ProSuite QA: Enhancing Geospatial Data Quality and Efficiency with Swiss Precision

Martin BRABEC, Australia, Emanuel MAHLER, Switzerland, Urs-Jakob RUEETSCHI, Switzerland

Key words: Quality assurance (QA), Geospatial data, Error detection, Data correction, Automation

SUMMARY

Quality in geographic data is becoming an increasingly important aspect of data production and processing. To derive accurate conclusions from geospatial analysis, it is essential to work with consistent, detailed, and high-quality datasets (Devillers, 2006). Dira GeoSystems is committed to empowering geospatial professionals with the tools they need to ensure data integrity, reduce errors, and ultimately improve the quality of geospatial datasets worldwide.

ProSuite QA is trusted by a diverse range of clients, from small organisations like forest offices in Germany to large institutions such as the Swiss national mapping agency. The software supports various use cases, from conducting basic consistency checks during data imports to performing parallelised validations in server environments, where digital landscape models with millions of objects and thousands of conditions are thoroughly tested.

Engineered with Swiss precision, ProSuite QA is designed to make geospatial data quality assurance as simple, efficient, and accessible as possible. With over 130 advanced test algorithms, it detects, identifies, and localises errors with unmatched precision. From common geometric and topological errors to complex attribute inconsistencies, ProSuite covers a broad range of potential issues while minimising the time and effort needed to resolve them.

Intuitive editing tools like Cracker, Chopper, and Reshape Along Existing Features further streamline data correction, enabling users to fix dataset problems quickly and accurately. What truly sets ProSuite QA apart, however, is its commitment to automation. Its Python API allows users to address recurring issues with minimal manual intervention, enabling greater efficiency and accuracy. This combination of powerful tools and automation supports the overall goal of improving the quality of spatial data worldwide.

ProSuite QA: Enhancing Geospatial Data Quality and Efficiency with Swiss Precision

Martin BRABEC, Australia, Emanuel MAHLER, Switzerland and Urs-Jakob RUEETSCHI, Switzerland

1. INTRODUCTION

As geospatial data becomes increasingly integral to critical industries such as urban planning, environmental monitoring, and infrastructure management, ensuring the accuracy and integrity of this data is essential. The complexity and scale of geospatial datasets can present significant challenges in maintaining high-quality standards. This paper introduces ProSuite QA, a Swiss-engineered software suite designed to address these challenges by providing a set of tools for precise error detection, identification, and correction in geospatial data. With over 130 test algorithms and a range of integrated editing tools, ProSuite QA aims to improve the efficiency and effectiveness of geospatial data quality assurance (QA) processes.

2. PROSUITE QA: FEATURES AND METHODOLOGY

ProSuite QA offers a comprehensive set of tools for geospatial data quality assurance, focusing on error detection, correction, workflow integration and automation. ProSuite is designed to seamlessly integrate with existing GIS environments, supporting both manual and automated workflows to address a variety of data quality issues and maintain high quality standards when datasets are updated.

2.1 ERROR DETECTION AND LOCALISATION

One of the primary functions of ProSuite QA is its ability to detect and localise errors in geospatial datasets with high precision. The tests are designed to quickly identify errors and provide detailed information both on their type and location, enabling users to address the issues effectively as illustrated in Figure 1. The integrated issue work list which is also shown in Figure 1 at the bottom right, steps the user through the errors providing support by various tools ensuring a fast and accurate fix.

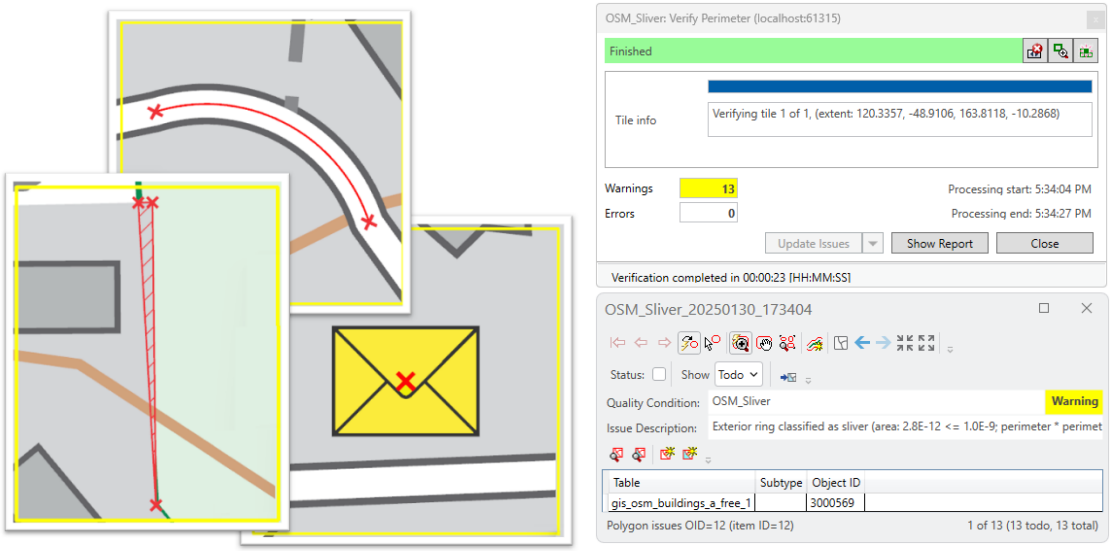

Figure 1: Precise localisation of errors and an integrated issue work list for fast and accurate corrections.

2.2 TEST CATEGORIES

ProSuite provides over 130 test algorithms that address a broad range of data issues, from common errors in simple features, as defined by the Open Geospatial Consortium Inc. (2011), to more complex issues such as geometric errors, topological inconsistencies, attribute mismatches, and 3D model inaccuracies. These tests can be grouped into the following categories:

- Attributes and relationships
- Single feature geometries
- Linear networks
- Proximity (distance between features)
- Topological conditions
- M values
- Z values
- 3D buildings (MultiPatch features)
- Polygon networks (boundary lines and centroids)
- Edge matching
- Table/feature class schemas

2.3 SELECTED EXAMPLES OF PROSUITE QA TESTS AND THEIR APPLICATIONS

In the following sections, a selection of ProSuite QA tests is illustrated, ranging from relatively simple to more complex applications. These tests ensure the accuracy and quality of geospatial data, addressing some of the various common issues in spatial data management. The names in brackets are the ones used within ProSuite QA.

2.3.1 MINIMUM SEGMENT LENGTH (QaSegmentLength)

The Minimum Segment Length test identifies line or polygon segments that are shorter than a definable minimum length. This is a very basic but important test because overly short segments can result in technical issues in some systems or negatively affect geometric processing and performance, particularly in large datasets. The test allows for evaluation in both 2D and 3D geometries, making it useful for a variety of feature types, including those with Z values.

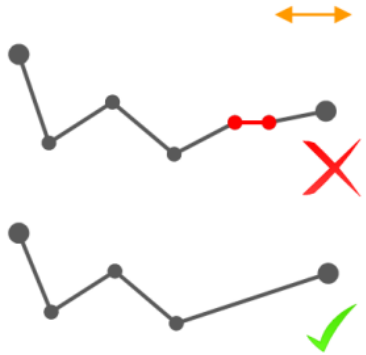

Figure 2: Minimum Segment Length test.

2.3.2 NO OVER- AND UNDERSHOOTS (QaNodeLineCoincidence)

The No Over- and Undershoots test detects point features or line endpoints that are too close to other lines or polygon boundaries but are not exactly coincident. This is critical for ensuring connectivity in networks and preventing errors in routing applications. For feature classes with Z values, the test can be applied in both 2D and 3D. It allows for flexibility in handling multi-level street crossings by using the XY tolerance of the spatial reference system to determine coincidence.

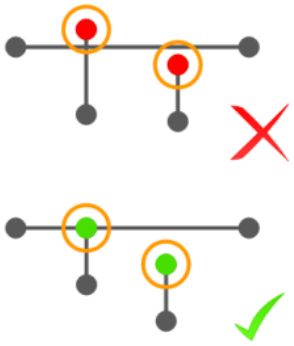

Figure 3: No Over- and Undershoots test.

2.3.3 MINIMUM DISTANCE BETWEEN LINES/POLYGONS (QaNotNear)

This test identifies sections of lines or polygon boundaries that fail to maintain a minimum distance to other features over a definable length. It is particularly useful for spotting duplicate or nearly coincident features. The minimum distance can be set to distinguish between legal line connections or crossings and errors where features are incorrectly captured too closely. This test is valuable in maintaining the integrity of large-scale datasets and preventing accidental feature overlap.

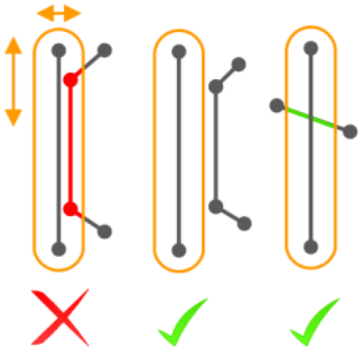

Figure 4: Minimum Distance between Lines/Polygons test.

2.3.4 CONSISTENT CONTINUIATION OF BORDER-CROSSING LINES (QaEdgeMatchCrossingLines)

The Consistent Continuation of Border-Crossing Lines test detects issues with line features that connect to border lines but do not have an expected continuation on the other side of the border. The test ensures that rules for topological and attribute consistency are followed when

lines cross borders. Attribute rules can be specified for matching fields, and the test checks for continuity in both geometry and attribute values across borders, which may be represented as lines or polygons. This is essential for ensuring consistency across administrative or technical boundaries and is typically applied when datasets from different organisations are integrated into one.

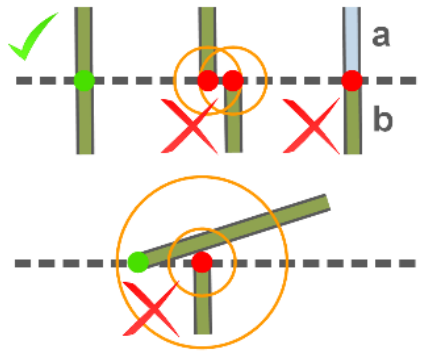

Figure 5: Consistent Continuation of Border-Crossing Lines test.

2.3.5 NO SPATIAL RELATIONSHIP ACCORDING TO THE 9 INTERSECTION MATRIX (9IM) (QaIntersectionMatrixOther, QaIntersectionMatrixSelf)

The No Spatial Relationship According to the 9IM test checks for specific spatial relationships between features, as defined by the 9IM (Egenhofer, 1993). The test can identify errors in intersection geometry, ensuring that the relationships between features are valid. It allows for exceptions based on attribute comparisons and can be applied to both multi-part intersection geometries and standard features. This test is particularly useful for complex spatial relationships, such as those found in cadastral or topographic data, where precise intersection rules are critical.

2.4 EDITING AND CORRECTION TOOLS

In addition to error detection, ProSuite QA includes a set of editing tools aimed at streamlining the data correction process. Tools such as Cracker, Chopper, and Reshape Along existing Features are designed to support the correction of errors by allowing users to edit features efficiently with plenty of visual feedback and convenient shortcuts. These tools facilitate the simultaneous editing of multiple features, which is particularly useful for correcting typical issues or complex data relationships. Moreover, these tools can be applied not only for error correction but also for general data editing tasks, thus improving the overall efficiency of geospatial data management.

2.5 WORKFLOW INTEGRATION

ProSuite is designed to integrate seamlessly into existing geospatial data workflows. It supports both interactive verification and correction within the ArcGIS Pro environment (Esri, 2010), as well as structured automated workflows where quality checks are applied at different stages of the data lifecycle as illustrated in Figure 2. Users can incorporate ProSuite into their processes from initial data collection to final delivery, ensuring errors are detected and resolved early. The tool's flexibility reduces the time spent on iterative corrections, maintaining data integrity throughout the workflow.

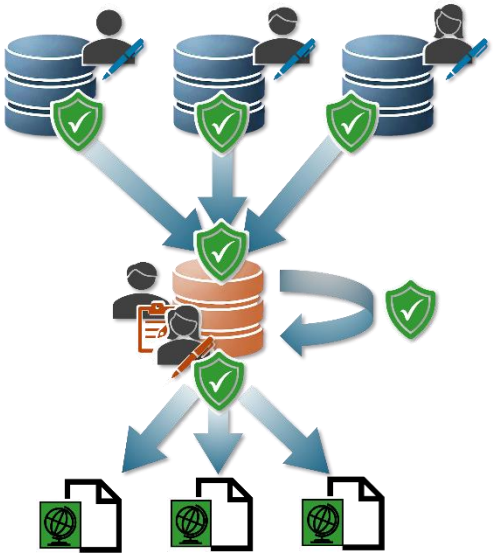

Figure 6: Quality checks in an automated workflow.

3. AUTOMATION

ProSuite QA also supports full automation through its Python API, enabling users to script and automate error detection and correction processes. This is particularly valuable for large-scale data management tasks where manual intervention would be inefficient. Automated checks can be scheduled or triggered within existing systems, ensuring continuous quality control without user input. The Python API has minimal dependencies, allowing integration with external applications and making ProSuite QA adaptable to diverse geospatial data environments.

3.1 DATA QUALITY CONDITIONS AND CUSTOMISATION

ProSuite QA allows users to define custom quality conditions tailored to specific data requirements. These conditions can be applied to both spatial and attribute data, ensuring that the data meets predefined standards before it is used for analysis or decision-making. By

allowing for custom configurations, ProSuite QA provides a flexible solution that can adapt to the unique needs of different organisations or projects. The collaboration between GIS specialists, technical staff and developers often results in a better understanding of the potential or real quality issues allowing for the implementation of even the most advanced requirements in an iterative process.

4. CONCLUSION

ProSuite QA offers a comprehensive approach to geospatial data quality assurance, addressing the growing need for consistent, detailed, and high-quality datasets in geospatial analysis. By combining error detection, correction tools, and automation, ProSuite QA directly enhances the overall quality of geospatial datasets, ensuring data integrity and reducing errors. Seamlessly integrating into existing workflows, the software supports both small organisations and large institutions in managing their datasets with greater efficiency and accuracy. This combination of manual and automated tools empowers geospatial professionals to tackle the challenges of maintaining high standards of data quality, ultimately improving the reliability of geospatial analysis across the globe.

REFERENCES

Devillers, R., & Jeansoulin, R. (Eds.). (2006). Fundamentals of spatial data quality. ISTE Ltd.

Egenhofer, M. J., Sharma, J., & Mark, D. (1993). A critical comparison of the 4-intersection and 9-intersection models for spatial relations: Formal analysis. In Auto-Carto XI.

Esri. (2010). Understanding geometric processing in ArcGIS (Esri Technical Paper, March 2010).

Open Geospatial Consortium Inc. (2011). OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture (Version 1.2.1, OGC 06-103r4).

BIOGRAPHICAL NOTES

Dr Martin Brabec is the National Manager at Dira GeoSystems in Australia. With a PhD in atmospheric physics and over 20 years of experience across various scientific disciplines, as well as a passion for GIS, he assists clients in producing high-quality geospatial data and creating visually compelling maps.

Emanuel Mahler is a software engineer and the founder of Dira GeoSystems. Currently, he is focused on developing the new topographic data production system (GoTop) for the Swiss national mapping agency. He holds an MSc in Geography and Informatics from the University of Zurich and the University of Sheffield.

Dr Urs-Jakob Rüetschi has a background and PhD in geographic information science. He provides software engineering for Esri Switzerland and Dira GeoSystems, with a special focus on cartographic processing and algorithmic design.

CONTACTS

Dr Martin Brabec
Dira GeoSystems Australia
Brisbane
AUSTRALIA
Tel. +61 494 066 467
Email: m.brabec@dirageosystems.com.au
Website: www. dirageosystems.com.au

Emanuel Mahler
Dira GeoSystems Switzerland
Geibelstrasse 35
CH-8037 Zürich
SWITZERLAND
Tel. +41 44 244 84 42
Email: e.mahler@dirageosystems.ch
Website: www. dirageosystems.ch

Dr Urs-Jakob Rüetschi
Dira GeoSystems Switzerland
Geibelstrasse 35
CH-8037 Zürich
SWITZERLAND
Tel +41 44 244 8 442
Email: u.ruetschi@dirageosystems.ch
Website: www.dirageosystems.ch