

**ПРОСТОРНА ДИСТРИБУЦИЈА НЕСИГУРНОСТИ ИНТЕРПОЛАЦИЈЕ:
ПРИМЕР ИЗОТЕРМНЕ КАРТЕ СРБИЈЕ (1991-2009)**

Милутин Пејовић^{1*}, Бранислав Бајат¹, Јелена Луковић²

¹Универзитет у Београду – Грађевински факултет, Одсек за геодезију и геоинформатику, Булевар краља Александра 73, 11000 Београд

²Универзитет у Београду – Географски факултет, Студентски трг 3/3, 11000 Београд

Сажетак: Све шира примена геостатистичких метода моделовања просторно дистрибуираних феномена захтева и евалуацију квалитета производа (карата) добијених њиховом применом. На примеру израде карте средњих годишњих температура ваздуха у Србији за период од 1991. до 2009. год., добијене на основу релативно малог просторног узорка за територију целе земље (110 метеоролошких станица) приказна је метода оцене њене несигурности тј. квалитета. Применом симулационог моделовања Монте Карло поступком оцењена је несигурност карте добијене кригинг интерполацијом.

Кључне речи: Монте Карло симулације, кригинг, коефицијент варијације, температура ваздуха.

Увод

За потребе описивања географских феномена у компјутерском окружењу неопходно је дефинисати одговарајуће концептуалне моделе који ће бити основа за развој логичких односно физичких модела просторних података. Концептуални модели се заснивају на физички фундаменталном разумевању процеса који се моделира. Два основна концептуална модела која у пракси обухватају све просторне појаве су: а) модел базиран на ентитетима где сматрамо да је простор састављен од ентитета (објеката) који се описују својим атрибутима или својствима б) модел базиран на пољима, где замишљамо да се испитивана појава мења у простору као нека непрекидна математичка функција или поље (Vunough, P.A. and McDonnell, R.A., 1998). У првом случају појаве се у окружењу Географских Информационих Система (ГИС) приказују кроз векторски модел података (тачке, линије и полигони) док у другом случају користимо растерски модел (гридна структура). Највећи број просторних појава који се анализира кроз разна научна или стручна истраживања представља управо оне феномене који се концептуално представљају пољима, попут климатолошких или педолошких променљивих величина, итд. Те појаве се за подручје од интереса приказују картама које физички представљају растерски приказ добијен применом различитих метода моделовања (обично је то просторна интерполација) на бази узоркованих вредности на одређеним локацијама. Овако добијене карте, тј. растери најчешће представљају улазне податке за даље просторне анализе. Корисници на овај начин добијене податке углавном користе не размишљајући о њиховом квалитету односно о њиховој тачности.

¹ Е-mail: mpejovic@grf.bg.ac.rs

Рад представља резултате истраживања у оквиру пројеката III47014, III43007 и TR 36009 које финансира Министарство просвете, науке и технолошког развоја Републике Србије

Проблем третирања утицаја квалитета просторних база података на резултате просторних анализа наметнуо се већ на самом почетку употребе оваквих типова података, па стога и први радови из ове области датирају из средине осамдесетих година прошлог века (Kumi-Boateng & Yakubu, 2010). Истраживања и развој поступака за оцену квалитета производа добијених разним просторним анализама значајно се увећава средином деведесетих са све широм применом ГИС апликација.

Та чињеница наметнула је потребу за израдом извештаја о квалитету ГИС производа, у циљу обезбеђивања детаљне информације корисницима на основу којих ће бити могућа оцена погодности ГИС података за одређене примене. С једне стране, термин "квалитет просторне базе података" је сувише уопштен, док с друге стране термин "грешка података" има негативну конотацију, с обзиром да се појам "грешка" односи на нешто што се може кориговати. Појам грешке се обично везује за одступање од истините вредности. У ГИС анализама, у којима се користе просторни подаци, не постоји начин да се дође до истинитих вредности. Непознавање истините вредности и поузданости мерених вредности исказује се несигурношћу. Зато је термин несигурност просторних база података, прихваћен као свеобухватнији, јер се односи на меру опсега вредности атрибута, од којих сваки од њих може бити прихватљив, и може бити резултат примене одређених метода мерења као и метода интерполације.

Интересантни су резултати велике анкете спроведене крајем прошлог века међу корисницима дигиталних модела терена (ДМТ), једне од најчешће коришћених просторних база (Wechsler, 1999). Анкета је обухватила 200 учесника (фирми и институција) из 25 земаља широм света (учествовале су и фирме из наше земље). Поред питања, ког су профила корисници ових података и у које се сврхе они користе, главна питања била су усмерена и на то да ли корисници у свом раду узимају у обзир квалитет, односно, несигурност ДМТ-а, и колико пажње поклањају овом проблему. На питање да ли приликом коришћења ДМТ података воде рачуна о њиховом квалитету, односно да ли у свом раду узимају у обзир и несигурност ДМТ-а, потврдно је одговорило само 22% учесника анкете, да понекад воде рачуна о томе 27% анкетираних, док 27% о томе нису никада ни размишљали.

У случају модела поља квалитет производа треба вредновати кроз евалуацију атрибутске тачности (Joksić, D. & Vajat, B., 2004). Она се дефинише као тачност квантитативне вредности атрибута, уколико је он дат нумеричком вредношћу у одређеном домену или као коректно идентификована класа уколико је сам атрибут категоријска променљива.

Код примена ГИС операција утицај несигурност улазних података на финални производ, а сами тим и квалитет истог тог производа може се оценити применом Тејлоровог модела првог или другог реда, Розенблутовим методом и методом Монте Карло симулација (Heuvelink G.V.M., 1998).

У овом раду разматрана је примена Монте Карло геостатистичких симулација над просторно ретким узорком средњих годишњих температура ваздуха у Србији за период 1991-2009. год. у циљу добијања изотермне карте као и могућности оцене и презентације квалитета овог производа.

Обрада података одрађена је у *open-source* програмском пакету R (R Development Core Team, 2008), док је за визуелизацију добијених резултата коришћен комерцијални ArcGIS софтвер.

Геостатистичке симулације

Као производ геостатистичке предикције просторног атрибута, добија се "слика" (*image, capture* *енгл.*) истог тог атрибута која представља појавни облике те

просторне случајне функције² $Z(u)$. У случају кригинга тај појавни облик представља најбољу линеарну непомерену оцену (BLUE³) просторног атрибута (Бајат Б. и Благојевић Д., 2007). Кригингом је, поред оцене, могуће израчунати и варијансу грешке кригинг оцене на свакој локацији $\hat{Z}(u_i) - Z(u_i)$, где је $\hat{Z}(u_i)$ оцењена, а $Z(u_i)$ измерена вредност на локацији u_i , која се често назива кригинг варијанса (Isaaks & Srivastava, 1989). Међутим, не ретко постоји потреба да се изведе закључак о тачности оцене просторног атрибута на некој локацији. Уобичајени приступ у оцењивању тачности сасатојао у конструисању "Гаусовског" интервала поверења којег чине кригинг оцене и одговарајуће кригинг варијансе (Isaaks & Srivastava, 1989). Варијанса кригинг оцене је независна од величина вредности на основу којих је оцењена али не и од просторне конфигурације тих вредности, то својство се назива "хомоседактивност". Тако се појављују случајеви да је оцењена вредност варијансе иста у простору малих разлика вредности атрибута као и у простору са великим разликама вредности атрибута. (Goovaerts, 1998).

Кригинг интерполација омогућава оцену просторног атрибута на некој локацији уз минималну варијансу оцене. У том погледу, кригинг оцена представља најбољу оцену јер се за резултат добија минимална локална варијанса грешке: $\hat{Z}(u_i) - Z(u_i)$ (Isaaks & Srivastava, 1989). Међутим мана оцењивања просторних атрибута под условом најмањих квадрата огледа се у такозваној заобљености површи. Заобљеност се манифестује кроз мању варијацију просторног атрибута него што је то заиста случај у природи, тј. немогућности интерпретације локалних минимума и максимума. Ефекат заобљености је искључиво завистан од просторне конфигурације узорака што значи да је мање изражен у околини тачака узорковања и повећава се са удаљењем од истих. Из тих разлога, карте добијене кригинг интерполацијом нису сасвим погодне за анализу просторних појава код којих екстремне вредности играју значајну улогу. У том случају, препоручује се примена методе симулације код којих је узета у обзир могућност појаве екстремних вредности на одређеним локацијама. Геостатистичка симулација представља просторну надоградњу концепта Монте Карло симулација. Оваква надоградња стохастичких техника омогућава дубљу анализу неког просторног феномена.

Насупрот кригинг оцени, стохастичка симулација просторних атрибута нема за циљ минимизацију грешке оцене, већ је фокусирана на репродукцију једнако вероватних манифестација тог атрибута, које за резултат дају хистограм узорака и вариограм конзистентан улазним подацима (Goovaerts, 1998). Да би се један случајни процес у простору правилно интерпретирао, у стохастичком смислу, непходно је познавати функцију која описује вероватноћу његове појаве. Вероватноћа појаве реализације просторног атрибута дефинише се вишедимензионалном функцијом расподеле (CDF⁴). Креирање једнаковероватних реализација случајних функција, на основу дефинисане вишедимензионалне функције расподеле је суштина Монте Карло симулације. (Hernandez, 1991). Изналажење модела функције расподеле просторне случајне функције представља суштину интерпретације просторног атрибута. Како је случајна функција $Z(u)$ у потпуности дефинисана својом вишедимензионалном функцијом расподеле, просторна симулација се може замислити као креирање реализације случајне функције случајним одабиром из функције расподеле, при чему се задржавају основне карактеристике просторне варијабилности. Постојање податка који би условили појаву одређене вредности неког атрибута на одређеној локацији,

² У даљем излагању, случајном функцијом ће се сматрати сет случајних процеса на локацијама $(Z(u_i); u_i \in A)$ и биће означавана великим словом $Z(u)$, а њена манифестација са малим словом $z(u)$.

³ Best Linear Unbiased Estimation

⁴ Cumulative Distribution Function

поделило је методе стохастичке симулације на две основне групе: а) условљене и б) неусловљене симулације. О методама неусловљених симулација неће бити нарочито речи у овом раду, већ ће само бити наведен њихов допринос формулацији индиректних метода условљених симулација. Преглед литературе у којој су обрађене методе неусловљених симулација дата је у раду (Hernandez, 1991). Суштина условљених симулације може се описати тежњом да се креира више једнако-вероватних реализација просторног атрибута, које ће задржати просторна варијабилност (зависност) условљену мереним вредностима на локацијама узорковања.

Методе условљене симулације могу се даље поделити на директне и индиректне методе (Hernandez, 1991). Индиректни приступ условљеном симулирању ослања се на чињеницу да је кригинг оцена независна од оцене кригинг грешке. Тиме се створила могућност замене кригинг грешке симулираном кригинг грешком добијеном неусловљеном симулацијом. (Zhang, J. & Goodchild, M. F., 2003)

Методе директне условљене симулације нашле су већу примену у пракси због лакше имплементације у софтверским решењима (Hernandez, 1991). У овом раду, акценат је стављен на методе директне условљене симулације, тачније на методе директне секвенцијалне симулације.

Условљена секвенцијална симулација

Метода секвенцијалне симулације заснована је на декомпозицији вишедимензионалне функције расподеле у производ једнодимензионалних условљених вероватноћа (Isaaks E., 1991) :

$$F(z_1 \dots z_n; u_1 \dots u_n) = F(z_1; u_1) \cdot F(z_2; u_2 | Z(u_1) = z_1) \cdots F(z_n; u_n | Z(u_\alpha) = z_\alpha, \alpha = 1, \dots, n-1) \quad (1)$$

Вредност вишедимензионалне случајне функције могуће је израчунати секвенцијално, случајним одабиром вредности из претходно дефинисаних условљених једнодимензионалних расподела, тако што ће се свака претходно срачуната вредност додавати у сет вредности које условљавају вероватноћу реализације следеће променљиве.

Секвенцијална симулација је један од опште прихваћених метода за просторну симулацију. Постоји више алгоритама за секвенцијалну симулацију просторног атрибута и сви су базирани на једном општем алгоритму који следи следеће кораке (Srivastava, 1996):

1. Случајним одабиром изабере се тачка у којој се жели симулирати вредност просторног атрибута,
2. Једном од познатих метода, моделира се локална условна расподела просторног атрибута,
3. случајним одабиром једне вредности из униформне расподеле, одабере се број од 0 до 1, који представља вероватноћу придружену будућој симулираној вредности атрибута,
4. на основу одабране вероватноће и функције локалне расподеле вероватноћа израчунава се вредност атрибута на тој тачки,
5. тако симулирана вредност сврстава се у групу датих вредности,
6. кораци од 1 до 5 се секвенцијално понављају док се не исимулирају све тачке грида.

С обзиром на данашњи ниво развоја софтверских и хардверских ресурса велики број овако генерисаних гридова данас не представља проблем У већини апликација он је обично 100, мада је Норе (1986) је доказао да је само 19 реализација симулација довољно да се добију статистички употребљиви резултати са 95% нивоом поверења.

Различити приступ моделовању функције локалне расподеле вероватноћа чини једину значајну разлику међу методама секвенцијалне симулације. Сходно томе, издвојила су се два приступа у моделовању функције расподеле: а) параметарски и б) непараметарски.

Параметарски приступ подразумева изналагање параметара који аналитички описују условну расподелу. Уколико је просторни атрибут концептуализован као вишедимензионална нормална (Гаусовска) случајна функција, облик функције условне расподеле дефинисан је средњом вредношћу и ковариационом функцијом (вариограмом). Насупрот параметарском, непараметарски приступ не подразумева претпоставку о облику или аналитичком изразу функције условне расподеле.

Апроксимација функције условне расподеле - параметарски приступ

У случају стационарне мултивариационе нормалне (Гаусовске) случајне функције, параметри условне расподеле случајне функције $Z(u_0)$, средња вредност и варијанса, могу бити идентификовани обичним кригингом (Isaaks & Srivastava, 1989). Параметарски приступ моделирања CDF, заснива се на тврдњи да је расподела вероватноћа једне случајне променљиве, из стационарне n -димензионалне функције случајних променљивих, нормална са средњом вредношћу једнакој кригинг оцени и варијанси која је једнака одговарајућој кригинг "варијанси". (Isaaks E. H., 1991).

$$E\{Z(u_0)|n\} = \sum_{\alpha=1}^n \lambda_{\alpha} \cdot Z(u_{\alpha}) + (1 - \sum_{\alpha=1}^n \lambda_{\alpha}) \cdot E\{Z(u)\} \quad (2)$$

$$C\{Z(u_0)|n\} = C\{Z(u)\} - \sum_{\alpha=1}^n \lambda_{\alpha} \cdot Cov(u_0, u_{\alpha})$$

где су коефицијенти λ_{α} решење (кригинг) система једначина:

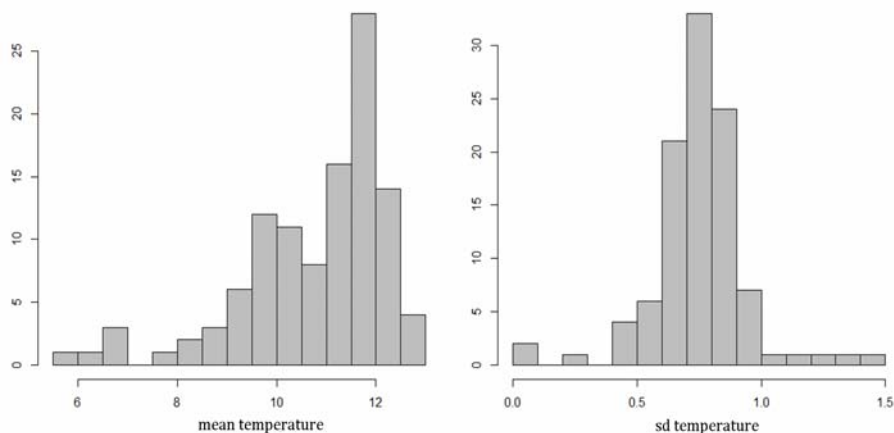
$$\sum_{\alpha, \beta=1}^n \lambda_{\alpha} \cdot C(u_{\beta} - u_{\alpha}) = C(u_0 - u_{\alpha}) \quad (3)$$

Тако добијени коефицијенти λ_{α} обезбеђују минималну вредност варијансе $C\{Z(u_0)|n\}$. Овај приступ подразумева нормалност и стационарност вишедимензионалне функције $Z(u)$. Међутим, уколико то није случај са улазним подацима, препоручује се трансформација улазних података како би се обезбедила нормалност расподеле.

Подаци

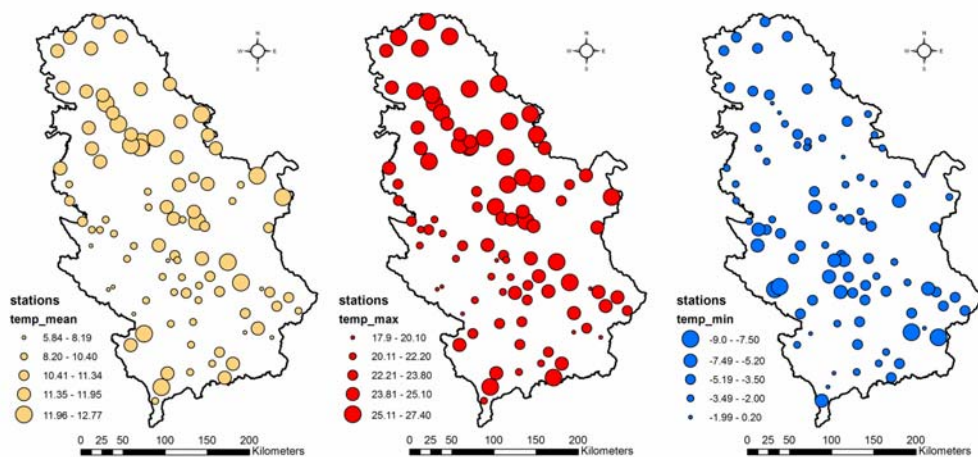
У раду су коришћене средње годишње температуре ваздуха у Србији, са 110 метеоролошких станица. Подаци су обезбеђени од стране Хидро-метеоролошког завода Србије. Сет коришћених података обухвата 28 синоптичких и 82 климатолошке станице. Иако је мрежа станица релативно добро распоређена на територији земље, важно је указати на недостатак станица у планинским областима Србије. Од укупног броја станица 87.3% се налази у распону од 0-500 m надморске висине, што представља укупно 62% територије државе. Између 500 m и 1000 m налази се 8.2% станица, а то је 27% површине Србије. Само 4.5% станица налази се на висинама већим од 1000 m. Иако Светска метеоролошка организације препоручује коришћење последњег стандардног периода (WMO, 1984) од 1961. до 1990. године, подаци у раду се односе на релативно новији период 1991-2009. Недостајући подаци у истраживаном периоду попуњени су коришћењем података са суседних станица. Претходна истраживања која су обухватила израду изотермне карте односе се најпре

на Атлас климе СФР Југославије, за период 1931-1960, а затим изотермну карту Србије публиковану од стране Дуцић и Радовановић (2005) за период 1961-1990. Средња вредност температуре над целим узорком за истраживани период је 10.8°C . Хистограми средњих температура по метеоролошким станицама као и њихове стандардне девијације приказани су на слици 1. Хистограм средњих температура је негативно закошен, што указује на већи број метеоролошких станица које имају средње температуре ближе горњој граници распона. Стандардне девијације средњих вредности температура по годинама имају средњу вредност $\pm 0.75^{\circ}\text{C}$ и имају приближно нормалну расподелу, где највећи број станица има стандардну девијацију у распону од 0.5°C до 1°C , што указује да у датом периоду није било великих варијација у средњим годишњим температурама.



Слика 1. Хистограми средњих годишњих температура ваздуха (лево) и стандардних девијација (десно) за период 1991-2009 год.

Просторна дистрибуција средњих, максималних и минималних вредности температуре по станицама дата је слици 2.



Слика 2. Карте са вредностима средњих (лево) максималних средњих (средина) и минималних средњих температура ваздуха (десно) за период 1991-2009 год.

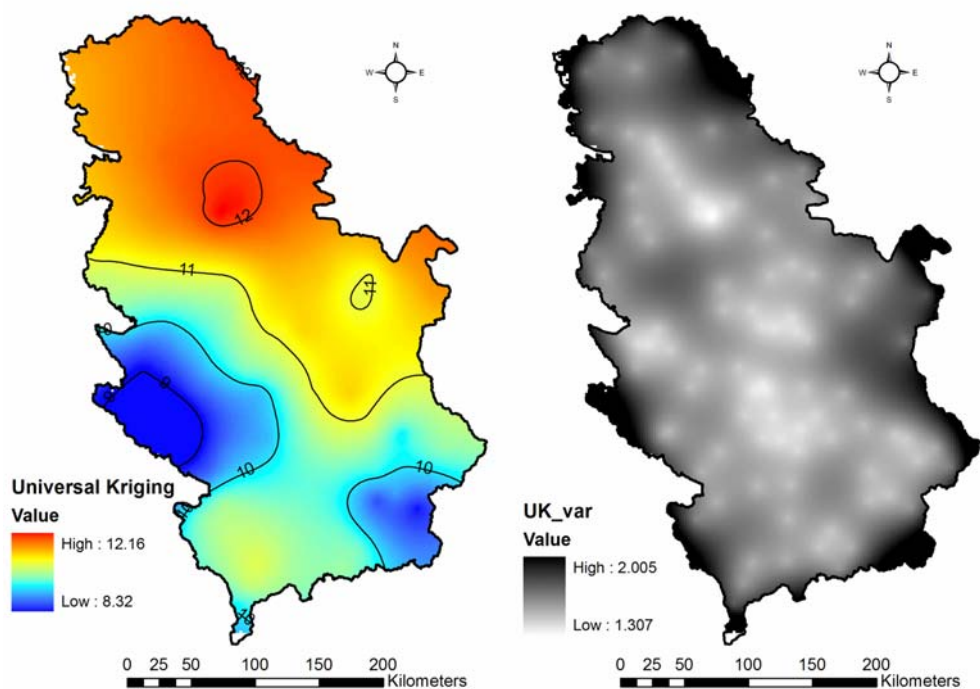
Резултати и дискусија

У првом кораку урађена је предикција средњих температура за период 1991-2009 применом методе универзалног кригинга код кога се тренд површи оцењује у функцији просторних координата. Добијена изотермна карта приказује заобљену површ на којој се не могу уочити екстремне локалне варијације (слика 3 лево).

Анализирајући добијену изотермну карту Србије за период 1991-2009 запажено је, као и у претходним истраживањима (Дуцић и Радовановић, 2005) да се област са средњом годишњом температуром ваздуха испод 10°C, простире углавном на надморској висини изнад 600 m. Ипак, важно је истаћи да се температуре испод наведеног прага уочавају и на нижим висинама што се може боље видети на карти добијеној симулацијама (слика 7 лево). То се односи на Бор, Жагубицу, Димитровград, Ужице и Приштину. У планинским областима изнад 1000 m надморске висине средње годишње температуре су око 6°C, а на висини изнад 1500 m око 3°C. Пажљивим одабиром и адекватном анализом одговарајућих профила било би могуће објаснити понашање средњих годишњих температура ваздуха у одређеним брдским и планинским појасевима. Са друге стране, температурне прилике у долинама Саве, Дунава, Тисе, Велике и Јужне Мораве имају средњу годишњу вредност између 11 и 12°C. Приказ дистрибуције максималних средњих годишњих температура ваздуха (слика 2) приказује да су у наведеним областима присутне највише средње годишње температуре ваздуха у посматраном периоду.

Осим надморске висине, на образовање температурних услова значајан удео имају морфологија и експозиција рељефа. Циркулациони процеси су динамичнији на стрмим падинама него у котлинама, у којима се ваздушне масе дуж задржавају. Пример који поткрепује утицај морфологије је Пештерска висораван која представља најхладнији део Србије. Све котлине, а нарочито оне високопланинске погодују образовању ниских температура и честих температурних инверзија. Ипак, у Сјеничкој котлини и на Пештерском пољу забележене су најниже температуре у Србији. Апсолутно минимална температура ваздуха од -39.5°C регистрована је 13. јануара 1985. године у Карајукића Бунарима у Пештерском пољу. Приказ минималних средњих температура ваздуха (слика 2) управо показује да се најниже средње годишње температуре ваздуха јављају у Рашкој и Власинској области. Према Ракићевићу (1971) ову су две од укупно четири области у којима се апсолутно минималне температуре спуштају испод -30°C. Хомољске планине и околина Вршца представљају преостала два тзв. пола хладноће. Метохијска котлина морфолошки погодује ујезеравању хладног ваздуха, али се на карти (слика 7) јасно види да преовлађују повољнији температурни услови. Наиме, маритимни утицаји у овом делу Земље условљавају нешто другачије температурне односе.

Уз добијену карту генерисана је и карта кригинг варијанси (слика 3 десно). Јасно се може уочити да су најмање вредности варијанси управо у локацијама које представљају метеоролошке станице, што је и очекивано јер кригинг предикција остварује најмање варијансе на мереним тачкама. Удаљавањем од мерених локација варијанса се постепено повећава. Највеће вредности варијанси су у граничним подручјима, што се може објаснити "ивичним ефектом", проблемом присутним код свих интерполационих метода, а који се може превазићи једино уколико имамо узорке изван граница картираног подручја.



Слика 3. Карте са моделованим вредностима средњих годишњих температура ваздуха за период 1991-2009 год. (лево) добијених применом универзалног кригинга (лево) и кригинг варијанси (десно)

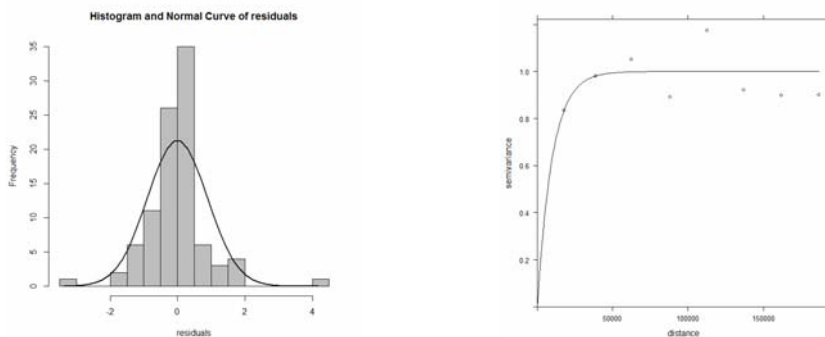
У циљу изналажења просторне дистрибуције несигурности просторног атрибута примењена је метода Гаусовске условљене секвенцијалне симулације, при чему су за улазне податке послужили резидуали добијени применом методе регресионог кригинга (Вајат, В. et al., 2012), чиме је обезбеђен услов нормалности мултивариационе просторне функције (слика 4 лево). Регресиони кригинг је интерполациона техника која користи методу регресије са помоћним (независним) предикторима у циљу оцене тренда зависне променљиве (величине коју интерполујемо). Код универзалног кригинга оцена тренда се рачуна само на основу координата узорка тражене променљиве. Код регресионог кригинга, након оцене тренда примењује се метода обичног кригинга над резидуалима који се добијају као остатак мерене вредности на датој локацији и оцењеног тренда.

Услед, природно високе корелисаности вредности температуре и надморске висине, као помоћни предиктор у поступку одређивања тренда искоришћен је SRTM дигитални модел терена (Самарџић, М. и Миленковић, М., 2010) који је за ову прилику са резолуције од 90 m ресемплован на резолуцију од 2 km.

Другим речима, креирање (симулирање) једнако вероватних манифестација средњих температура, замењено је симулирањем појава резидуала, који, додавањем на одређен тренд, представљају могуће појавне облике просторне расподеле температура. Просторна зависност резидуала моделована је вариограмом резидуала (слика 4 десно).

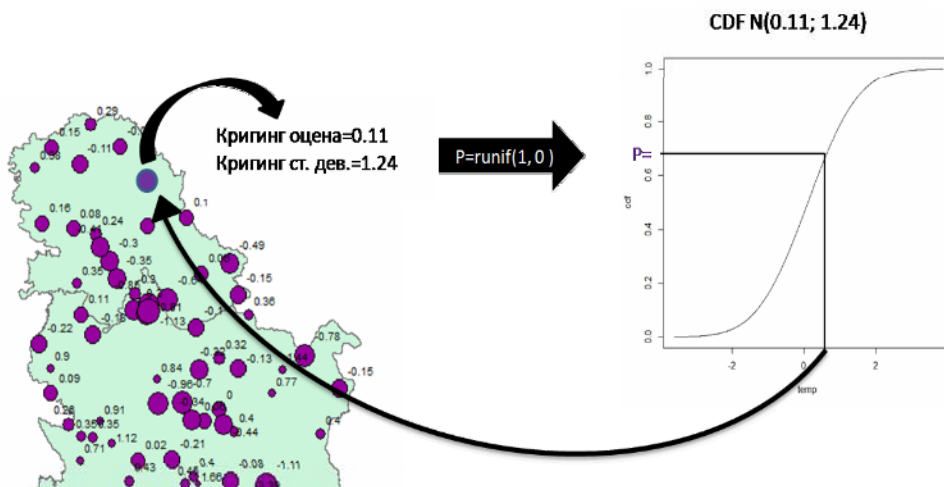
Кригинг систем калибрисан је над подацима средњих температура којима је придружена надморска висина, од којих је, у оквиру методе регресионог кригинга одузет тренд. Применом симулационог поступка над резидуалима, у свакој ћелији

грида (ресемпловани SRTM) генерисано је 100 могућих вредности резидуала случајним одабиром из функције локалне расподеле вероватноћа.



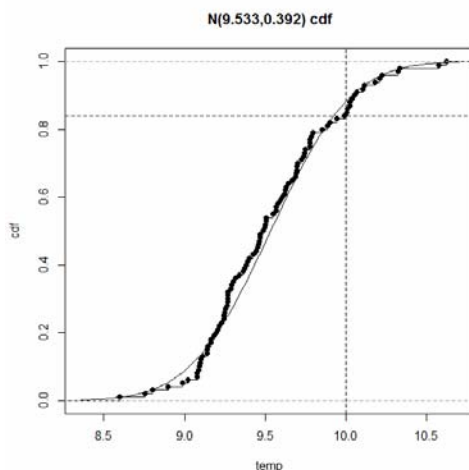
Слика 4. Хистограм и модел расподеле резидуала од тренда (лево), вариограм резидуала (десно)

Симулациони поступак спроведен је према алгоритму описаном у претходним поглављима, а који је имплементиран у оквиру "gstat" пакета у R окружењу. Графички приказ једног корака секвенцијалног алгоритма симулације над резидуалима приказан је на слици 5.



Слика 5. На основу кригинг оцене резидуала у насумично одабраној ћелији грида, формира се функција локалне расподеле, из које се случајним одабиром, на основу случајно одабраног једног броја из униформне расподеле (0,1) који представља вероватноћу, одређује једна могућа вредност.

Додавањем добијених резидуала на тренд добијено је 100 могућих вредности температура у свакој ћелији грида. Добијене вредности за једну ћелију грида послужиле су и за нумеричку апроксимацију функције расподеле на карактеристичним тачкама. (слика 6). На слици се види, да се са вероватноћом од скоро 90% може тврдити да је средња температура бити мања од 10 степени.

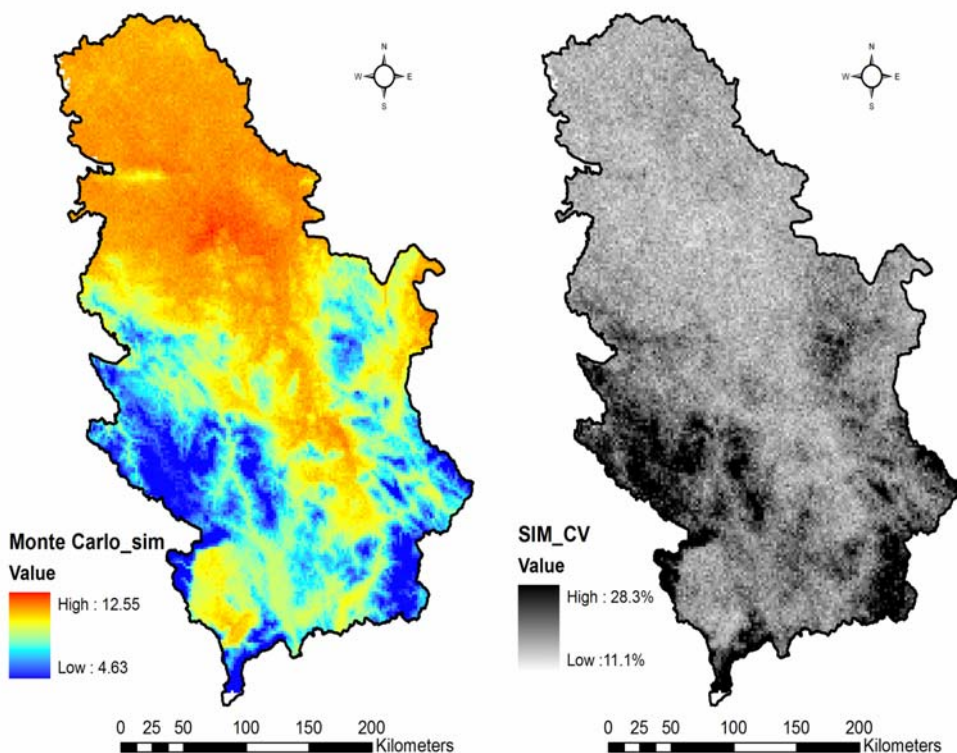


Слика 6. Модел емпиријске функције расподеле.

Осредњавањем симулираних вредности температура добијена је карта највероватнијих вредности (слика 7). У сврху анализе просторне дистрибуције грешака интерполације креиране су карте варијанси универзалног кригинга, као и карта коефицијента варијације (CV) добијеног на основу симулираних вредности температура. Коефицијент варијације као стандардна статистичка мера, даје информацију о варијабилности узорка у односу на његову средњу вредност, и рачуна се као количник стандардне девијације σ и аритметичке средине μ .

$$CV = \frac{\sigma}{\mu} \quad (4)$$

У нашем случају могуће срачунати вредности за CV за сваку ћелију грида на основу вредности које су добијене за 100 симулација, тако добијене вредности за CV осликавају несигурност финалне (осредњене) карте добијене Монте Карло симулацијама (Philips, D.L. & Marks D.G, 1996) која је резултат искључиво примењене интерполационе технике (регресионог кригинга у нашем случају). Вредности CV обично се дају у процентима. Највеће вредности за CV добијене су на подручјима са ниским средњим годишњим температурама (слика 7 десно). Поређењем карте са CV вредностима са стандардном картом кригинг варијанси, може се лако уочити да не постоје велике вредности за CV које би настале као резултат ивичног ефекта, као код већине интерполационих метода.



Слика 7 Карте са моделованим вредностима средњих годишњих температура ваздуха за период 1991-2009 год. (лево) добијених применом Монте Карло симулација (лево) и коефицијената варијације (десно).

Закључак

Примена геостатистичких метода интерполације данас је све присутнија у разним истраживањима и научним дисциплинама, то је и резултовало појавом различитих модалитета ових метода (обични, универзални, регресиони, индикаторски кригинг, итд). Једна од предности геостатистичке интерполације је базирана на томе што поред просторне предикције променљиве могуће је добити и визуелну интерпретацију варијанси интерполованих величина. Главни недостак геостатистичких интерполација је што угланом заобљавају локалне варијације променљиве. Овај недостатак може се превазићи применом геостатистичких симулација која резултује површима комплексније текстуре. Осредњена вредност из низа симулираних вредности може се третирати као највероватнија вредност.

Поред тога из симулираних вредности може се рачунањем коефицијента варијације добити и оцена несигурности интерполације, која се користи као оцена квалитета добијеног производа или као улазна несигурност уколико се исти тај производ користи у даљим ГИС анализама.

Литература:

- Бajat, Б., Благојевић, Д. (2007). Примена геостатистике у просторним анализама. *Гласник Српског географског друштва*, 87(2), 143-152.
- Bajat, B., Pejović, M., Luković, J., Manojlović, P., Ducić, V., Mustafić, S. (2012) Mapping average annual precipitation in Serbia(1961–1990) by using regression kriging. *Theoretical and Applied Climatology*. DOI 10.1007/s00704-012-0702-2
- Burrough, P.A. and McDonnell, R.A. (1998). *Principles of Geographical Information Systems Spatial Information systems 2nd ed.*. New York, Oxford University Press.
- Deutsch, C. V. (1992). *Annealing techniques applied to reservoir modeling and the integration of geological and engineering (well test) data*. Ph.D. Thesis. Stanford, California: Stanford University, http://scrfs.stanford.edu/documents/Theses/SCRF-Theses/1990-1999/1992_PhD_Deutsch.pdf
- Дуцић, В., Радовановић, М. (2005) *Клима Србије*, Завод за уџбенике и наставна средства, Београд, стр.212
- Goovaerts, P. (1998). Geostatistics in soil science: state-of-the-art and perspective. *Geoderma*, 1-45.
- Hernandez, G. J. (1991). *A stochastic approach to the simulation od block conductivity fields conditioned upon data measured at a smaller scale*. Ph.d. Thesis. Stanford, California: Stanford University, http://scrfs.stanford.edu/documents/Theses/SCRF-Theses/1990-1999/1991_PhD_Hernandez.pdf
- Heuvelink, G.V.M. (1998). *Error Propagation in Environmental Modelling with GIS*. Research Monographs in Geographical Information Science, Taylor&Francis.
- Hope, A.C.A. (1968). A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society B* 30, 582-598
- Isaaks, E. H., & Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics*. New York City, Oxford University Press
- Isaaks, E. H. (1991). *The application of Monte Carlo methods to the analysis of spatially correlated data*. Ph.D. Thesis. Stanford, California: Stanford University. http://scrfs.stanford.edu/documents/Theses/SCRF-Theses/1990-1999/1990_PhD_Isaaks.pdf
- Joksić, D., Bajat, B. (2004). Elements of Spatial Data Quality as Information Technology Support for Sustainable Development Planning. *SPATIUM* 11, 77-83.
- Joksić, D., Bajat B. (2005). Probability maps as a measure of reliability for intervisibility analysis. *SPATIUM* 12, 22-27.
- Kumi-Boateng, B., Yakubu, I. (2010). Assessing the Quality of Spatial Data. *European Journal of Scientific Research* 43(4), 507-515
- Philips, D.L., Marks D.G. (1996). Spatial uncertainty analysis: propagation of interpolation errors in spatially distributed models. *Ecological Modelling* 91, 213-229.
- Ракићевић, Т. (1971). Утицај рељефа на доњу границу температуре на примеру Сјенице и Златибора. *Зборник радова ПМФ, св. XVIII*, Београд.
- Самардић, М., Миленковић, М. (2010). Shuttle Radar Topography Mission – Доступност података и остварена тачност. *Гласник Српског географског друштва*, 90(1), 51-72.
- R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria
- Srivastava, M. R. (1996). *An Overview of Stochastic Spatial Simulation. Spatial accuracy assessment in natural Resources and Environmental Sciencee*. Fort Collins: U.S. Department of Agriculture.
- Wechsler, S. (1999). Results of the DEM User Survey <http://web.syr.edu/~srperlit/survey.html> .
- WMO (1984) World Meteorological Organization, 1984: *Technical Regulations, Vol. I*. WMO Publication No. 49. Geneva, Switzerland.
- Zhang, J., Goodchild, M.F. (2003). *Uncertainty in Geographical Information* Taylor & Francis, London.

SPATIAL DISTRIBUTION OF INTERPOLATION UNCERTAINTY: CASE STUDY OF ISOTHERM MAP OF SERBIA (1991-2009)

MILUTIN PEJOVIĆ^{1*}, BRANISLAV BAJAT¹, JELENA LUKOVIĆ²

¹*University of Belgrade – Faculty of Civil Engineering, Department of Geodesy and Geoinformatics, Bulevar kralja Aleksandra 73, 11000 Belgrade*

²*University of Belgrade – Faculty of Geography, Studentski trg 3/3, 11000 Belgrade*

Abstract: More widely used geostatistical methods for modeling distributed phenomena requires the evaluation of the quality of the product (maps) obtained by their application. The method of evaluation of uncertainty i.e. quality of the map was described on the example of a map of mean annual air temperature in Serbia for the period from the years 1991 to 2009 that was obtained from a relatively small number of samples for the whole country area (110 meteorological stations). The uncertainty of the map, obtained by kriging interpolation was evaluated by applying a Monte Carlo simulation modeling method.

Key words: Monte Carlo simulations, kriging, coefficient of variation, air temperature

Introduction

For the purpose of describing geographical phenomena in the computer environment, it is necessary to define appropriate conceptual models that should be the basis for the development of logical and physical models of spatial data. Conceptual models are based on the fundamental understanding of the physical processes being modeled. Two basic conceptual models that include all physical phenomena in practice are: a) models based on the entities where we consider that the space consists of physical entities (objects) with specific attributes or properties; b) field models, where we imagine that examined phenomena changes over the space as a continuous mathematical function or field (Burrough, PA and McDonnell, RA, 1998). In the first case described above, geographical phenomena are presented as vector data (points, lines and polygons) in Geographic Information Systems (GIS) environment, whereas raster models (grid structure) are used in the second case. Various reports that spatially analyze a great number of physical phenomena include variables such as climate and soil as fields in the conceptual model. The phenomena are presented as maps that represent the physical raster image obtained by the different modeling methods (basically a spatial interpolation) of the samples in particular locations. Such obtained maps i.e. rasters often represent the input data prepared for further spatial analysis. However, the users of such obtained products mostly apply the maps without reconsidering their quality or their accuracy.

Investigations that analyze the influence of spatial database quality on the results of spatial analysis dated back to the mid-eighties (Kumi-Boateng & Yakubu, 2010). Further research and development of procedures for assessing the quality of GIS products obtained as the results of various spatial analysis significantly increased in the mid-nineties with the broader usage of the GIS applications.

The demand for the preparation of reports on the quality of GIS products was imposed in order to provide users with detailed information that could help to evaluate the applicability of GIS data for particular uses. For example, it has been found that the term "the quality of the spatial database" is too general, and the commonly used term "data error" has a negative connotation due to the fact that it refers to something that can be corrected.

* ⁵ E-mail: mpejovic@grf.bg.ac.rs

The concept of error is related to the deviation from the true value. In GIS analysis where spatial data is used, there is no way to estimate the true value of the data. The outset of the true values and reliability of the input values is expressed as an uncertainty. Therefore, the term uncertainty of spatial databases is accepted as a comprehensive term because it pertains to the measure of the range of attribute values, where each of them may be acceptable.

Interesting results were obtained from a large survey of users of digital terrain models (DTM), which is one of the most commonly used spatial databases conducted at the end of last century (Wechsler, 1999). The survey incorporated 200 participants (companies and institutions) from 25 countries around the world (some companies in our country took part in this poll). Aside from queries about the professional profiles of the users and the purposes for which their data was used, main issues also focused on whether people in their workplace take into account the quality, i.e., the uncertainty of DTM, and also ascertained how much attention is paid to this problem. When asked whether they take care about the quality and uncertainty of DTM in their work, only 22% of people answered affirmatively, 27% sometimes take care about quality of DTM, while 27% never consider this issue.

In conceptual field models, the quality of the product should be determined by evaluating attribute accuracy (Joksić, D. & Bajat, B., 2004), which is defined as the accuracy of quantitative attribute values if they are given as a numerical value in a certain domain or as a correctly identified class if the attribute represents a categorical value. In GIS operations, the impact of the uncertainty of input data on the final product can be evaluated using a variety of methods including the Taylor method of the first or second order the Rosenblueth's method and the method of Monte Carlo simulation (Heuvelink, G.B.M., 1998).

In this paper we considered the application of Monte Carlo geostatistical simulations on the spatially sparse sample of average annual air temperatures in Serbia for the time period from 1991 to 2009 with the aim to obtain an isotherm map and to both evaluate and visualise its quality.

The data processing was done in *open-source* R software (R Development Core Team, 2008) whereas the obtained results were visualised in commercial ArcGIS software.

Geostatistical Simulation

A geostatistical estimation of a spatial attribute is an image (capture) of a spatial realization of this attribute, which is just one possible realization of the spatially random function⁶ $Z(u)$. The kriging prediction is the Best Linear Unbiased Estimation⁷ (Bajat, B. & Blagojević D, 2007) which provides the minimum local error variance. Besides the predicted values, kriging provides the error variance related to the estimate on any unsampled location $\hat{Z}(u_i) - Z(u_i)$, where $\hat{Z}(u_i)$ is the estimated value and $Z(u_i)$ is a measured (sampled) value. This error variance is often described as kriging variance (Isaaks & Srivastava, 1989) and often one also needs to include information about the uncertainty (reliability and accuracy) of a geostatistical estimation on any unsampled location. Traditional approaches for modeling local uncertainty are based on constructing Gaussian-type confidence intervals, which consist of kriging estimation and related kriging variance (Isaaks & Srivastava, 1989). Kriging variance is independent of the actual (measured) data values and depends only on data spatial configuration, a situation refers to as "homoscedasticity". There are certain cases when the estimated variance has the same value in the locations with a smaller data variation, as well as in locations with bigger data variations (Goovaerts, 1998)

⁶ Random function is a set of random process at locations u_i over the field $A(Z(u_i); u_i \in A)$, and will be denoted $Z(u)$, but realization will be denoted as $z(u)$.

⁷ Best Linear Unbiased Estimation

Kriging estimates provide the best estimates of spatial attributes at unsampled location in the least-square sense, because it provides the minimum of the local error variance (Isaaks & Srivastava, 1989). However, a shortcoming of the least-squares criterion is the generated smoothing effect of estimated surfaces. The smoothing effect is presented through a smaller variation of spatial attributes than is actually the case in nature, i.e. inability of local minimum and maximum interpretations. The smoothing effect only depends on spatial configuration of sampled data, which means that it is less expressed near the sampled location and increases with increasing distance from sampled data. For that matter, Kriging maps are not adequate for spatial analysis where extreme values play important roles, like in risk analysis. In this case, it's recommended to apply spatial (geostatistical) simulation methods which takes into account the probability of extreme values appearing and therefore provides more realistic results. Geostatistical simulation is an extension of the Monte Carlo simulation concept in a spatial sense, and provides a deep analysis of distribution patterns/trends of spatial attributes.

Unlike kriging, spatial simulation does not aim at minimizing a local error variance but instead focuses on reproducing equally probably realizations of attributes, where statistical details including the sample histogram or the semivariogram model, honor the overall statistics of data values (Goovaerts, 1998). Correct interpretation of any random field requires knowing the mathematical function which describes the probability of its realization. Such stochastic processes are completely characterized by multivariate cumulative distribution functions (CDF⁸). The ability to generate realizations of a spatially random function with a given multivariate distribution is a key issue to spatial analyses that use the Monte-Carlo method (Hernandez, 1991). On the other hand, defining the multivariate CDF is one of the most important tasks in the process of spatial simulation. Since a random function $Z(u)$ is completely characterized by its multivariate CDF, we can describe the spatial simulation as a technique that is drawing at random from the multivariate CDF of $Z(u)$ where the spatial distribution of $Z(u)$ remains unchanged.

However, if simulated values are conditioned by actual (or neighbor) data, the simulation is determined to be a conditional simulation. The simulation is determined to be unconditional if the simulated values remain insensitive to neighbor data. Direct methods of unconditional simulation are not considered in this paper, however their contribution to the formulation of indirect methods for assessing conditional simulations are discussed. A great literature review of unconditional simulation methods can be found a thesis written by Hernandez (Hernandez, 1991). The aim of conditional simulations is to create many equally probable realizations of spatial attributes that keep spatial distribution and other statistics unchanged.

Methods of conditional simulation can be divided into direct or indirect methods (Hernandez, 1991). The indirect approach is based on the fact that kriging estimation and kriging variance are independent. Thus, it is possible, to replace kriging variance with a simulated one that is generated by unconditional simulation (Zhang, J. & Goodchild, M. F., 2003).

Methods of direct conditional simulations are more commonly used in practice since they are simply implemented in software packages (Hernandez, 1991). In this paper, the focus is on the method of direct conditional simulation, namely the method of direct sequential simulation.

Conditional sequential simulation

⁸ Cumulative Distribution Function

The method of sequential conditional simulation is based on decomposition of multivariate CDF into a product series of univariate conditional distribution (Isaaks E., 1991).

$$F(z_1 \dots z_n; u_1 \dots u_n) = F(z_1; u_1) \cdot F(z_2; u_2 | Z(u_1) = z_1) \cdots F(z_n; u_n | Z(u_\alpha) = z_\alpha, \alpha = 1, \dots, n-1) \quad (1)$$

If all univariate conditional distributions in Eq. 1 are known, then one realization of the multivariate random function $Z(u)$ can be computed sequentially by randomly selecting values from each of pre-defined univariate conditional distribution.

Sequential simulation is the most frequently used method for spatial simulation. There are many algorithms for sequential simulation of spatial attributes and all are based on the same basic algorithms which follow the steps listed below (Srivastava, 1996):

1. choose at random a node to be simulated,
2. model the conditional CDF at the first location,
3. choose at random one value between 1 and 0 from uniform distribution, which represents a probability attached to simulated value,
4. depending on the estimated conditional CDF and chosen probability, draw the attribute value at chosen location.
5. include the newly simulated value in the set of conditioning data
6. repeat steps 1 through 5 until all grid nodes have a simulated value.

Nowadays, many simulation images can be generated relatively quickly due to the enormous increases in computer processing speed, memory and data storage capacity. In most cases, the number of images generated can be over 100, but Hope (1986) proved that only 19 realizations are enough to get statistically useful results, with 95% confidence level. The only significant difference between the various methods of sequential simulation is the different approach in modeling local CDF. Thus, there are two main approaches: a) parametric and b) nonparametric (indicator approach)

The parametric approach determines parameters of pre-assumed conditional CDF. If the random function $Z(u)$ has multivariate normal (or Gaussian) model, it is fully determined by the mean and variance (or by covariate function or variogram). The nonparametric approach does not determine the conditional CDF through its parameters, but rather the conditional CDF values are estimated directly from data for various threshold values.

Approximating of CDF - parametric approach

If the random function model $Z(u)$ of random fields is stationary and its multivariate CDF is Gaussian, then the parameters of univariate conditional CDF, mean and variance, can be identified by ordinary kriging (Isaaks & Srivastava, 1989). The parametric approach of modeling CDF is based on the assumption of a normal probability of a random variable with the mean equal to the simple Kriging estimation and the variance equal to the corresponding ordinary kriging variance (Isaaks E. H., 1991):

$$E\{Z(u_0) | (n)\} = \sum_{\alpha=1}^n \lambda_\alpha \cdot Z(u_\alpha) + (1 - \sum_{\alpha=1}^n \lambda_\alpha) \cdot E\{Z(u)\} \quad (2)$$

$$C\{Z(u_0) | (n)\} = C\{Z(u)\} - \sum_{\alpha=1}^n \lambda_\alpha \cdot Cov(u_0, u_\alpha)$$

where coefficients λ_α are solution of Kriging system equations:

$$\sum_{\alpha, \beta=1}^n \lambda_\alpha \cdot C(u_\beta - u_\alpha) = C(u_0 - u_\alpha) \quad (3)$$

The coefficients λ_α calculated from this system of equations provide a minimum variance $C\{Z(u_0)|n\}$ condition. This approach typically requires a prior normal score transform of data to ensure that at least the univariate distribution histogram is normal.

Data

The average annual temperature data from 110 meteorological stations were used in the paper. The data set is provided by Republic Hydro-Meteorological Service of Serbia and consists of 28 synoptical and 82 climatological stations. The stations network is very well distributed over the country with lack of stations in the mountainous areas. From the total number of stations, 87.3% are placed between altitudes of 0 m and 500 m that occupy 62% of the country's area. Another 8.2% of stations are located within altitudes that range from 500 m-to 1000 m, which covers 27% of country's area. Only 4.5% of stations are on altitudes higher than 1000 m. Although World Meteorological Organization (WMO, 1984) recommends the last standard period from 1961 to 1990, in this paper we have used the recent period from 1991-2009. Missing data points are fulfilled using neighboring stations. Previous investigations of isotherm maps in the region were done for the period 1931-1960 in the Atlas of Climate SFR Yugoslavia, as well as for Serbia (1961-1990) by Ducić and Radovanović (2005). The average temperature over the investigated period was 10.8°C. Histograms showing average temperatures as well as standard deviations are shown on figure 1. The histogram showing average temperatures has a negatively skewed distribution, which implies that a great number of meteorological stations with average temperatures are closer to the upper limit. The standard deviation has average value of $\pm 0.75^\circ\text{C}$ with relatively normal distribution. Most of the stations have standard deviation between 0.5°C and 1°C, which indicates that there are no large variations in average annual temperatures.

Figure 1 Histograms of average temperatures (left) and standard deviations (right) for the period 1991-2009.

Spatial distribution of average, maximal and minimal temperatures is shown on figure 2.

Figure 2 Maps with average (left) maximum average (middle) and minimum average temperatures (right) in the period 1991-2009

Results and Discussion

The prediction of average temperatures for the period 1991-2009 was performed using the method of universal kriging in which the trend surface is estimated as a function of spatial coordinates. The isotherm map shows smooth surface on which extreme local variations are not noticed (Fig. 3 left).

Analysis of the isotherm map of Serbia for the period 1991-2009 shows that the area with mean annual air temperature below 10°C extends generally at altitudes above 600 m. Our conclusion is in line with conclusions made in previous studies (Ducić and Radovanović, 2005) using different methods. However, it is important to add that temperature within the given threshold limits was observed at lower altitudes and is more obvious on a map obtained by simulations (Figure 7) around the areas of Bor, Žagubica, Dimitrovgrad, Užice and Priština. Mountainous terrain with altitudes above 1000 m had mean annual temperatures of around 6°C, while at altitudes above 1500 m it is around 3°C. Careful selection of appropriate profiles and their adequate analysis may help to explain the mean annual temperatures in mountainous areas. On the other hand, the temperature

conditions in the valleys of the Sava River, Danube River, Tisza River, Velika Morava

River and Južna Morava River have mean temperature between 11 and 12°C. The distribution of maximum mean annual air temperature (Figure 2) shows that these areas present the highest mean annual temperatures in the investigated period.

Apart from the criteria of altitude, morphology and aspect are also significant factors of temperature conditions. Circulation processes are more dynamic on steep slopes than in valleys, where the air masses are prolonged. An example that supports the influence of morphology is the coldest part of Serbia- the Pešter plateau. All valleys, particularly those in higher altitudes are suitable for cold air masses and frequent temperature inversions. However, the lowest temperatures in Serbia were recorded in the Pešter plateau and Sjenica valley. An absolute minimum temperature of -39.5°C was registered on 13 January 1985 at the Karajukića bunari station in Pešter plateau. The minimum average air temperature (Figure 2) shows that the lowest mean annual temperatures occur in Raška and Vlasina area. According to Rakićević (1971), these are two out of four areas in which the absolute minimum temperatures drop below -30 ° S. The Homoljske Mountains and areas around the town of Vršac are the other two cold areas. The Metohijska valley is morphologically suitable for developing cold air masses but the map (Figure 7) clearly shows favorable temperature conditions prevailing. The influence of the sea in this part of the country can cause slightly different temperature conditions.

Figure 3 Maps showing modled values of average annual air temperatures for the period 1991-2009 obtained by Universal kriging (left) and corresponding kriging variense (right)

A map of kriging variance was also generated (Fig. 3 right) in addition to the the resulting map of prediction. It can be clearly seen that the minimum values of variance occurs in locations of a weather stations, which was expected because the minimum kriging prediction variance are achieved at the measured points. However, the variance gradually increases with distance from those locations. The highest values of the variances are in the border areas and can be explained as an "edge effect". That problem exists for all interpolation methods and it can only be minimized if samples are taken outside of mapped areas.

A conditional sequential Gaussian simulation was applied in order to determine the spatial distribution of uncertainty for spatial attributes. The residuals obtained by the method of regression kriging (Bajat, B. et al., 2012) were used as input data and provided the normality condition of multivariate spatial function (fig. 4 left). Regression kriging is the interpolation technique that uses the regression method on ancillary (independent) predictors to determine the trend of the target variable (interpolated variable). The universal kriging method calculates the trend only on the basis of the coordinates of sample values of the target variable. After the trend has been determined, regression kriging can be applied to the ordinary kriging method on residuals, which are obtained by subtracting sampled values and estimated trend values at a given locations.

Due to the naturally high correlation between the temperature and altitude, an SRTM Digital Terrain Model (Samardžić M. and Milenković M., 2010) was used as an ancillary predictor to determine the trend. The original grid resolution of the SRTM DTM was resampled from 90 m to 2 km resolution.

Generating equiprobable realizations of the mean temperature is obtained by applying a simulation of the residual realization. Adding simulated residuals on a determined trend provides a probable realization of temperature.

Spatial dependence of residuals was modeled using a residual variogram that is depicted in figure 4 to the right. The kriging system was calibrated based on average temperature data for corresponding heights. The temperature trend was identified depending on the height within the regression kriging method. The simulation procedure applied on

residuals was provided at each grid cell (resampled SRTM) and a total of 100 possible realizations of residuals were generated by randomly selecting values from local CDF.

Figure 4 Histogram and distribution model of residuals (left), variogram of residuals (right)

The simulation procedure was performed in accordance with the algorithm described in the previous chapters, which is implemented in the "gstat" package in the R software environment. Graphical representation for one of the sequential steps of the simulation algorithm of the residuals is shown in Figure 5.

Figure 5The function of the local distribution was generated based on kriging estimation of residuals in randomly selected grid cell.

The 100 equiprobable values of the temperature in each grid cell were calculated by adding obtained residuals to the trend value. The calculated values for one grid cell were used for the numerical approximation of the distribution function at characteristic points (Figure 6). The figure shows that one can say that the average temperature is less than 10 degrees with a confidence of almost 90%.

Figure 6. Model of empirical function of distribution.

The most probable isotherm map was obtained by averaging simulated values of temperature (Figure 7 left). are created The Universal kriging variance map was created in order to analyze the spatial distribution of interpolation errors and, at the same time, the map of the coefficient of variation (CV) was calculated on the basis of simulated values of temperature. The coefficient of variation is a standard statistical measure which provides information on the variability of the sample with respect to its mean value and is calculated as the ratio of the standard deviation σ and mean μ .

$$CV = \frac{\sigma}{\mu} \quad (4)$$

In our case, it is possible to calculate the value of CV for each grid cell based on the values obtained by 100 simulations so that resulting values of CV reflect the uncertainty of the final (averaged) map obtained by Monte Carlo simulations (Philips, D.L. & Marks D.G, 1996). The obtained CV value is solely a result of applied interpolation techniques (regression kriging in our case) and is usually reported as a percentage. The highest values of CV were obtained in areas with low averaged annual temperatures (Figure 7 right). By comparing the map of the CV values with standard map of Kriging variance, it can be easily seen that the CV values do not arise as a result of the edge effect, as it often occurs when most interpolation methods are applied.

Figure 7 Maps of modeled values of average annual air temperature for the period 1991-2009 obtained by Monte Carlo simulations (left) and coefficient of variation (right).

Conclusion

The application of geostatistical interpolation methods today is increasingly present in the various studies and scientific disciplines and has resulted in the appearance of disparate modalities of these methods (ordinary, universal, regression, indicator kriging, etc). One of the advantages of geostatistical interpolation is that, in addition to spatial prediction

of the variables, the visual interpretation of interpolation variances is also available. The main shortcoming is that geostatistical interpolation smoothes local variations of the variable. This disadvantage can be overcome by geostatistical simulation, which generates more complex texture of output surfaces. The averaged value from the series of simulated values can be considered as the most probable value.

Also, it is possible to obtain the uncertainty for the estimation of interpolation by calculating the coefficients of variation from simulated values. The determination of these uncertainties could be used to evaluate the quality of the obtained product as well as the uncertainty of the input data if the product is used in further GIS analysis.

References

See References on the page 42