Стручни критички осврт

Јован Д. Ковачевић, *Картирање шумске вегетације на основу података сателитског осматрања Земље коришћењем техника машинског учења*. Докторска дисертација. Београд: Грађевински факултет Универзитета у Београду, студијски програм Геодезија и геоинформатика, 2021, 22 предстране без броја + 166 стр. (основни текст на штампаним странама 1–123; литература од стр. 124, прилози од стр. 138, биографија на стр. 159). Ментор в. проф. др Жељко Цвијетиновић; комисија за преглед, оцену и одбрану: Д. Михајловић, Љ. Брајовић, М. Самарџић-Петровић, Д. Лакушић, Б. Радић. Отворени приступ: НаРДуС.

I Предмет, грађа и мерило

Дисертација узима за предмет методологију картирања типова шумске вегетације на целој територији Србије из вишевременских снимака мисије Sentinel-2, надзираним непараметарским техникама машинског учења. Задатак је постављен онако како га уже области фотограметрије и даљинске детекције и геоинформатике данас препознају као тежак: не сегментисати пиксел на шуму и нешуму, него разликовати типове шумске вегетације на националној размери, из бесплатних оптичких података, аутоматизованим поступком који се може поновити. Рад тај задатак и уоквирује без прецењивања, изричито признајући да „универзално решење за картирање типова шумске вегетације на глобалном нивоу и за било који ниво детаљности и даље не постоји“ (стр. 6) и да „изналажење оптималног решења подразумева дефинисање локално прилагођене методологије, којом се обухватају све специфичности и карактеристике подручја истраживања“ (стр. 6). Тежиште рада је методолошко и инжењерско — развој, тестирање и оцена поступка — што ће бити важно при мерењу свих замерки које следе.

Пре свега осталог ваља изрећи мерило, изричито и без изузетка. Рукопис је предат Служби за студентска питања 15. децембра 2021. године, а три прописане изјаве у прилогу датиране су 13. децембра 2021; реферат Комисије носи датум 24. децембра 2021. Рад се стога овде мери искључиво према литератури, стандардима, прописима, продуктима и знању који су постојали и били доступни пре средине децембра 2021. Ништа настало касније није мерило: ни доцнија издања стандарда, ни продукти објављени после тог тренутка, ни научни радови који су уследили. То правило важи и за списак литературе овог осврта — свака његова јединица објављена је пре рока, изузев саме дисертације, реферата Комисије и записа репозиторијума, који служе идентификацији рада, а не његовом мерењу.

Наводи из дисертације преносе се дословно, у наводницима, ћирилицом или латиницом онако како стоје у самом раду, с бројем штампане стране у загради. Тврдње о продуктима, стандардима и прописима сравњиване су с изворним текстовима у верзији важећој средином децембра 2021. Целовит текст рада коришћен је из репозиторијумског примерка, а оцена Комисије и подаци о поступку из њеног реферата.

II Шта рад доноси

Осврт је дужан да почне оним што дисертација јесте, јер је то и највећи део онога што она вреди.

Прво и најважније, рад изводи целовиту, аутоматизовану методологију и примењује је на целој територији Србије (88 361 km²), у јединственом 20-метарском растерском гриду и државној пројекцији, над 89 гранула Sentinel-2 нивоа обраде Level-2A из вегетационе сезоне 2019. То није студија случаја на неколико стотина квадратних километара него продукт националне размере, што сам рад с правом истиче као кључну разлику у односу на упоредиве студије сличне тачности (стр. 103).

Друго, рад доноси промишљену операционализацију временске компоненте. Две репрезентације — имплицитна, у којој три сезонска интервала улазе као засебни атрибути (196 454 опажања, 35 атрибута), и експлицитна, у којој датум опажања улази као атрибут (1 231 783 опажања) — постављене су тако да се могу директно поредити: „Резултујући скуп података са имплицитном временском дефиницијом има 196 454 опажања, што је шест пута мање опажања у односу на експлицитно дефинисану временску компоненту“ (стр. 71). Реч је о самосталном и добро операционализованом методолошком доприносу, чија је сврха баш да покаже да се фенолошка информација може искористити и када недостају опажања услед облачности.

Треће, рад пажљиво гради референтну грађу и балансирање класа. Тренинг скуп од 743 полигона („Резултујући скуп података се састоји од 743 полигона који представљају седам шумских класа и једну класу жбунасте вегетације“, стр. 62) векторизован је ручно око локација из базе о распрострањености станишта Србије, а балансирање је изведено прилагођеним алгоритмом SMOTE с прагом од 30 %: „Као праг балансираности је одабрана вредност од 30%, а додатно је уведен и услов како би се очувао редослед заступљености класа из почетног скупа података“ (стр. 77). Очување стварне неравнотеже класа, уместо њиховог потпуног изједначавања, промишљено је и зрело решење.

Четврто, валидациона поставка је методолошки исправна на два места осетљива на оптимистичку пристрасност. Балансирање се спроводи само над скуповима за тренирање, независно по итерацији — „балансирање класа вршено само над

скуповима за тренирање класификационог модела, независно за сваку итерацију из пет пута поновљене петоструке крос-валидације са изостављањем“ (стр. 95) — чиме се синтетички примерци не преливају у валидацију; а крос-валидација изоставља комплетне локације, а не насумичне примерке, што уважава просторну зависност узорка. Обе одлуке одговарају ономе што методолошка литература препоручује (Chawla и др., 2002; Roberts и др., 2017).

Пето, рад је интелектуално поштен према сопственим бројевима. Тамо где разлике нема, одбија да прогласи победника — „Због тога на основу тако малих разлика није могуће тврдити који приступ даје тачније резултате“ (стр. 83); свестан је да крос-валидација прецењује квалитет карте и због тога уводи независну валидацију — „Коришћена крос-валидација са изостављањем комплетних локација може дати информације о тачности модела, али су срачунате метрике тачности генерално пристрасне када је реч о оцени квалитета карте на комплетном подручју истраживања (Olofsson и др., 2014)“ (стр. 109); и не проглашава карте готовим продуктом — „Све ово сугерише да коначне карте типова шумске вегетације захтевају додатну валидацију пре него што могу бити пуштене у ширу употребу“ (стр. 116). Таква самокритичност најјача је одбрана рада од највећег дела примедаба које следе.

Најзад, ваља рећи и оно што надилази појединачне доприносе: где год су се тврдње рада могле сравнити с изворним подацима и с унутрашњом рачуницом, у највећем делу су се показале тачним и доследним. Полазни бројеви (743 полигона по класама, 196 454 и 1 231 783 опажања) слажу се кроз цео рад; матрице конфузије независне валидације интерно су доследне, а из њих се тачно реконструишу произвођачеве и корисникове тачности, укупна тачност и капа; спецификације продуката (Sentinel-2, EU-DEM v1.1) верно су пренете. Та поузданост чињеничне основе није споредна околност — она је оно што уопште омогућава да се о раду расправља овако детаљно. Формално, рад испуњава и услове за одбрану: садржи три прописане изјаве (Упутство за формирање репозиторијума докторских дисертација, Сенат УБ, 2011), а услов претходно објављеног, садржински повезаног рада првог аутора испуњен је радом Kovačević и др. (2020) у часопису *Remote Sensing* (Правилник о докторским студијама на Универзитету у Београду, 2021, чл. 34), који је у дисертацији и цитиран (стр. 21).

III Циљеви, претпоставке и метод

Аналитички лук рада јасан је: од избора и припреме података, преко двеју временских репрезентација, избора атрибута, балансирања и просторно-временске агрегације оцењених вероватноћа, до четири узастопна експеримента оцењена петоструком крос-валидацијом с изостављањем комплетних локација, поновљеном

пет пута. Метод одговара најави — три технике машинског учења (стабло одлучивања, случајна шума, градијентно појачавање), пет метода избора атрибута и прилагођени SMOTE стварно су примењени и вредновани.

Једно место, међутим, тиче се начина на који су постављени циљеви. Они су у уводу дати квалитативно, као шест парцијалних задатака, без изричито исказаног нумеричког прага успеха и без формално оборива хипотеза; критеријуми оцене (укупна тачност, капа, крос-валидација) уведени су тек у поглављима 3 и 5, а успех се на крају мери поређењем с упоредивим студијама, не унапред задатим прагом. Реч је о формалном питању структуре увода, које не доводи у питање мерљивост резултата, тим пре што рад сам поставку ограничава — оптималност изричито назива локалном: „проглашени оптимални параметри методологије се могу сматрати локално прилагођеним, применљивим само за дефинисано подручје истраживања и за обезбеђен скуп података“ (стр. 122). Ту, међутим, треба приметити и напетост с рефератом, о којој в. одељак XI: реферат исте циљеве накнадно представља као истраживачке хипотезе, док у самом тексту рада хипотеза под тим именом нема.

Из истог круга је и статус двеју носећих претпоставки. Претпоставка о довољности три сезонска интервала — „Претпоставка је да ова три временска интервала довољно наглашавају разлике у фенолошким обрасцима сваког типа шумске вегетације“ (стр. 66) — посредно је подржана резултатима, али није директно проверена (нпр. поређењем с гушћом временском серијом). Претпоставка о неосетљивости резултата на хиперпараметре — „претпоставља се да промена параметара неће значајно утицати на одређивање оптималних параметара методологије“ (стр. 80) — усвојена је, а не тестирана, јер варирање хиперпараметара техника машинског учења није вршено. Ниједна од тих претпоставки не руши резултат, али обе остају ограде које би будући рад требало да испита, што и сам закључак признаје.

IV Преглед сателитских система и литературе о шумама

Прегледни део (поглавље 2) структуриран је око јасног, признатог оквира — три нивоа информација о шумама по Boyd-у и Danson-у (2005) — што уводу даје аналитичку окосницу уместо пуког набрајања, а преглед сензора коректно упућује читаоца на ауторитативне изворе уместо да претендује на исцрпност. Преглед литературе о картирању типова шума из Sentinel-2 тачно преноси кључне налазе упоредивих студија, укључујући бројке из Grabska и др. (2019) — „Grabska и др. (2019) су постигли укупну тачност већу од 90% на истраживаној површини од 305 km2 за девет типова шумске вегетације користећи 18 Sentinel-2 снимака без облака“ (стр. 103).

Прегледне табеле сателитских сензора (Табеле 1 и 2), међутим, носе низ ситних непрецизности које не додирују методологију ни резултате, али их у монографском издању вреди отклонити. Група веома високе резолуције дефинисана је на стр. 14 прагом мањим од једног метра, а међу њене представнике увршteни су и RapidEye (6,5 m) и ZiYuan III (2,1–5,8 m), који по сопственој Табели 1 нису субметарски; исти праг разликује се од дефиниције „веома високе резолуције“ као мање од 5 m на стр. 11. Почетна година мисије SAOCOM у Табели 2 наведена је као 2015, иако је та мисија лансирана тек 2018; за Sentinel-1 и ALOS-2 реч је о непрецизности од једне године. Вредност најфиније панхроматске резолуције Landsat-а дата је као до 10 m, док је најфинији панхроматски Landsat канал 15 m. Тврдња да су снимци Landsat-9 доступни била је у децембру 2021. преурањена (јавно преузимање података почело је тек 10. фебруара 2022). Комерцијални називи мисија местимично су писани нестандардно („Worldview“ уместо WorldView, „COSMO-Skymed“ уместо COSMO-SkyMed, „Pleiades“ уместо Pléiades). Табеле не наводе извор компилованих спецификација по реду, него се ослањају на опште референце у пратећем тексту, па порекло појединачних вредности није директно проследиво — што иначе не би било вредно помена да табеле не садрже и понеку нетачну вредност. Најзад, у уводу је формулација „више од 2000 активних сателита који врше осматрање Земље“ (стр. 10) непрецизна, јер бројка одговара укупном броју активних сателита, а не само оних намењених осматрању Земље. Све су то грешке прегледа, а не поступка; наводе се ради потпуности, а не зато што дирају тезу.

V Технике машинског учења, избор атрибута, балансирање и мере тачности

Теоријски приказ методолошког апарата (поглавље 3) прегледан је, систематичан и добро поткрепљен изворима: ентропија и Гинијев индекс, критеријумска функција CFS и симетрична неодређеност, историјски развој генетских алгоритама, SMOTE с изричито наведеним ограничењем методе, капа коефицијент и мере тачности из матрице конфузије верно су пренети из меродавних извора. Ослонац на Olofsson-а и др. (2014) за оцену тачности — „капа коефицијент слагања (енгл. kappa coefficient of agreement) (Olofsson и др., 2014)“ (стр. 52) — управо је препоручена литература уже области. Први закон географије тачно је цитиран, преведен и приписан Tobler-у (1970): „Све је повезано са свиме, али су блиске појаве међусобно више повезане него оне удаљене“ (стр. 41).

У уџбеничком приказу постоји неколико коректносних омашки које, треба одмах рећи, не задиру у резултате, будући да је реч о поступцима који у самом истраживању или нису коришћени (нпр. степен информативности атрибута — gain ratio) или су изведени готовим пакетом. У једначини 3.4 (стр. 35) степен информативности нормализован је ентропијом расподеле класа, а не интринзичном

информацијом поделе по атрибуту, па тако дефинисана мера не отклања пристрасност према вишевредносним атрибутима, што је наведени мотив; како се избор атрибута у раду спроводи методама CFS и омотач-методама, а не gain ratio-ом, последице по резултате нема. У једначини 3.8 (стр. 37) постоји недоследност предзнака градијентног члана (типографска омашка у слогу), а у једначини 3.6 (стр. 37) обичан резидуал изједначен је с негативним градијентом, што важи тачно само за квадратну функцију губитка; будући да је коришћен пакет xgboost, приказ не мења резултате. Скраћеница SU употребљена је двоструко — за симетричну неодређеност и за удео случајног слагања у капи — што је формална недоследност нотације, недвосмислена из контекста. Најзад, изворни рад Friedman (2001) о градијентном појачавању није наведен као засебна библиографска јединица, него се приказ ослања на меродаван преглед Hastie и др. (2009), уз изричито именовање родоначелника у тексту. Ниједна од ових ставки не носи резултат; све су уредничке или приказне природе.

Три методолошке примедбе теже су и вреди их изрећи прецизно, не строже него што материја допушта.

Прва је да рад ни уз једну меру перформанси не наводи меру расипања између пет понављања крос-валидације (стандардну девијацију, интервал поверења) ни тест значајности разлика, па се разлике реда неколико десетих делова процента укупне тачности и неколико стотих делова капе не могу поуздано разликовати од случајне варијације. За ту врсту исказивања варијабилности не постоји један именовани стандард који би је изричито налагао, па се примедба води као методолошки аргумент, поткрепљен општом праксом ресемплинга (Kuhn и Johnson, 2013; Hastie и др., 2009). Њен домет треба тачно оивичити: она погађа само изборе засноване на тако малим разликама — пре свега правило агрегације — док носећи избори (искључење стабла одлучивања као слабијег, увођење суседства, балансирање класа) почивају на разликама знатно изнад тог прага, а рад сам, на месту где су разлике занемарљиве, одбија да пресуди (стр. 83). Отуда је оправдано ублажити и једну реченицу самог рада — да проста аритметичка средина незнатно надмашује преостала два правила агрегације — будући да су сва три правила практично изједначена и да избор међу њима не мења оптимални модел.

Друга је да је избор репрезентативних атрибута изведен засебном (двоструком) крос-валидацијом над целим скупом података, а не угнежђен у спољну евалуациону петљу, што начелно може оптимистички проценити изабране подскупове атрибута. То је стварно ограничење: класичан доказ да избор атрибута мора бити унутар сваке итерације крос-валидације дали су Ambroise и McLachlan (2002), а канонско место о томе носи и уџбеник који рад сам користи (Hastie, Tibshirani, Friedman, 2009, одељак 7.10.2); за просторно-временске податке исти захтев износе Meyer и др.

(2018). Последица по носећи закључак је, међутим, безначајна, јер рад као најбоље ионако усваја комплетан скуп атрибута, који том цурењу није подложан — те би угнежђена процена била методолошки чистија, а да главни исход не би променила.

Трећа се тиче просторне независности. Изостављање комплетних локација у крос-валидацији мера је коју методолошка литература баш и препоручује против просторне зависности узорка (Roberts и др., 2017), и рад ју је применио — али оно не уклања просторну аутокорелацију међу кластерисаним полигонима исте класе, па је крос-валидациона тачност вероватно оптимистична (Ploton и др., 2020). Ово ограничење рад делом предупређује самом поделом по локацијама, а — што је важније — увођењем независне валидације, чији низак исход управо и чини оптимизам крос-валидације видљивим.

Балансирање и мере тачности, узети у целини, методолошки су добро постављени; примедбе овог одељка тичу се исказивања несигурности и чистоће процене, а не исправности носећег поступка.

VI Подручје, подаци и припрема

Припрема података (поглавље 4) описана је концептуално детаљно и с бројевима који омогућавају проверу: критеријум одабира 89 гранула (мање од 20 % облачности, покривеност целе територије у три сезонска интервала), скуп 20-метарских канала са SCL картом, три сезонска интервала, извођење топографских карактеристика и повезивање класа с националним прописом. Класификациона шема од осам типова изричито је и проследиво везана за EUNIS дефиницију шуме (стр. 1) и за Правилник о стаништима („Службени гласник РС“, бр. 35/2010), а рад се унапред ограђује од синтаксономских захтева тиме што фитосоциолошка имена уводи као помоћна: „свим класама су додељена и асоцијативна неформална фитосоциолошка имена, без синтаксономског значења“ (стр. 62). Та ограда је умесна и потврђује је кодекс фитосоциолошке номенклатуре, чија формална правила важе тек када се тврди синтаксономски статус (Theurillat и др., 2021).

Неколико места овог дела вреди прецизирати. Атрибути висине и из ње изведене топографске карактеристике рачунати су из продукта EU-DEM v1.1, чији метаподаци изричито наводе да је реч о дигиталном моделу површи (DSM), који над шумом одражава и висину крошњи (EEA/Copernicus, 2016); назив висина терена је утолико непрецизан, мада је ефекат на резултате занемарљив, јер је битно релативно висинско раслојавање класа, а не апсолутна кота, и јер је EU-DEM био најбољи расположиви бесплатни продукт — препорука је само да се атрибут прецизније именује. У дискусионој реченици на стр. 74 групе канала описане су лабаво — блиски инфрацрвени за B05–B08a и средњи инфрацрвени за B11/B12 — што одступа од назива у властитој Табели 5 и од званичног корисничког приручника мисије, у

коме су B05–B07 црвена ивица (red edge), а B11/B12 краткоталасни инфрацрвени (SWIR) опсег (ESA, 2015); канали су, додуше, недвосмислено идентификовани ознакама, па је реч о терминолошком, не о суштинском одступању. Постоји и мања недоследност у опису временског обухвата (стр. 66 наводи да месеци после октобра нису разматрани, док поједине грануле трећег интервала имају датуме с почетка новембра), као и уводне географске непрецизности (формулација да шуме покривају више од трећине Србије прекорачује податак цитираног извора, који Србију сврстава у средње шумовите земље с уделом од 29,1 %). Слика 16 приказује рефлектансу без јединице и без назначеног фактора скалирања и без наслова временске осе, што би требало допунити; будући да график служи приказу релативних фенолошких трендова, тумачење и закључци тиме нису доведени у питање (за захтеве приказа в. Slocum и др., 2009). Ниједна од ових ставки не дира резултате, али их у монографском издању треба ускладити.

Озбиљнија, начелна замерка овог дела тиче се поновљивости. Методологија је концептуално потпуно описана, али за потпуно независну репродукцију недостају: навод семена случајности, верзије програмског језика R и коришћених пакета, пуни хиперпараметри прилагођеног SMOTE и табеларни списак идентификатора 89 гранула (који су дати само графички). За те ставке постоји јасан писани темељ добре праксе: правило да се при поступцима с насумичношћу забележи семе случајности и да се верзионишу скрипте (Sandve и др., 2013, правила 4 и 6), начела налазивости и доступности података (Wilkinson и др., 2016), цитирање софтвера с бројем верзије (Smith, Katz и Niemeyer, 2016), цитирање скупова података с трајним идентификатором и тачном верзијом (Data Citation Synthesis Group, 2014), као и критеријуми поновљивости прилагођени управо геоинформатици (Nüst и др., 2018; Kedron и др., 2021). Део ове замерке ублажава то што случајна шума добро ради и с подразумеваним хиперпараметрима (Probst и др., 2019), што су тренинг и референтни подаци грађа трећих лица, и што се део детаља ослања на већ објављени рад аутора (Kovačević и др., 2020). Ипак, табеларни попис гранула, верзије пакета, семе случајности и кратка изјава о доступности били би најплоднија допуна целог рада и наводе се као препорука, не као недостатак који руши резултате.

VII Експерименти и оптимални параметри методологије

Четири узастопна експеримента доследно су спроведена и уредно извештена; нумеричке тврдње о предностима имплицитне репрезентације и о доприносу балансирања реконструишу се из табела и прилога без размимоилажења, а детаљни прилози с мерама перформанси свих тестираних модела омогућавају унутрашњу проверу (укључујући, на пример, потпун изостанак најмањих класа код стабла одлучивања). Оптимални модели — случајна шума, суседство 5×5, агрегација

простом аритметичком средином и балансиран скуп — дају крос-валидациону укупну тачност од 85,43 % и капу 0,79 за имплицитну, односно 83,62 % и 0,78 за експлицитну репрезентацију.

Две ограде овог дела већ су начелно образложене у одељку V и овде добијају конкретан облик. Прва: оптималност треба разумети као најбоље унутар намерно суженог, узастопног поступка с фиксираним хиперпараметрима, у коме су могуће пропуштене интеракције између компоненти; рад то ограничење изричито наводи и образлаже смањењем броја тестирања и заједничком основом све три технике, па је реч о транспарентној огради, а не о грешци — за похлепно сужавање простора претраге не постоји писани стандард који би га забрањивао, те се примедба и води само као оцена домета оптималности. Друга: стубичасти дијаграми (слике 18–22) немају бројчано означене вредности и на једној оси мешају репрезентацију и правило агрегације, па се мале разлике не могу очитати са слике; за експерименте 2–4 то надокнађују нумеричке табеле прилога, али за први експеримент дијаграми остају једини извор укупне тачности и капе, што отежава независну проверу баш тог корака. Уз то стоји и неколико редакцијских неусаглашености у ознакама (упут на прилог с табелама укупне тачности и капе за први експеримент не одговара садржају тог прилога; ознака правила агрегације оптималног модела разликује се између Табеле 14 и одговарајућег прилога), при чему саме вредности остају тачне, а разлика међу правилима агрегације је реда мерне несигурности. Све то треба уредити, али ништа од тога не мења ни избор оптималног модела ни носеће закључке.

VIII Коначна карта и валидација независним подацима

Најзрелији је онај део рада у коме се коначне карте суоче с независним подацима. Валидација независним скупом од 25 235 полигона шумских одсека државних шума даје знатно ниже вредности него крос-валидација (укупна тачност 64,02 % и капа 0,41 за имплицитну, односно 60,59 % и 0,37 за експлицитну карту). Рад тај пад не прикрива, него га разлаже: „постоји оправдана сумња да је значајан део одступања последица спроведене конверзије информације о идентитету шумских састојина у усвојене класе типова шумске вегетације“ (стр. 115), уз изричито навођење свих ограничења референтног скупа — конверзије шумарске типологије у осам класа, обухвата само државних шума, изостанка једне класе („у референтном скупу података нема класе која одговара шибљацима бора кривуља (8 – Pinion mugo), па је ова класа изузета из валидације“, стр. 110), неравномерног распореда полигона и разлике у шумској маски. Рад отворено именује и зависност излазних карата од маске: „Једна од идентификованих слабих тачака методологије јесте да поузданост излазних карата зависи од коришћене шумске маске“ (стр. 122).

Управо зато што је овај део тако поштен, примедбе које стоје тичу се начина исказивања, а не прикривања. Прво, поређење са сличним студијама и оцена применљивости методологије наслањају се на крос-валидационе мере, док знатно ниже вредности независне валидације у то поређење нису укључене; било би прецизније да рад на истом месту супостави обе групе вредности, тим пре што управо независни, вероватносни узорак — а не крос-валидацију, па ни просторну крос-валидацију — струка сматра правим начином оцене тачности карте (Olofsson и др., 2014; Stehman и Foody, 2019; Wadoux и др., 2021). Ова последња потпора иде и у корист рада: она показује да ни просторна крос-валидација не би била прави начин оцене карте, чиме независна валидација коју је рад спровео добија тежину исправног корака, а не пропуста. Друго, права тачност коначних карата остаје негде између оптимистичне крос-валидације и независне валидације оптерећене конверзијом типологије, па није тачно одређена — што рад не прикрива, будући да пресудну оцену свесно препушта додатној валидацији.

Треће, површине и удели класа (Табела 17) исказани су као тачкасте вредности без мере неодређености, иако цитирани стандард добре праксе препоручује интервале поверења за параметре тачности и површине и процену површина из референтног узорка (Olofsson и др., 2014, препоруке v и vi; Olofsson и др., 2013); пуна примена тог поступка захтевала би вероватносни узорак и теренску кампању изван обима рада, али вреди је навести као правац. Четврто, поређење удела класа с Националном инвентуром шума води се у процентима унутар шумске површине сваког скупа, што је исправно, али рад не именује изричито основицу за прерачун удела у квадратне километре; како је укупна шумска површина карте (маска око 43 %) знатно већа од инвентурне (29,1 %), потцењеност удела појединих класа иде упоредо с вишом апсолутном површином, па би те две основице требало експлицитно раздвојити — за то нема именованог стандарда, реч је о логичко-рачунском аргументу из саме грађе. Најзад, неколико ситних нумеричких неусаглашености треба усагласити: наведена укупна тачност експлицитне карте (60,59 %) одступа за око 0,02 процентна поена од вредности прерачунате из њене матрице конфузије (60,61 %); између збира референтних пиксела и сума колона матрица постоји необјашњена разлика реда хиљаду пиксела; описни број полигона (25 235) не поклапа се са збиром одговарајуће колоне Табеле 18 (24 277), што се разложно приписује изостављању једне класе и полигона несводивих на осам класа при конверзији и растеризацији, али рад ту редукцију не наводи изричито. Ниједна од тих неусаглашености не мења пикселску основицу валидације ни срачунате метрике, али их у монографском издању треба отклонити.

IX Слике, табеле, једначине и језик

Слике и табеле у највећем делу су доследне и прослеђиве, а квантитативна основа рада конзистентна је кроз цео текст. Примедбе се своде на неколико поновљивих ставки већ поменутих на својим местима: дијаграми без бројчаних вредности и без наслова појединих оса (слике 16, 18–22), једначине с приказним омашкама (3.4, 3.6, 3.8) и двострука ознака SU. Њима треба додати низ ауторских словних омашки, чисто лектурске природе, које не мењају смисао али указују на недовршену коректуру: „Sentnel-2“, „Имлицитна“, „енлг.“ уместо „енгл.“, „Carpinus orentalis“ уместо исправног облика Carpinus orientalis, „Коначна карте“ у натписима слика 25 и 26, као и упаривање имена једне класе с фитосоциолошким именом које у властитим табелама припада другој класи (стр. 74), при чему исправно упаривање јасно стоји на више других места.

Посебно место заслужује запис бројева. Кроз цео рад децимални знак је тачка (нпр. 85.43), а раздвајач хиљада размак (нпр. 1 231 783). Овде мерило треба поставити прецизно, а не строго: међународни стандард јединица допушта и тачку и зарез као децимални знак и управо размак прописује као раздвајач цифара у групама од по три (BIPM, 2019, т. 5.4.4), па запис бројева у раду није повреда тог стандарда — раздвајање хиљада размаком чак му тачно одговара. Приговор о децималном знаку стога није приговор против међународне норме него против српске правописне норме, која на језику на коме је рад писан тражи децимални зарез (Пешикан, Јерковић и Пижурица, 2010). Реч је, дакле, о доследно спроведеном, недвосмисленом избору који одступа од националне ортографске норме, а не о нетачности вредности; усклађивање са српском нормом било би пожељно у монографском издању.

Терминологија машинског учења на српском (нпр. омотач-методе за wrapper methods, ивични црвени за red edge) нема потврђен јединствен писани стандард који би те облике прописивао или забрањивао, па се о њима не може судити позивом на стандард; рад доследно даје енглески оригинал у загради, што је при непостојању устаљеног српског термина прихватљива пракса и уклања сумњу у значење. Задржавање изворних назива мисија и продуката (Sentinel, Copernicus) у ћириличком тексту такође је бранљиво као стручна пракса (Пешикан, Јерковић и Пижурица, 2010, т. 195).

X Научни апарат

Научни апарат је, узет у целини, уредан и савремен: систем аутор–година примењен је доследно, усклађеност позивања и пописа изразито је висока, нема посредног цитирања ни двоструких јединица, а попис од 203 јединице савремен је (претежни

део из последњег десетлећа). Темељна дела уже области — сензор Sentinel-2, стабло одлучивања и CART, случајна шума, SMOTE, капа, добра пракса оцене тачности, избор атрибута и генетски алгоритми — цитирана су из изворних јединица.

Преостале примедбе претежно су формалне и лако отклоњиве. Једно позивање нема одговарајућу јединицу у попису — „Lu и др. (2017)“ на стр. 103, на месту где се сопствени резултат пореди с туђим — и то је стварни пропуст који треба исправити додавањем тачне референце. Једна јединица наведена је у попису, а нигде се у тексту не цитира. Јединице пописа немају датум приступа, а јединице софтвера, R пакета и мрежних продуката ни број верзије ни адресу, што умањује поновљивост за верзионисане ресурсе; ту важе иста писана начела као у одељку VI (ISO 690:2021; APA, 2019; Smith, Katz и Niemeyer, 2016; Data Citation Synthesis Group, 2014). Две јединице формално су непотпуне (скраћен назив часописа без броја свеске, односно књига без издавача и места), а код двеју одредаба из Службеног гласника ознаке a/b примењене су недоследно, што их различит број гласника ипак недвосмислено разликује. Све су то допуне за уједначавање апарата, а не садржинске грешке; ниједна не дира носеће закључке.

XI Однос рада и реферата комисије

Реферат Комисије оцењује рад као научни и стручни допринос у ужим областима фотограметрије и даљинске детекције и геоинформатике, наводи пет научних доприноса и предлаже да се рад прихвати и упути на јавну одбрану. Тај део оцене, у делу који се тиче обима, метода и концепта, одговара стању у тексту: обим (188 страна, основни текст на 123 стране, шест поглавља, 29 слика, 20 табела, 203 библиографске јединице) слаже се с радом, а пет доприноса стварно су спроведени и поткрепљени резултатима.

Две оцене реферата, међутим, вреди одмерити према стању у тексту. Прва: реферат наводи да резултати „потврђују иницијално постављене истраживачке хипотезе“, док у самом раду формално постављених, оборивих хипотеза под тим именом нема — рад поставља главни и шест парцијалних циљева и неколико претпоставки. Уз то, потврда почива на крос-валидационим мерама, које рад сам проглашава пристрасним за оцену карте (стр. 109), док независна валидација даје знатно ниже вредности; оцена реферата је, дакле, нешто одлучнија него што поставка рада допушта. Друга је обична омашка реферата, а не рада: реферат на једном месту закључке смешта у седмо поглавље — „У седмом поглављу дати су закључци докторске дисертације“ — иако рад има шест поглавља и закључак је поглавље 6 (стр. 117–123), што и сам реферат на другим местима тачно наводи. Ниједна од ове две ставке не умањује вредност рада; прва је питање прецизности формулације, друга уредничка омашка реферата.

XII Шта остаје

Кад се све сабере, добија се оцена коју вреди изрећи у једном комаду, без ублажавања и без претеривања.

Дисертација Јована Ковачевића солидан је и самостално изведен методолошки рад који ужим областима фотограметрије и даљинске детекције и геоинформатике доноси нешто што им је недостајало: целовиту, аутоматизовану методологију картирања типова шумске вегетације из вишевременских Sentinel-2 података, примењену на целој територији Србије, и коначне карте за осам типова, са сопственом операционализацијом временске компоненте, промишљеним балансирањем класа и валидационом поставком која на два кључна места избегава два позната извора оптимистичке пристрасности. Тамо где се могла проверити из себе саме, чињеничка и рачунска основа рада у највећем делу држи. Изнад свега, рад је уздржан према сопственим резултатима: не проглашава победника где разлике нема, признаје да крос-валидација прецењује квалитет карте, спроводи и отворено објављује неповољну независну валидацију и изричито тражи додатну валидацију карата пре шире употребе — што је знак научне зрелости.

Слабости су, по оцени овог осврта, четвороструке и ниједна не руши носеће тезе. Прва је исказивање несигурности: ниједна мера перформанси није праћена расипањем између понављања крос-валидације ни тестом значајности, па се најситније разлике међу моделима не могу поуздано разликовати од случајне варијације — што погађа само изборе засноване на таквим разликама, док носећи избори почивају на разликама знатно већим. Друга је чистоћа процене: избор атрибута није угнежђен у спољну евалуациону петљу (Ambroise и McLachlan, 2002; Hastie и др., 2009, 7.10.2), а крос-валидација, иако исправно изостављена по локацијама, не уклања просторну аутокорелацију међу кластерисаним полигонима (Roberts и др., 2017; Ploton и др., 2020) — обе с безначајним ефектом на главни закључак, јер рад усваја комплетан скуп атрибута и уводи независну валидацију. Трећа је поновљивост: изостају семе случајности, верзије софтвера и пакета, табеларни попис гранула и изјава о доступности, што се данас с правом тражи као добра пракса (Sandve и др., 2013; Nüst и др., 2018; Kedron и др., 2021) и што овде остаје препорука за допуну, а не мана која дира резултате. Четврта је оцена карте: применљивост и поређење са студијама наслањају се на крос-валидацију, а не на независни вероватносни узорак који струка сматра правим мерилом тачности карте (Olofsson и др., 2014; Wadoux и др., 2021), при чему рад ту разлику не крије него је сам износи. Уз то стоји низ ситних, претежно уредничких и приказних омашки (прегледне табеле сензора, једначине 3.4/3.6/3.8, осе дијаграма, децимални знак, лектура, једно позивање без јединице), које у монографском издању треба отклонити.

Изнад свега тога стоји једна правилност: рад је уздржанији и брижљивији у мерењу и у извештавању о сопственим ограничењима него у прегледним и приказним деловима. Управо зато највећи део онога што треба поправити не тражи ново истраживање него доследно спровођење онога што је рад већ започео.

По оцени овог осврта, у евентуалном монографском издању или наставку истраживања најплодније би биле следеће измене. Прво, уз сваку меру перформанси приказати расипање између пет понављања крос-валидације (стандардну девијацију или интервал поверења) и, где је могуће, тест значајности, а тврдње о „оптималним“ моделима ускладити с тако исказаном несигурношћу. Друго, избор атрибута угнездити у спољну евалуациону петљу, ради чистије процене. Треће, приложити табеларни попис 89 гранула, верзије R-а и пакета, семе случајности и хиперпараметре прилагођеног SMOTE, уз кратку изјаву о доступности кода и (у мери у којој то допуштају власници) података. Четврто, оцену тачности коначних карата допунити независним вероватносним узорком с исказаним интервалима поверења за тачност и површине класа, и на истом месту супоставити крос-валидационе и независне вредности. Пето, отклонити прегледне и приказне омашке (табеле сензора, једначине, осе дијаграма, лектура, децимални знак усагласити са српском нормом, додати недостајуће и допунити непотпуне библиографске јединице). Уз ту врсту редакције — која у највећем делу не тражи ново истраживање него довршетак већ отворених линија — ова дисертација остаје вредан и употребљив допринос, чија би методологија и карте имали шта да кажу и струци и корисницима изван уже теме.

Литература

Предмет осврта

Ковачевић, Јован Д. *Картирање шумске вегетације на основу података сателитског осматрања Земље коришћењем техника машинског учења* [докторска дисертација]. Грађевински факултет Универзитета у Београду, Београд, 2021. (НаРДуС, https://nardus.mpn.gov.rs/handle/123456789/21969.)

Реферат Комисије за преглед, оцену и одбрану докторске дисертације Јована Д. Ковачевића, Грађевински факултет Универзитета у Београду, 24. децембар 2021.

Оцена тачности карата, крос-валидација и просторна валидација

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., Wulder, M. A. „Good practices for estimating area and assessing accuracy of land change“. *Remote Sensing of Environment*, 148, 2014, 42–57. DOI 10.1016/j.rse.2014.02.015.

Olofsson, P., Foody, G. M., Stehman, S. V., Woodcock, C. E. „Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation“. *Remote Sensing of Environment*, 129, 2013, 122–131. DOI 10.1016/j.rse.2012.10.031.

Stehman, S. V., Foody, G. M. „Key issues in rigorous accuracy assessment of land cover products“. *Remote Sensing of Environment*, 231, 2019, 111199. DOI 10.1016/j.rse.2019.05.018.

Roberts, D. R. и др. „Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure“. *Ecography*, 40(8), 2017, 913–929. DOI 10.1111/ecog.02881.

Ploton, P. и др. „Spatial validation reveals poor predictive performance of large-scale ecological mapping models“. *Nature Communications*, 11, 2020, 4540. DOI 10.1038/s41467-020-18321-y.

Wadoux, A. M. J.-C., Heuvelink, G. B. M., de Bruin, S., Brus, D. J. „Spatial cross-validation is not the right way to evaluate map accuracy“. *Ecological Modelling*, 457, 2021, 109692. DOI 10.1016/j.ecolmodel.2021.109692.

Избор атрибута, балансирање и подешавање модела

Ambroise, C., McLachlan, G. J. „Selection bias in gene extraction on the basis of microarray gene-expression data“. *PNAS*, 99(10), 2002, 6562–6566. DOI 10.1073/pnas.102102699.

Hastie, T., Tibshirani, R., Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2. изд. New York: Springer, 2009 (одељак 7.10.2). DOI 10.1007/978-0-387-84858-7.

Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., Nauss, T. „Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation“. *Environmental Modelling & Software*, 101, 2018, 1–9. DOI 10.1016/j.envsoft.2017.12.001.

Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P. „SMOTE: Synthetic Minority Over-sampling Technique“. *Journal of Artificial Intelligence Research*, 16, 2002, 321–357. DOI 10.1613/jair.953.

Probst, P., Wright, M. N., Boulesteix, A.-L. „Hyperparameters and tuning strategies for random forest“. *WIREs Data Mining and Knowledge Discovery*, 9(3), 2019, e1301. DOI 10.1002/widm.1301.

Kuhn, M., Johnson, K. *Applied Predictive Modeling*. New York: Springer, 2013. DOI 10.1007/978-1-4614-6849-3.

Поновљивост рачунарског и геопросторног истраживања; цитирање података и софтвера

Sandve, G. K., Nekrutenko, A., Taylor, J., Hovig, E. „Ten Simple Rules for Reproducible Computational Research“. *PLoS Computational Biology*, 9(10), 2013, e1003285. DOI 10.1371/journal.pcbi.1003285.

Wilkinson, M. D. и др. „The FAIR Guiding Principles for scientific data management and stewardship“. *Scientific Data*, 3, 2016, 160018. DOI 10.1038/sdata.2016.18.

Smith, A. M., Katz, D. S., Niemeyer, K. E. „Software citation principles“. *PeerJ Computer Science*, 2, 2016, e86. DOI 10.7717/peerj-cs.86.

Data Citation Synthesis Group. *Joint Declaration of Data Citation Principles* (ур. M. Martone). San Diego: FORCE11, 2014.

Nüst, D., Granell, C., Hofer, B., Konkol, M., Ostermann, F. O., Sileryte, R., Cerutti, V. „Reproducible research and GIScience: an evaluation using AGILE conference papers“. *PeerJ*, 6, 2018, e5072. DOI 10.7717/peerj.5072.

Kedron, P., Li, W., Fotheringham, S., Goodchild, M. „Reproducibility and replicability: opportunities and challenges for geospatial research“. *International Journal of Geographical Information Science*, 35(3), 2021, 427–445. DOI 10.1080/13658816.2020.1802032.

Стандарди јединица, записа, картографског приказа и научног апарата

Bureau International des Poids et Mesures. *The International System of Units (SI)*, 9. изд. Sèvres: BIPM, 2019 (т. 5.4.4).

Пешикан, М., Јерковић, Ј., Пижурица, М. *Правопис српскога језика*, измењено и допуњено издање. Нови Сад: Матица српска, 2010.

ISO 690:2021, *Information and documentation — Guidelines for bibliographic references and citations to information resources*, 4. изд. Женева: ISO, јун 2021.

American Psychological Association. *Publication Manual of the American Psychological Association*, 7. изд. Washington, DC: APA, 2019.

Slocum, T. A., McMaster, R. B., Kessler, F. C., Howard, H. H. *Thematic Cartography and Geovisualization*, 3. изд. Upper Saddle River: Pearson Prentice Hall, 2009.

Продукти, прописи и номенклатура

European Space Agency. *Sentinel-2 User Handbook*, ESA Standard Document, Issue 1 Rev 2, 24. јул 2015.

European Environment Agency / Copernicus. *EU-DEM v1.1* (дигитални модел површи), април 2016.

Copernicus Land Monitoring Service. *High Resolution Layer Tree Cover & Forests — Product User Manual* (референтна 2018), 2020.

Правилник о критеријумима за издвајање типова станишта, о типовима станишта, осетљивим, угроженим, ретким и за заштиту приоритетним типовима станишта и о мерама заштите за њихово очување. „Службени гласник РС“, бр. 35/2010.

Theurillat, J.-P., Willner, W., Fernández-González, F., Bültmann, H., Čarni, A., Gigante, D., Mucina, L., Weber, H. „International Code of Phytosociological Nomenclature. 4th edition“. *Applied Vegetation Science*, 24(1), 2021, e12491. DOI 10.1111/avsc.12491.

Формални оквир докторских студија (важећи у децембру 2021)

Правилник о докторским студијама на Универзитету у Београду (пречишћен текст 2021), „Гласник Универзитета у Београду“, бр. 191/2016, 212/2019, 215/2020, 217/2020, 228/21 и 230/21 (чл. 34).

Упутство за формирање репозиторијума докторских дисертација, Сенат Универзитета у Београду, 06 број 3222/XXVI-11, 14. децембар 2011.