

ABSTRAK

DESAIN ALUR KERJA SEMI-OTOMATIS UNTUK PEMBERSIHAN DAN KLASIFIKASI ALIRAN ORDE 4 DI DAERAH ALIRAN SUNGAI OPAK

Oleh
Ameliana Nurul Aisyah
NIM: 117220020
(Program Studi Sarjana Teknik Geomatika)

Dalam pengelolaan Daerah Aliran Sungai, kualitas data jaringan sungai menjadi aspek penting sebagai dasar analisis dan perencanaan. Namun, proses pembersihan dan identifikasi aliran orde 4 di DAS Opak masih banyak dilakukan secara manual, sehingga memerlukan waktu yang lama, berpotensi menimbulkan kesalahan, dan kurang konsisten secara spasial. Oleh karena itu, penelitian ini bertujuan untuk merancang alur kerja secara semi-otomatis dalam pembersihan dan identifikasi aliran orde 4 di DAS Opak.

Data yang digunakan meliputi data aliran orde 4 dan batas DAS dari BBWS Serayu Opak, data penggunaan lahan dan jaringan sungai dari BIG, serta citra satelit dan Google *Street View* untuk tahap interpretasi manual. Tahap pembersihan dilakukan dengan mengintegrasikan *tools* QGIS ke dalam *script* PyQGIS untuk menghilangkan *noise* duplikasi sehingga satu aliran di lapangan direpresentasikan oleh satu *polyline*. Tahap klasifikasi dilakukan menggunakan metode *Random Forest* dengan parameter *geometry*, *struktur* jaringan, dan konteks penggunaan lahan serta divalidasi menggunakan *stratified k-fold cross validation*. Tahap *post-classification* dilakukan untuk penyempurnaan klasifikasi jenis aliran berdasarkan interpretasi visual secara manual.

Hasil *stratified k-fold cross validation* yaitu rata-rata akurasi sebesar 0,819, *F1-score* sebesar 0,810, dan kappa sebesar 0,62 yang menunjukkan performa model cukup stabil yang cenderung lebih baik dalam mengidentifikasi kelas sungai. Tahap evaluasi akhir menunjukkan nilai akurasi sebesar 0,873, *presisi* sebesar 0,961, *recall* sebesar 0,824, *F1-score* sebesar 0,887, dan kappa sebesar 0,7444 menunjukkan bahwa model mampu melakukan generalisasi dan bekerja baik saat diuji menggunakan keseluruhan dataset. Tahap *post-classification* berhasil memperbaiki 179 kesalahan klasifikasi (7,6%) dan menghapus 8 *segmen* aliran yang tidak relevan. Sehingga, total akhir data yaitu 2.327 aliran dengan aliran sungai sejumlah 1.271 dan saluran buatan sejumlah 1.056. Secara keseluruhan, alur kerja semi-otomatis mampu meningkatkan efisiensi, konsistensi, dan objektivitas dibandingkan metode manual, meskipun masih memerlukan koreksi manual akibat kemiripan karakteristik fitur antar kelas.

Kata kunci: Aliran Orde 4, DAS Opak, Klasifikasi, *Random Forest*

ABSTRACT

DESIGN OF A SEMI-AUTOMATIC WORKFLOW FOR CLEANING AND CLASSIFYING 4TH-ORDER STREAMS IN THE OPAK RIVER BASIN

By

Ameliana Nurul Aisyah

NIM: 117220020

(Geomatics Engineering Undergraduated Program)

In watershed management, the quality of river network data is a critical aspect serving as the basis for analysis and planning. However, the process of cleaning and identifying fourth-order streams in the Opak Watershed is still largely performed manually, resulting in time-consuming procedures, potential errors, and a lack of spatial consistency. Therefore, this study aims to design a semi-automated workflow for the cleaning and identification of fourth-order streams in the Opak Watershed.

The data used includes fourth-order stream data and watershed boundaries from BBWS Serayu Opak, land cover and river network data from BIG, as well as satellite imagery and Google Street View for the manual interpretation stage. The cleaning stage was performed by integrating QGIS tools into a PyQGIS script to remove duplicate noise so that a single stream in the field is represented by a single polyline. The classification stage was performed using the Random Forest method with parameters including morphometry, geometry, network structure, and context of land cover, and validated using stratified k-fold cross validation. A post-classification stage was conducted to refine the stream type classification based on manual visual interpretation.

The results of the stratified k-fold cross-validation show an average accuracy of 0,819, an F1-score of 0,810, and a kappa of 0,62 indicating that the model's performance is quite stable and tends to be more effective at identifying river classes. The final evaluation stage showed an accuracy of 0,873, a precision of 0,961, a recall of 0,824, an F1-score of 0,887, and a kappa 0,744 indicating that the model is capable of generalization and performs well when tested using the entire dataset. The post-classification stage successfully corrected 179 classification errors (7.6%) and removed 8 irrelevant flow segments. Thus, the final dataset totaled 2,327 flows, comprising 1,271 river flows and 1,056 artificial channels. Overall, the semi-automated workflow improved efficiency, consistency, and objectivity compared to manual methods, although it still required manual corrections due to the similarity of feature characteristics across classes.

Keywords: 4th-Order Streams, Opak Watershed, Classification, Random Forest