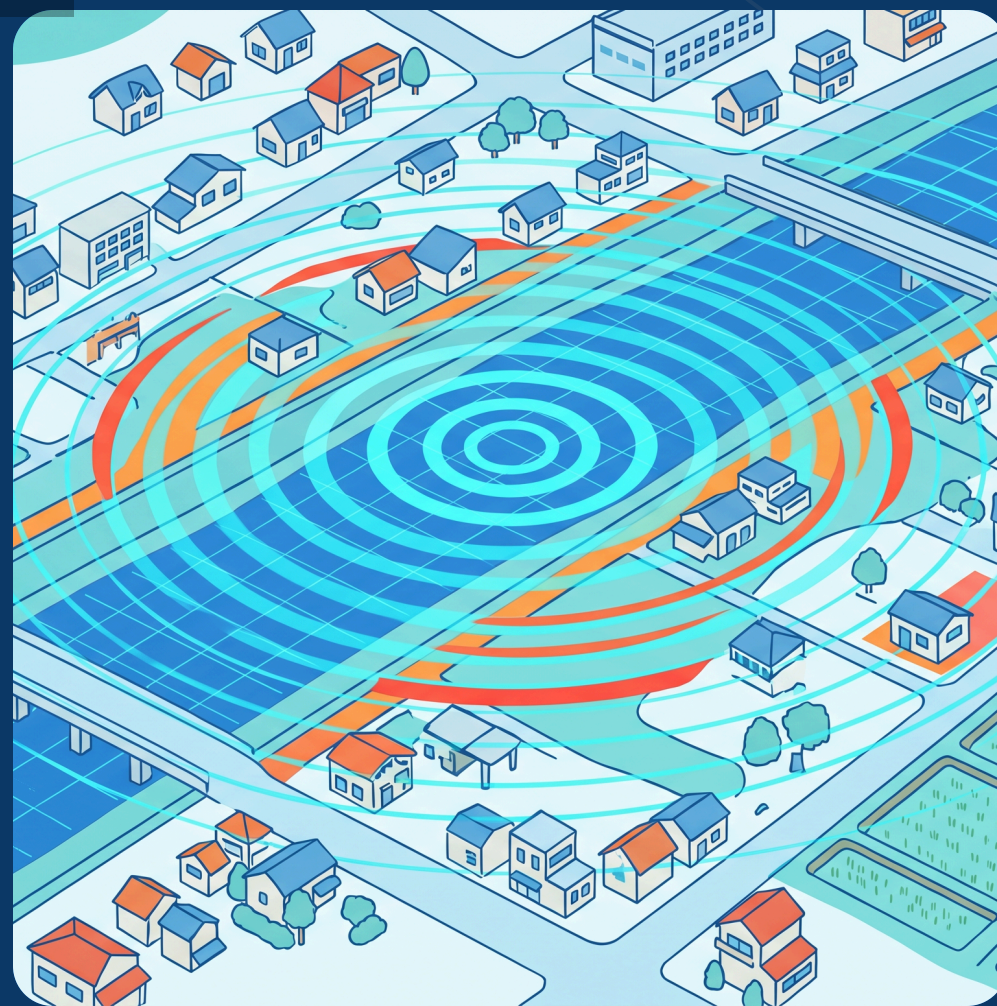


河川定期縦横断測量データの 利活用促進に向けた 保存形式標準化案

全国規模の氾濫リスク評価に向けて～

石原 道秀・星野 剛・山田 朋人

北海道大学・土木研究所寒地土木研究所



研究背景：全国氾濫リスク評価におけるデータボトルネック

データ前処理の膨大な労力と標準化による自動化の可能性

① 全国規模の河道網不定流計算の必要性

激甚化する水害への対策として、デジタル技術を活用した全国規模のリスク評価が急務となっている。

SIP第3期

流域治水DX

ケタ違いDX

② データ前処理の課題（ボトルネック）

現状、定期縦横断測量データは命名規則や書式が不統一であり、手作業でのデータ整理・統合が必須となっている。

圧倒的な手作業コスト

1流域のデータ整理・接続設定に 約2ヶ月

全国109水系換算で 約17年相当の労力

データ前処理が全国評価の最大の障壁

③ 標準化による自動化の可能性

データ保存形式（ファイル名・文字コード・CSV書式・階層）を標準化し、機械可読性を担保する。

標準化のインパクト

手動でのデータ結合処理を 機械的に完全自動化

広域の河道網不定流計算モデルの 即時構築

現実的な時間・コストで全国評価を実現

現状課題の構造化（4要件の観点から）

全国氾濫リスク評価の自動化を阻害するデータ構造のボトルネック

① 検索性の欠如

Lack of Searchability

現状の課題

ファイル名が暗号化（例: WZ001A4001.CSV）されており、河川コードやKP等のメタデータが欠如している。

技術的影響

数千件のファイルから対象データを抽出・特定するプロセスが自動化できず、前処理に膨大な人的工数が発生。

② 一意性の担保不可

Lack of Uniqueness

現状の課題

同一の距離標・測線に対して一意の識別子が定義されておらず、更新によるファイルの上書きや重複が混在。

技術的影響

版（リビジョン）の取り違いによる最新データの欠落や、不整合なデータセットを用いた不定流計算エラーを誘発。

③ 機械可読性の低さ

Low Machine Readability

現状の課題

文字コード（Shift-JIS等）やCSVの列名・列順序が事務所ごとに不統一で、独自形式のExcelフォーマット等も混在。

技術的影響

スクリプトによる自動パース処理が頻繁に失敗し、特殊文字（㎡等）の文字化けや多言語解析環境への適用を阻害。

④ GIS/DB連携の困難

Difficult Interoperability

現状の課題

空間参照系（EPSGコード等）の明示的記述がなく、全国数千のファイルが階層構造を持たないフラットな状態で散在。

技術的影響

国土数値情報河川データ等の空間ポリゴン（.shp）との自動的な空間結合が不可となり、関係データベース構築が困難。



提案の全体像（3つの標準化施策）

測量データの機械可読性・一意性・GIS連携を確保する包括的標準化

③ ディレクトリ階層化

全国規模の検索性とDB連携を担保する
一意な階層構造

定期縦横断測量データ/
[水系コード] (例: 810108) /
[河川コード] (例: 8101080024) /

▶ 施策①：構造化ファイル名

8101080024_KP24.2_20210315.csv

▶ 施策②：距離標座標データ

8101080024_LXLYRXRY.csv

▶ 推奨追加：特徴点情報

8101080024_FeaturePoints.csv

① 命名規則の構造化とUTF-8化

ファイル名の構造化

[河川コード]_[KP]_[測量日].csv

- 河川コード: 一意な識別子 (例: 8101080024)
- KP: 距離標 (例: KP24.2)
- 測量日: 年月日 (例: 20210315)

文字コードの標準化

UTF-8

特殊文字（㎡、㎥等）や多言語に対応。
プログラムからの機械可読性を担保。

② 距離標座標CSV書式の統一

空間的な相補関係を確立し、GIS連携を可能にする必須8項目の定義

空間・位置情報

- EPSGコード (例: 6668)
- 左岸X座標、左岸Y座標
- 右岸X座標、右岸Y座標

属性・時間情報

- 河川コード
- 距離標 (KP)
- 測量日

メタデータの構造化: 適用期間や測地系などの情報もヘッダー等で標準化



提案①：命名規則とUTF-8化の詳細仕様

ファイル名構造の定義と文字コード標準化の技術的根拠

1. ファイル名構造の定義

✖ 現行の課題：暗号化されたファイル名

「WZ001A4001.CSV」等では、内部データを開くまで対象河川や測量時期が特定できず、バッチ処理等による一括自動処理が不可能。

✓ 標準化仕様

[河川コード]_[KP]_[測量日].csv

例：8101080024_KP24.2_20210315.csv

≡ 各要素の技術的意義

- **河川コード**：国土数値情報河川データ（空間ポリゴン）とリレーショナルに結合するための主キーとして機能。
- **KP（距離標）**：小点对応（小数第一位を基本）を含めることで、同一河川内の測線を一意に特定。
- **測量日**：YYYYMMDD形式により、経年変化モニタリングや最新版の自動抽出など、版管理を担保。

2. 文字コードのUTF-8化

✖ 現行の課題：Shift-JISによる制限

- 測量業者の作成環境に依存し、機種依存文字の混入リスクが高い。
- 単位記号（㎡、㎥等）など水理計算に付随する特殊文字が文字化けする。
- Python等を用いた機械処理パイプラインにおいて、デコードエラー（UnicodeDecodeError）が頻発し自動化を阻害。

✓ UTF-8採用の技術的根拠

🔧 完全な機械可読性の確保

文字コードの自動判定の失敗を排除し、クラウド環境（Linuxベース）や多様なシステム間でのシームレスなデータ連携を実現。

✖ 特殊文字・多言語の安全性

水理水文データの表現に必要なすべての記号類を安全に保持可能。

🌐 国際標準（ISO/IEC 10646）への準拠

OGCが策定するWaterML 2.0等の国際水文データ標準との相互運用性を確保するための基本要件。



提案②：距離標座標CSV書式統一の詳細

空間・時間情報と座標系を明示し、GIS/DB連携の基盤を構築

⚠ 現行方式の技術的課題

列名・順序の不統一

作成事務所により列名（例: KP/キロ程/距離標）や順序が異なり、プログラムによる一括パース（読み込み）が困難。

座標系の不明瞭さ

JGD2000、JGD2011、任意座標等が混在し明記されていないため、GISマッピング時に数十メートルの位置ズレが発生するリスク。

🌐 座標系明示化の重要性

全国規模の氾濫解析において、データ自体の空間参照系（Spatial Reference System）の特定は不可欠です。

✓ 自動GIS連携の実現

EPSGコードを列として保持することで、QGIS等のプラットフォームやDB構築時にメタデータの自動認識・投影変換が可能になります。

✓ 国土数値情報との空間結合

同一の測地系・投影法への一括変換が担保され、河道網（ポリゴン・ライン）との確実な空間結合（Spatial Join）を実現します。

📄 距離標座標CSV 標準フォーマット仕様

機械可読性を担保し、空間解析に必要十分な属性を必須項目として定義します。

列名	データ型 / 推奨値・例	技術的意義・備考
河川コード	<code>String</code> (例: 8101080024)	国土数値情報河川データとのリレーションキー
KP	<code>Float</code> (例: 24.2)	距離標（キロポスト）。縦断測量の基準
左岸X, 左岸Y	<code>Float</code>	横断面図の空間配置に使用。経緯度または平面直角座標
右岸X, 右岸Y	<code>Float</code>	
EPSGコード	<code>Integer</code> (例: 6668)	最重要: 座標参照系を指定。 JGD2011 (6668) を標準として推奨
測地系	<code>String</code> (例: JGD2011)	人間可読性向上のための測地系名称
適用開始日	<code>YYYYMMDD</code>	地形変化に伴う版管理・履歴管理（時系列解析）
適用終了日	<code>YYYYMMDD</code> (最新: 99991231)	

💡 実装上のポイント

- 全ての座標値は同一のEPSGコードに基づくこと。
- 空白（Null）セルはパースエラーを防ぐため、適切な欠損値処理（例: NaN指定）を行うこと。
- BOM付きUTF-8の使用を推奨し、プログラムでの自動エンコーディング判定を確実にする。



提案仕様比較表（現行 vs 提案）

命名規則・文字コード・CSV書式・ディレクトリ構成の4観点で対比

項目	現行方式	提案方式
命名規則	✕ 暗号化されたファイル名 改善必要 WZ001A4001.CSV 距離標一覧.xlsx 標杭座標.pdf • 河川コードが不明 • 測量日・KP情報が含まれていない	✓ 構造化されたファイル名 推奨 8101080024_KP24.2_20210315.csv [河川コード]_[KP]_[YYYYMMDD].csv • 8101080024 : 河川コード • KP24.2 : 下流から24.2km • 20210315 : 測量日
文字コード	✕ Shift-JIS 非推奨 • 特殊文字（㎡、㎥）が文字化け • 多言語対応不可 • Excelで開くと文字化け	✓ UTF-8 推奨 • 特殊文字対応（㎡、㎥、℃） • 多言語対応可能 • 国際標準（ISO/IEC 10646）
CSV書式	✕ 列名・順序が不統一 改善必要 • 列名：KP, 左岸X, 左岸Y（事務所A） • 列名：番号, X座標, Y座標（事務所B） • 座標系：JGD2000, WGS84混在	✓ 標準化されたCSV 推奨 • 列名：河川コード, KP, 左岸X, 左岸Y, 右岸X, 右岸Y, 測量日, EPSGコード • 座標系：EPSG:6668（JGD2011）統一 • 必須項目：8項目（論文表1準拠）
ディレクトリ	✕ フラット構造 改善必要 • すべてのファイルが同じフォルダ • 検索困難（ファイル数：数千件） • バックアップ・管理が困難	✓ 階層化構造 推奨 定期縦横断測量データ / └─ 水系コード / └─ 河川コード / └─ 河川コード_KP24.2_20210315.csv └─ 河川コード_LXLYRXRY.csv



国土数値情報河川データとの統合手順

測量データと地図データの相補関係（表1）および接続結果（表2）

① 入力データ

国土数値情報河川

.shp 等の空間データ

- 河川コード・名称
- 水系情報
- 合流点・分岐点
- 上流・下流関係（位置）

+

定期縦横断測量

標準化された .csv

- 横断面形状（幅・深さ）
- 距離標（KP）
- 左岸・右岸座標

② 変換処理

1. 座標抽出・照合

「河川コード」をキーとして、両データセットの空間的な相補関係（論文表1）を紐付け



2. 接続判定（フィルタリング）

合流点等の複雑な接続関係を整理し、不定流計算に不要な無効ノードを除外



3. ネットワーク構成・出力

節点（ノード）と路線（リンク）による河道網ネットワークを構築し、統合CSVを生成

③ 接続結果



統合河道網データベース

（不定流計算システム用入力）

合流点接続の自動化実績（表2）

直轄一級 × 指定一級	2,487 箇所
指定一級 × 区間外	1,123 箇所
二級 × 指定区間外	859 箇所
直轄一級 × 区間外	788 箇所

全国規模の氾濫リスク評価の
自動モデリングが実現



定量的効果：接続箇所2,487の自動処理（表2データ）

データ標準化による河道網ネットワーク構築の自動化と工数削減効果

品 合流点接続の自動処理実績（論文表2）

国土数値情報河川データとの統合において、標準化された測量データを用いることで、複雑な合流点の接続関係（ノード・リンク）をプログラム処理により自動構築。

直轄一級 × 指定一級 2,487 箇所

指定一級 × 区間外 1,123 箇所

二級 × 指定区間外 859 箇所

直轄一級 × 区間外 788 箇所

✓ 技術的意義

計5,257箇所に及ぶ接続処理を、手動の空間照合からスクリプトベースの自動処理へ移行。
「河川コード」による一意な識別と、距離標データの書式統一がこれを可能にした。

🎯 標準化による自動処理の効果

1流域あたりのデータ前処理工数

従来（手作業）

提案仕様（自動化）

約 2ヶ月



約 1日

全国規模（100流域想定）でのスケール効果

従来換算

プログラム一括処理

約 17年



100日

⚙️ システム実装への展開

🌊 河道網不定流計算システム

- 本流・支川が合流する複雑なネットワークの自動構築
- 背水影響を考慮した高精度な水位予測が可能に
- データ更新に伴うモデルの再計算が容易（CI/CD化）

🗺️ GIS/DBシステムへの組み込み

- EPSGコードの明記により、異なる空間データとの即時重畳
- 流域データプラットフォームの基盤データとして活用
- デジタルツインや3D可視化のインプットとして直結



実装上の課題と対応策

技術・運用のハードルと段階的導入によるロードマップ

⚠ 現状の実装ハードル

📦 ① 技術的課題

- 河川コード欠損データの処理と補完
- 合流点の曖昧性（物理的接続とデータの接続の齟齬）
- 既存データ（Shift-JIS、非標準列）の一括変換

👥 ② 運用上の課題

- 測量業者への新仕様の周知と教育
- 既存システム（自治体固有ツール等）との互換性確保
- 移行期間中の新旧データ混在への対応

✓ 課題解決に向けた対応策（③）

📁 新規データからの段階的適用

- 国交省ガイドラインの改訂と標準化仕様の明記
- 新規発注業務の特記仕様書への反映による確実な取得

⇄ 既存データの効率的移行

- 公式フォーマット変換ツールの開発とオープンソース提供
- 文字コード変換・列名マッピング処理の自動化スクリプト配布
- 各種マニュアル整備と技術サポートデスクの設置

🏠 ④ 導入フェーズ計画

Phase 1: 新規データ標準化

- 直轄一級河川の新規測量から適用
- ガイドライン制定と運用テスト
- 変換ツールのベータ版検証



Phase 2: 既存データ一括変換

- 過去のストックデータの一括変換
- 統合河道網データベースの試験構築
- 指定一級河川への段階的拡張



Phase 3: 全国展開

- 二級河川・指定区間外への全面拡張
- 全国の氾濫リスク評価自動化運用
- 国際標準化への展開準備



国際標準化への展開可能性（WaterML2.0等）

本提案仕様に基づく水文データ国際標準（OGC/ISO）への適合と今後のロードマップ

🔑 国際標準化の必須要件（前提）

✔ 文字コードのUTF-8化

多言語対応および機械処理における国際的な互換性確保に不可欠

✔ EPSGコードの明示的付与

空間参照システム（測地系・投影法）のグローバルな特定に必須

→ 本提案はこの2要件を完全に満たしている

🔧 技術的適合性・データ相互運用性

• OGC WaterML 2.0対応への容易な変換

観測・時系列データ（Time Series / Observations）の国際標準規格であるWaterMLフォーマットへの、スクリプトベースでのシームレスな出力が可能。

• ISO 19100シリーズ（地理情報標準）への準拠性

EPSGの構造化付与を通じて、空間・地理情報に関する国際標準（ISO 19100系）の構造モデルに適合したデータセットとして利用可能。

• 国際河川データベース（GRDC等）との相互運用

フォーマットの統一化により、全球流出観測データセンター等の国際水文プラットフォームに対するデータの提供・統合が飛躍的に容易化。

🏠 水防災分野の政策・国際標準形成

ISO/TC113（水文観測）との整合

国土交通省が推進する同技術委員会での国際標準化活動に対し、技術的裏付けとなる具体的なデータ構造と日本のユースケースを提供。

インフラシステム海外展開の推進

「流域治水DX」など日本の高度な水防災技術を、標準化されたデジタルデータ基盤とともにパッケージとして海外へ輸出する基盤となる。

🗺️ 国際標準化への展開ロードマップ

Phase 1

国内データ保存形式の標準化と実装（本研究）

河川定期縦横断測量データの保存形式・運用ルールの策定と全国展開

Phase 2

国際データ規格への変換スキーム構築

構造化された統一CSVから WaterML 2.0 等への自動出力・API連携の開発

Phase 3

ISO等への提案および国際標準採択

日本発の汎用的かつ高精度な河川・水文データ管理モデルとして国際舞台へ普及



まとめと今後の展望

データ標準化による全国氾濫リスク評価の自動化に向けて

📖 研究の総括と限界

① 研究成果の要点

- 命名規則・CSV統一・ディレクトリ階層化による3つの標準化施策を提案
- 直轄・指定一級河川等、計2,487箇所の複雑な接続関係の自動統合を実証

② 技術的意義

- 全国規模での「河道網不定流計算」の実行基盤となる統合ネットワークを構築
- データ前処理工数を約96%削減（1流域2ヶ月→1日）する定量的実証

③ 本研究の限界と課題

- 現存ストックデータにおける「河川コード欠損」の効果的な補完・修復手法の確立
- 物理的な合流点とデータ上の接続点の間に生じる空間的曖昧性のシステムの解消

🏠 社会実装と今後の展開

④ 実装への道筋

- 国土交通省「河川定期縦横断測量要領」等への本仕様の反映と標準化ガイドラインの策定
- 公式フォーマット変換ツールの開発およびオープンソースとしての提供
- 新規取得データからの段階的導入と、ストックデータの一括変換パイプライン構築

⑤ 今後の展開

- OGC WaterML 2.0等へのデータ変換による、異なるプラットフォーム間の相互運用性確保
- ISO/TC113等を通じた水防災分野の国際標準化提案の推進
- 統合河道網データベースを基盤としたAI浸水予測の高度化

論文の核心

データ標準化による
全国氾濫リスク評価の自動化の実現