

# LACOVAL – Open-Source-Portal zur Validierung der Landbedeckung

Christoph PERGER<sup>1</sup>, Nicole VOJTECH<sup>2</sup>, Roland GRILLMAYER<sup>1</sup>, Martina DÜRAUER<sup>2</sup>,  
Jürgen WEICHSELBAUM<sup>3</sup> und Christopher DRESEL<sup>1</sup>

<sup>1</sup>Fachhochschule Wiener Neustadt · christoph.perger@fhwn.ac.at

<sup>2</sup>FOTEC Forschungs- und Technologietransfer GmbH, Wiener Neustadt

<sup>3</sup>GeoVille GmbH, Innsbruck

## Zusammenfassung

Die Evaluierung der Datenqualität anhand aussagekräftiger Qualitätsmaße stellt einen wesentlichen Teilprozess bei der Produktion von Geodaten. Obwohl eine standardkonforme Dokumentation der Datenqualität immer mehr an Bedeutung gewinnt, wird in der Praxis dieser Teilprozess oft unzureichend berücksichtigt. Diese Tatsache lässt sich auf die zeitintensive Arbeit und die damit verbundenen höheren Kosten, die durch eine wenig zufriedenstellende Softwareunterstützung noch weiter verschärft wird, zurückzuführen. Um dieser Situation entgegenzuwirken, wurde von der Europäischen Weltraumorganisation ESA das Projekt LACOVAL ins Leben gerufen welches das Ziel hat, Softwaretools zu entwickeln, um lokale, regionale und nationale Landbedeckungs- und Veränderungsdatensätze in standardisierter Form zu validieren. Im Rahmen des Beitrags wird die Systemarchitektur sowie die einfache Handhabung der entwickelten Validierungssoftware und deren Funktionsumfang dargestellt.

## 1 Ausgangslage

Bei der fernerkundungsgestützten Produktion von Geodaten stellt der Validierungsprozess einen der wesentlichen Arbeitsschritte dar. Aussagen zur Datenqualität von Fernerkundungsprodukten, wie Landbedeckungs- und Landnutzungskartierungen, sind sowohl für die Datenproduzenten als auch für die Endkunden unverzichtbare Informationen. Einerseits ist der Datenproduzent verpflichtet die geforderten Genauigkeitskriterien der Produktspezifikation im Rahmen der Produktion zu dokumentieren, andererseits ermöglicht es den Kunden eine Abschätzung der Eignung der Produkte für seine spezifischen Anforderungen. Im Allgemeinen beschreibt der Begriff „Validierung“ Prozesse, bei denen die Genauigkeit von satellitengestützten Landbedeckungsdaten evaluiert und die dazugehörigen Unsicherheiten quantifiziert werden (JUSTICE et al. 2000).

Die derzeitigen Entwicklungen von Methoden zur Validierung sind meist auf spezielle Projektanforderungen bzw. an Produkte ausgerichtet. LACOVAL setzt an diesem Punkt an und bietet standardisierte Validierungsworkflows als Onlinelösung an.

## 2 Zielsetzung

LACOVAL (Land Cover and Land Cover Change Validation) ist ein von der Europäischen Weltraumorganisation (ESA) finanziertes Projekt und hat das Ziel, Methoden und Software-Prototypen für die Validierung von satellitenbasierten Landbedeckungsdaten und Veränderungsdaten zu entwickeln. Der Fokus liegt dabei auf der Validierung von Landbedeckungs- und Veränderungsdaten, die einen lokalen, regionalen bzw. nationalen Maßstabbereich abdecken.

LACOVAL besteht aus zwei Teilsystemen:

- Das Open Source LACOVAL Online-Tool (LACOVAL Portal), das frei verfügbar und einfach verwendbar sein wird. Für die Validierung der Landbedeckungs- und Veränderungsdaten werden Standardmethoden und online verfügbarer Referenzdaten (Google Earth oder Web Map Services) zur Verfügung gestellt.
- Das LACOVAL offline Tool, das in die RegioCover Prozesskette der Fa. GeoVille (WIECHSELBAUM 2007) zur Erstellung von Landbedeckungs- und Landnutzungsprodukten integriert wird und ein erweitertes Portfolio an Validierungsmethoden zur Verfügung stellt.

Das LACOVAL Portal baut auf bestehenden Komponenten des Validierungsportals Geo-Wiki.org auf. Dieses Portal wurde zur Validierung der globalen Landbedeckung entwickelt und nutzt die Methodik des *Crowdsourcing*, um neue Erkenntnisse zur Qualität von globalen Landbedeckungsdaten zu erlangen. Insgesamt besitzt die Plattform 1800 registrierte Benutzer, die an Projekten mit unterschiedlichen thematischen Fragestellungen mitarbeiten können (FRITZ et al. 2012, PERGER et al. 2012, SCHEPASCHENKO et al. 2012).

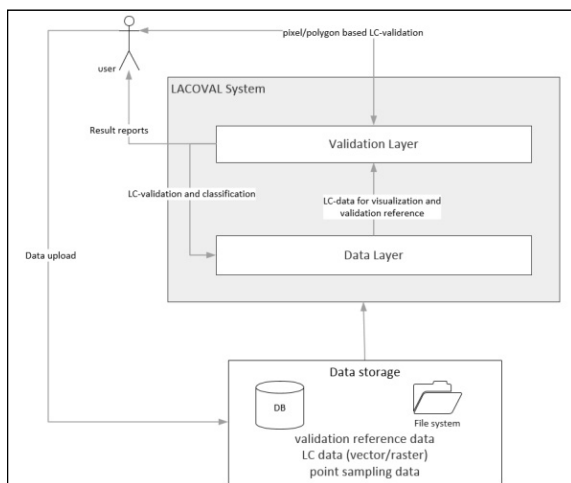
## 3 Umsetzung

Die Architektur des LACOVAL Portals setzt zum Teil auf bereits in Geo-Wiki bzw. im ASAP 7 Projekt LandSpotting entwickelten Softwarekomponenten auf. Diese Komponenten sind modular aufgebaut und können flexibel konfiguriert werden, um grundlegende GIS-Funktionalitäten im LACOVAL Portal bereitzustellen. Dazu gehören Basiswerkzeuge wie Vergrößern und Schwenken der Ansicht als auch Overlays verschiedener Datenquellen. Ebenfalls enthalten sind Werkzeuge zur Darstellung der Legende, zur Generierung von Abfragen von Informationen und Daten sowie Formulare zur Dateneingabe.

Das LACOVAL Portal weist eine Dreischichtenarchitektur auf: Datenspeicherung, Daten- und Validierungsschicht (Abb. 1).

Die Validierungsschicht fordert Landbedeckungsdaten, die zuvor auf das Portal geladen werden können, von der Datenschicht an. Anschließend wird der Untersuchungsbereich entweder direkt vom hochgeladenen Landbedeckungsdatensatz abgeleitet oder optional bei der Erstellung des Validierungsprozesses mittels Bounding Box definiert. Der Benutzer interagiert mit dieser Schicht und validiert innerhalb des Untersuchungsbereichs. Zentrale Komponenten in der Validierungsschicht steuern den Validierungsworkflow und ermöglichen die Erstellung von Samples für den Validierungsprozess. Beim Sampling kann entweder ein systematisches Punktraster mit definiertem Punktabstand (Meter oder Dezimalgrad) oder ein Zufallspunktsample gewählt werden. Des Weiteren können auch bereits generierte

Punkte ins Portal hochgeladen werden. Als Besonderheit ist auch die Validierung von kontinuierlichen Variablen auf Pixelebene möglich, wie sie derzeit für die hochauflösenden Landbedeckungsdaten im europäischen GMES/Copernicus Land Monitoring erzeugt werden.



**Abb. 1:**  
Schematische Darstellung der  
Dreischichtenarchitektur des  
LACOVAL Portals

Die Validierung kann entweder ohne Kenntnis der bereits durchgeführten Klassifizierung (*blind*), oder per Überprüfung der Plausibilität dieser Klassifizierung erfolgen. Für Karten, die die Veränderung in der Landbedeckung darstellen, gibt es zusätzlich die Möglichkeit das Sampling ausschließlich innerhalb der Veränderungsflächen durchzuführen.

Abbildung 2 zeigt das LACOVAL Portal bei der Validierung eines LISA (Landinformationssystem Austria) Landbedeckungsdatensatzes im Gebiet Traun/Oberösterreich. Die linke Spalte führt den Benutzer durch den Validierungsprozess (Punkte 1-4). Im Folgenden wird näher auf diese vier Hauptfunktionalitäten des LACOVAL Portals eingegangen.

### 3.1 Hochladen des Datensatzes für die Validierung

Der Upload der Landbedeckungsdatensätze und Referenzdaten wird über das Open-Source-Tool GeoNode durchgeführt. Vektor- und Rasterdaten können als Shapefile bzw. GeoTiff in der Projektion WGS84 (EPSG:4326) hochgeladen werden. Shapefiles müssen darüber hinaus zwingend die Spalte „lccode“ (land cover code) als *short integer* enthalten. Die Anzeige des Datensatzes kann bei Shapefiles mittels SLDs definiert werden, Rasterdaten müssen über eine Colormap verfügen.

### 3.2 Erstellung des Validierungssamples

Mittels Assistent wird in fünf Schritten das Validierungssample generiert. Dazu wird zuerst der Datensatz ausgewählt und parametrisiert um zwischen Landbedeckungs-, Veränderungsdatensatz oder kontinuierlicher Landbedeckung (z. B. Versiegelungsgrad) zu unterscheiden. Im nächsten Schritt erfolgt die Definition der Nomenklatur durch Angabe der Klassencodes und Klassennamen. Anschließend kann der Untersuchungsbereich der Vali-

dierungen definiert werden. Die Definition kann entweder direkt aus dem zuvor definierten Landbedeckungsdatensatz abgeleitet und übernommen oder alternativ mittels Polygon auf einer Karte definiert werden.



**Abb. 2:** LACOVAL Portal – LISA Landbedeckungsdaten im Gebiet Traun/OÖ

Den wichtigsten Schritt stellt die Definition des Validierungssamples dar. Es wird zwischen Punkt- und Pixelvalidierung unterschieden, wobei Punkte bei klassifizierten Raster- und Vektordatensätzen sowie Pixel bei kontinuierlichen Datensätzen verwendet werden. Danach werden die Samples generiert. Für die räumliche Verteilung der Samples stehen drei Möglichkeiten zur Verfügung:

- Zufallsverteilung:  
Eine zu definierende Anzahl an Punkten wird zufällig innerhalb des Untersuchungsgebietes generiert.
- Systematisches Punktsample:  
Es werden Punkte in einem zu definierenden Abstand in Meter voneinander entfernt generiert.
- Zufallsselektion:  
Eine zu definierende Anzahl von Pixeln wird zufällig aus dem kontinuierlichen Datensatz ausgewählt.

Bei den generierten Samples besteht außerdem die Möglichkeit existierende Stichproben hinzuzufügen (*Augmentation*). Dazu muss ein Shapefile mit den Punkten zuvor im GeoNode Portal hochgeladen worden sein. Weiters kann die Anzahl an zu generierenden Punkten auf 0 gesetzt werden, um nur die existierenden Samplepunkte zu verwenden.

Abschließend wird die Methode, die bei der Validierung verwendet werden soll, definiert. Bei der *blind-Validierung* hat der Benutzer keine Information über die Klassifizierung des Samples und interpretiert somit das vorliegende Bildmaterial hinsichtlich der vorliegenden Landbedeckung (Google Earth und Referenzdaten). Bei der *Plausibilitätsprüfung* wird die

Landbedeckungsklasse des Samples angezeigt und der Benutzer entscheidet über die Richtigkeit dieser Klassifizierung.

### 3.3 Validierung des Samples

Beim Start des Validierungsprozesses wird automatisch auf das ausgewählte Sample gezoomt. Je nach gewählter Validierungsmethode besteht das Validierungsformular entweder aus einem Schalter zur Bestätigung oder Ablehnung der thematischen Klasse oder aus einem Dropdown-Feld zur Auswahl der Klasse. Durch Abschicken des interpretierten Samples wird automatisch der nächste Stichprobenpunkt aktiviert und auf den passenden Kartenausschnitt gezoomt. Bei der Validierung stehen dem Benutzer alle verfügbaren Referenzdaten zur Verfügung, die bei frei wählbarer Transparenz aktiviert und mit den Google-Earth-Daten überlagert werden können.

### 3.4 Validierungsbericht erstellen

Nach Abschluss der Validierung des gesamten generierten Samples (oder zur Zwischenkontrolle) kann der Validierungsbericht im Microsoft Excel (.xlsx) Format heruntergeladen werden. Dieser Bericht beinhaltet standardkonforme Messungen zur Genauigkeit (Konfusionsmatrix, Gesamtgenauigkeit mit Konfidenzintervallen, Benutzer- und Produzenten-Genauigkeit, Kappa Statistik, sowie Cross-Plots/RMSE/MAE). **Tabelle 1** zeigt einen einfachen Validierungsbericht nach Plausibilitätsprüfung eines LISA Landbedeckungsdatensatzes zur Veranschaulichung der User Accuracy-Berechnung.

**Tabelle 1:** Beispielhafter Auszug eines Validierungsberichtes

Klassifikation	Korrekt	Inkorrekt	Zeilensumme	User's Accuracy
Gebäude	9	0	9	100 %
Offener Boden	8	2	10	80 %
Bestockte Fläche	19	5	24	79,17 %
Gebüsch	9	1	10	90 %
Grünbewuchs	18	1	19	94,74 %
Spaltensumme	63	9	72	
<b>OVERALL ACCURACY</b>				<b>87,5 %</b>

Die Rohdaten der Validierung können als Google-Earth-KML-Datei heruntergeladen werden. Zur Auswahl stehen der Export aller Samples, ungeachtet der Fertigstellung der Validierung oder nur die validierten Samples. Der erstellte Download beinhaltet neben der räumlichen Information zwei Attributfelder. Ein Feld mit der Klassifizierung des originalen Datensatzes und ein Feld mit der Validierungsinformation, das im Falle einer noch nicht erfolgten Validierung (bei Export aller Samples) leer bleibt.

## 4 Resultate und Ausblick

Die vorliegende Arbeit beschreibt die Erstellung des Prototyps eines Webportals zur Validierung von lokalen, regionalen und nationalen Landbedeckungs- und Veränderungsdaten-

sätzen. Besonderes Augenmerk bei dieser Lösung wurde auf die Definition des Validierungssamples und auf die Validierungsberichte gelegt. Jeder User kann eigene Vektor- und Rasterdatensätze sowohl zur Validierung als auch als Referenz für den Validierungsprozess über das Webportal hochladen. Nach dem Hochladen der zu validierenden Datensätze und Referenzdaten wird das Sample generiert, welches zufällig oder systematisch verteilt sein kann. Diese Samples können entweder aus Punkten oder aus Pixel bestehen. Für kontinuierliche Datensätze steht die Zufallsauswahl von Pixel zur Verfügung. Zusätzlich zu generierten Samples können Punkte aus einem hochgeladenen Shapefile zur Validierung herangezogen werden. Mittels der zwei Validierungsmethoden Neuinterpretation (*blind validation*) oder Plausibilitätsprüfung wird das erstellte Sample validiert. Zum Abschluss können sowohl der Validierungsbericht, der standardkonforme Messungen zur Genauigkeit enthält, wie auch die Rohdaten der Validierungen heruntergeladen werden.

Das Ergebnis von LACOVAL bildet ein Alpha-Prototyp, welcher einer breiten Community zu Testzwecken online kostenlos zur Verfügung gestellt wird. Das im Projekt ebenfalls entwickelte ArcGIS-Expertentool wird zukünftig für operationelle Kartierungsprojekte eingesetzt, z. B. im Bereich des österreichischen (Land Information System Austria) und europäischen Landmonitorings (GMES/Copernicus) sowie in durch die ESA finanzierten Erdbeobachtungsprogrammen (z. B. TIGER-NET in Afrika).

Um der wenig zufriedenstellenden Situation der fehlenden Softwareunterstützung im Rahmen des Validierungsprozesses entgegenzuwirken und die standardkonforme Dokumentation der Datenqualität voranzutreiben, wäre eine Weiterentwicklung des Prototypen hin zu einem Produktivsystem ein wünschenswerter und sowohl für die Datenproduzenten als auch Konsumenten wertvoller Beitrag zur Erhöhung der Wiederverwendbarkeit, des richtigen Einsatzbereiches und der Wertsicherung und von Landbedeckungsdaten.

## Literatur

- FRITZ, S., MCCALLUM, I., SCHILL, C., PERGER, C., SEE, L., SCHEPASCHENKO, D., VAN DER VELDE, M., KRAXNER, F. & OBERSTEINER, M. (2012), Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software*, 31, 110-123.
- JUSTICE, C. O., BELWARD, A., MORISSETTE, J., LEWIS, P., PRIVETTE, J. & BARET, F. (2000), Developments in the ‘validation’ of satellite sensor products for the study of land surface. *International Journal of Remote Sensing*, 21, 3383-3390.
- PERGER, C., FRITZ, S., SEE, L., SCHILL, C., VAN DER VELDE, M., MCCALLUM, I. & OBERSTEINER, M. (2012), A campaign to collect volunteered geographic information and land cover and human impact. In: JEKEL, T. et al. (Eds.), *GI Forum 2012: Geovisualisation, Society and Learning*. Wichmann, Berlin/Offenbach, 83-91.
- SCHEPASCHENKO, D., FRITZ, S., MCCALLUM, I., SHVIDENKO, A., PERGER, C., SCHILL, C., SEE, L., KRAXNER, F. & OBERSTEINER, M. (2012), A platform to visualize, analyze and improve biomass datasets, *Folia Forestalia Polonica, Series A – Forestry*, 54 (2), 137-139.
- WEICHSELBAUM, J. (2007), Intelligent Geoinformation to Go! – GeoVille Group Uses GIS to Automatically Transform Images into Affordable and Precise Geoinformation Products. *ESRI ArcNews Online*: <http://www.esri.com/news/arcnews/summer07/articles/intelligent-geoinformation.html>.