Savšek, S. 2013. Pomen testiranja hipotez v deformacijski analizi = The significance of hypothesis testing in deformation analysis. Geodetski vestnik 57, 3: 465-478.

DOI: http://dx.doi.org/10.15292/geodetski-vestnik.2013.03.465-478
http://drugg.fgg.uni-lj.si/ 4487/

POMEN TESTIRANJA HIPOTEZ V DEFORMACIJSKI ANALIZI

THE SIGNIFICANCE OF HYPOTHESIS TESTING IN DEFORMATION ANALYSIS

Simona Savšek

UDK: 519.87:528.1

Klasifikacija prispevka po COBISS-u: 1.02

IZVLEČEK

V geodeziji na podlagi kakovostnih opazovanj ter ustreznega matematičnega in stohastičnega modela pridobimo zanesljive rezultate. Izpolnjevanje nekaterih meril in zahtev pri merskih in računskih postopkih ugotavljamo z ocenjevanjem in testiranjem značilnih parametrov, pri čemer so postopki statističnega testiranja hipotez nepogrešljivi. V deformacijski analizi s statističnimi testi preizkušamo domneve o parametrih, ugotavljamo skladnost opazovanj in matematičnega modela, odkrivamo grobe pogreške v opazovanjih ter ugotavljamo skladnost domnevno stabilnih točk med terminskimi izmerami in določamo točke s statistično značilnimi premiki. Naloga geodeta je, da za obdelavo kakovostnih merskih podatkov uporabi pravilen in zanesljiv matematični model, ki zagotavlja visoko stopnjo zaupanja v rezultate in tako zmanjšuje naročnikovo tveganje. Postopki statističnega testiranja hipotez pomagajo pri odločitvah, ne morejo pa nadomestiti merskih in računskih postopkov v geodeziji.

ABSTRACT

In geodesy, reliable results are obtained with quality observations and the corresponding mathematical and stochastic models. One establishes if individual criteria and requirements during the measurement and calculation procedures are met by evaluating and testing typical parameters respectively – procedures of statistical hypothesis testing being indispensable part of the process. In deformation analysis, statistical tests are applied for testing the assumptions about the parameters, for assessing the conformity of observations and the mathematical model, for detecting of gross errors in observations, for identifying the conformity of allegedly fixed points between the epochs and for determining the points with statistically typical displacements. For the processing of quality measurement data, the task of a surveyor is to use a correct and reliable mathematical model, which provides a high degree of confidence in the results and reduces a client's risk. Methods of statistical hypothesis testing help experts to make decisions, but it is not possible to use them to replace measurement and calculation procedures in geodesy.

KLJUČNE BESEDE

statistično testiranje hipotez, stopnja značilnosti testa, porazdelitvena funkcija, kritična vrednost, dejansko tveganje

KEY WORDS

statistical hypothesis testing, significance level, distribution function, critical value, actual risk

1 UVOD

V geodeziji s kalibrirano mersko opremo in preizkušenimi merskimi postopki pridobimo veliko opazovanj, kar nam omogoča številne analize. Stalen tehnološki razvoj prinaša vse natančnejše in kakovostnejše podatke, hkrati pa se postavlja vprašanje o zanesljivosti pridobljenih podatkov in njihovi obdelavi. Analize geodetskih opazovanj so tako že zgodaj začele vključevati verjetnostni račun kot matematično osnovo statistike, na podlagi katere določimo verjetnost nekega dogodka.

Ker posamezen podatek ni nikoli popolnoma gotov in vsebuje premalo informacij, podatke razvrščamo v razrede in jih vzorčimo. Ugotavljamo značilne vrednosti vzorca in mere pričakovane vrednosti, na podlagi katerih testiramo množico pridobljenih merskih podatkov. Odločitve geodetskega izvajalca vedno temeljijo na domnevah, ki jih na podlagi testiranja hipotez sprejmemo ali zavrnemo. Kolikor večja je verjetnost, da so rezultati pravilni in nepristranski, toliko manjše je tveganje za naročnika. V članku želimo predstaviti pomen poznavanja postopkov statističnega testiranja pri obdelavi geodetskih opazovanj in izračunu zanesljivih rezultatov.

2 STATISTIČNO TESTIRANJE HIPOTEZ

V geodeziji pojem Gauss-Markov model (GMM) razumemo kot zvezo med opazovanji, neznankami in konstantami v mreži, njihove medsebojne korelacije ter matematično in statistično teorijo izravnave po metodi najmanjših kvadratov. GMM lahko v splošnem zapišemo (Caspary, 2000):

$$(\mathbf{l} - \mathbf{l}_0) + \mathbf{v} = \mathbf{A}\hat{\mathbf{x}}$$

$$\Sigma_{\mathbf{ll}} = \sigma_0^2 \mathbf{Q}_{\mathbf{ll}},$$

kjer je $\mathbf{l}$ – vektor opazovanj ($n \times 1$), $\mathbf{l}_0$ – vektor izračunanih približnih vrednosti opazovanj ($n \times 1$), $\mathbf{v}$ – vektor popravkov opazovanj ($n \times 1$), $\mathbf{A}$ – matrika koeficientov enačb popravkov – modelna matrika ($n \times u$), $\hat{\mathbf{x}}$ – vektor popravkov približnih vrednosti neznank po izravnavi ($u \times 1$), $\Sigma_{\mathbf{ll}}$ – kovariančna matrika opazovanj ($n \times n$), σ_0^2 – a priori referenčna varianca in $\mathbf{Q}_{\mathbf{ll}}$ – matrika kofaktorjev opazovanj ($n \times n$), n – število opazovanj, u – število neznank.

V postopku izravnave se običajno omejimo na preprost model, ki najbolje opisuje funkcijske zveze med količinami v mreži, predpostavke o tipu porazdelitve ter o korelacijah med njimi. GMM temelji na nizu predpostavk, da:

- opazovanja niso grobo pogrešena,
- matematični model pravilno opisuje zvezo med opazovanji in neznankami,
- izbran stohastični model primerno opisuje stohastične lastnosti opazovanj.

Za ocenjevanje kakovosti geodetske mreže uporabljamo postopke statističnega testiranja hipotez. Osnovni namen statističnega testiranja je ugotovitev, ali izbrani matematični in stohastični model *ne ustrezata* dejanskemu stanju (Harvey, 1991). Pomembno je tudi, da odkrijemo mogoče grobe pogreške v opazovanjih, ki jih ne smemo vključiti v izravnavo. V postopku deformacijske analize je ujemanje matematičnega modela z dejanskim stanjem še posebej pomembno, saj napačne predpostavke ali celo napake v modelu lahko interpretiramo kot deformacije, kar ima za posledico napačne ali celo nevarne odločitve.

Testiranje hipotez izvedemo z globalnimi in lokalnimi testi. Prvi podajo zgolj splošno informacijo o prisotnosti pogreškov v rezultatih merjenih količin ali o neskladju točk v dveh terminskih izmerah. Z njimi lahko potrdimo domnevo o prisotnosti grobih pogreškov v opazovanjih ali nestabilnih točk v mreži, brez možnosti njihovega lociranja. Z lokalnimi testi pa odkrijemo, katero opazovanje je grobo pogrešeno ali katera točka izkazuje značilen premik.

V postopku deformacijske analize je preverjanje GMM zelo pomembno, saj se napake v modelu ali slabo izbran model lahko napačno razumejo kot deformacije, kar vodi do napačnih sklepov. Statistično testiranje se uporablja na primer kot pomoč pri odločitvi, ali je posamezen popravek dovolj velik, da ga obravnavamo kot premik, ali ne.

Navadno predpostavimo, da so opazovanja normalno porazdeljena, kar lahko zapišemo kot $\mathbf{l} \sim N(\mathbf{Ax}, \Sigma_{\mathbf{ll}})$, kjer je vektor $\mathbf{x}$ vektor pravih vrednosti neznank, $\Sigma_{\mathbf{ll}}$ pa pripadajoča kovariančna matrika opazovanj. Po linearizaciji in izravnavi lahko zapišemo naslednje zveze, iz katerih je razvidno, katere količine so normalno porazdeljene:

$$\hat{\mathbf{x}} \sim N(\mathbf{x}, \Sigma_{\hat{\mathbf{x}}\hat{\mathbf{x}}})$$

$$\mathbf{v} \sim N(\mathbf{0}, \Sigma_{\mathbf{vv}}), \Sigma_{\mathbf{vv}} = \Sigma_{\mathbf{ll}} - \mathbf{A}\Sigma_{\hat{\mathbf{x}}\hat{\mathbf{x}}}\mathbf{A}^T.$$

$$\mathbf{l} + \mathbf{v} \sim N(\mathbf{Ax}, \Sigma_{\hat{\mathbf{l}}\hat{\mathbf{l}}}), \Sigma_{\hat{\mathbf{l}}\hat{\mathbf{l}}} = \mathbf{A}\Sigma_{\hat{\mathbf{x}}\hat{\mathbf{x}}}\mathbf{A}^T$$

2.1 Ničelna in alternativna hipoteza

V postopku testiranja hipotez postavimo ničelno in alternativno hipotezo. Ugotavljamo, ali obstaja dovolj statističnih dokazov, da lahko alternativno hipotezo štejemo za pravilno. V splošnem za vsa testiranja velja, da ničelne hipoteze ne moremo zavrniti, če je testna statistika T manjša od kritične vrednosti T_α, ki jo izračunamo ali preberemo iz tabel, v nasprotnem primeru pa jo zavrnemo in sprejmemo alternativno hipotezo. Od vrste porazdelitve, po kateri se porazdeljuje testna statistika T, je odvisno, kako jo izračunamo ali iz katerih tabel vzamemo vrednosti za T_α (normalna, studentova, porazdelitev χ^2, porazdelitev F idr.).

Rezultate testiranja hipotez lahko strnemo v dve trditvi:

1. Zavrnitev ničelne hipoteze pomeni sprejetje alternativne hipoteze ob stopnji tveganja α. Ugotovimo lahko, da obstaja dovolj statističnih dokazov, in alternativno hipotezo lahko sprejmemo kot pravilno.
2. Če ne zavrnemo ničelne hipoteze, ugotovimo, da ni dovolj statističnih dokazov, in alternativne hipoteze ne moremo sprejeti kot pravilne.

Splošna linearna hipoteza ima obliko (Caspary, 2000):

$$H_0 : \mathbf{H}^T\mathbf{x} - \mathbf{g} = \mathbf{0},$$

kjer je $\mathbf{H}^T$ – matrika reda $(m \times u)$, ki izpolnjuje pogoja: $rang(\mathbf{H}) \leq m$, $1 \leq m \leq u$. Enačba predstavlja m linearnih enačb oblike $\mathbf{h}_i^T\mathbf{x} - g_i = 0$, $i \in \{1, 2, ..., m\}$.

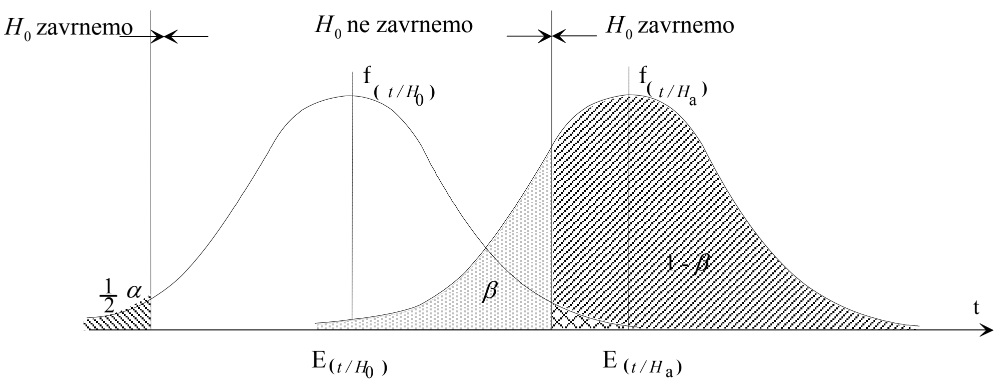
Vrstice $\mathbf{h}_i^T$ matrike $\mathbf{H}^T$ morajo biti linearne kombinacije vrstic matrike koeficientov enačb popravkov $\mathbf{A}$ v GMM, kar zahteva obstoj takšne matrike $\mathbf{L}$, ki izpolnjuje pogoj $\mathbf{LA} = \mathbf{H}^T$. Vektor $\mathbf{g}$ vsebuje konstante, ki izhajajo iz hipotez. Zapis posebne *ničelne hipoteze* v prvem koraku postopka testiranja sledi transformaciji osnovne oblike linearne hipoteze, na primer $H_0 : x_i - x_k = 0$.

Iz oblike $\mathbf{x} = (x_1, x_2, ..., x_i, x_j, x_k, ..., x_u)^T$ sledi $\mathbf{h}_i^T = (0, 0, ..., -1, 0, +1, ...0)$ in $\mathbf{g} = \mathbf{0}$. Torej lahko zapišemo ničelno hipotezo oblike:

$$H_0 : \mathbf{h}^T\mathbf{x} = 0.$$

Mejno vrednost za zavrnitev ničelne hipoteze imenujemo *kritična vrednost*. Kritična vrednost torej postavi mejni okvir za zavrnitev hipoteze in je določena z izbiro *stopnje značilnosti testa* α. Verjetnost, da presežemo kritično vrednost kljub izpolnjenim predpostavkam ničelne hipoteze, je enaka α. Z drugimi besedami, α pomeni verjetnost za neupravičeno zavrnitev pravilne ničelne hipoteze. Če je ničelna hipoteza H_0 zavrnjena, sprejmemo *alternativno hipotezo* H_a.
Značilne oblike alternativnih hipotez:

$$\begin{aligned} H_a :\ & \mathbf{h}^T\mathbf{x} \neq g \\ & \mathbf{h}^T\mathbf{x} > g \\ & \mathbf{h}^T\mathbf{x} < g \\ & \mathbf{h}^T\mathbf{x} = \overline{g} \neq g \end{aligned}$$



Slika 1: Dvostranski test v odvisnosti od stopnje značilnosti testa in verjetnosti za napako II. vrste glede na postavljeno alternativno hipotezo (Caspary, 2000)

Po izravnavi opazovanj je za ugotavljanje skladnosti med opazovanji in modelom koristno uporabiti statistične teste. Mogoči vzroki za morebitno neskladje, ki ga izkazuje statistični test, so:

1. **Funkcionalni model**, kar lahko pomeni, da so vzrok za neujemanje med opazovanji in matematičnim modelom lahko neupoštevani sistematični pogreški instrumenta ali neustrezno izbran model refrakcije ipd. Praviloma naj bodo enačbe v matematično-funkcionalnem modelu vsaj za stopnjo natančnejše od natančnosti opazovanj.
2. **Stohastični model**, kar lahko pomeni napačne predpostavke o a priori varianci, utežeh opazovanj in korelacijah med opazovanji.
3. **Opazovanja**, kar lahko pomeni grobe pogreške v opazovanjih, napake pri vnosu vhodnih podatkov, pogrešek centriranja idr.
4. **Izračuni**, pri katerih se lahko pojavijo numerična nestabilnost pri računskih operacijah in kopičenje napak pri zaokroževanju idr.

2.2 Stopnja tveganja in kritična vrednost

Po postavitvi ničelne in alternativne hipoteze moramo iz dejanskih opazovanj izračunati primerno testno statistiko T. Testna statistika se lahko porazdeljuje po eni od znanih porazdelitvenih funkcij ali pa porazdelitveno funkcijo določimo s simulacijami (Savšek-Safić, 2002). V splošnem presoja postavljenih hipotez temelji na primerjavi vrednosti testne statistike T z njeno kritično vrednostjo T_α glede na izbrano stopnjo značilnosti testa α, pri čemer je treba upoštevati tip porazdelitvene funkcije in vrsto statističnega testa (enostranski ali dvostranski).

Največjo dovoljeno stopnjo tveganja za napako prve vrste α imenujemo *stopnja značilnosti testa*. Njeno velikost predpišemo glede na posledice napačne odločitve. Čeprav se pri postopku testiranja uporabljajo stroga matematična pravila – kar nam da občutek, da gre za rigorozen postopek –, je izbira stopnje značilnosti popolnoma poljubna. Ne obstaja niti prava niti napačna stopnja značilnosti, kakor tudi ne merilo ali pravilo za izbor njene najbolj prave vrednosti. Pomembno je poudariti, da se je o stopnji tveganja treba odločiti pred nadaljnjimi vmesnimi izračuni, ki bi lahko vplivali na njen izbor, in da so pri tem zelo pomembne izkušnje. Izbira primerne stopnje značilnosti testa je odvisna od namenov testiranja. Pri odkrivanju grobih pogreškov ali testu skladnosti lahko privzamemo popolnoma drugačne vrednosti stopnje značilnosti testa. Običajne vrednosti stopnje značilnosti so med vrednostmi α = 10 %, α = 5 %, α = 1 % in α = 0,1 %. Učinkovitost izbrane stopnje značilnosti testa je odvisna tudi od občutljivosti testne statistike (Caspary, 2000).

Statistični test predpisuje pravilo za določitev podmnožice prostora vzorcev, ki jo imenujemo *kritično območje preizkusa* ali *območje zavrnitve* ničelne hipoteze ob stopnji tveganja α. Če dani vzorec pripada tej podmnožici, ničelno hipotezo zavrnemo, sicer pa tega ne moremo storiti.

Nadaljnji korak v postopku testiranja je določitev *kritične vrednosti* T_α pri izbrani stopnji značilnosti α. Za predpisano stopnjo značilnosti običajno iz tabel, grafov ali z računalniškimi grafi določimo ustrezno kritično vrednost T_α. Če ničelno hipotezo zavrnemo, je tveganje za napako prve vrste manjše od α. Danes nam računalniški programi omogočajo izračun (in ne tabelaričnega odčitavanja) *mejne stopnje značilnosti*, pri kateri lahko v konkretnem primeru ničelno hipotezo še zavrnemo, torej *stopnjo tveganja za napako I. vrste* v primeru zavrnitve pravilne ničelne hipoteze.

Pri izbrani stopnji značilnosti testa α primerjamo vrednost izračunane testne statistike T s kritično vrednostjo T_α. Nastopita lahko dva primera:

1. $T \leq T_\alpha$: ničelne hipoteze ne zavrnemo. V tem primeru statistično testiranje ne kaže na statistično značilno neujemanje opazovanj in izbranega matematičnega modela;
2. $T > T_\alpha$: ničelno hipotezo zavrnemo. V tem primeru statistično testiranje kaže na statistično značilno neujemanje med opazovanji in izbranim matematičnim modelom.

V postopku statističnega testiranja presojamo o pravilnosti postavljenih hipotez, zato so mogoči štirje primeri (preglednica 1):

- Ničelna hipoteza je dejansko pravilna in je ne zavrnemo: *odločitev je pravilna*.
- Ničelna hipoteza je pravilna, a jo zavrnemo – sprejmemo alternativno hipotezo; verjetnost takega dogodka je enaka stopnji značilnosti testa α – naredimo napako I. vrste: *odločitev je napačna*.

- Ničelna hipoteza je dejansko napačna in jo zavrnemo – sprejmemo alternativno hipotezo: *odločitev je pravilna*.
- Ničelna hipoteza je napačna, a je ne zavrnemo; verjetnost takega dogodka je enaka vrednosti β – naredimo napako II. vrste: *odločitev je napačna*. V statistiki imenujejo verjetnost zavrnitve napačne ničelne hipoteze »moč statističnega testa«, označi se z 1 – β.

Trditev/odločitev	**Ne moremo zavrniti** H_0	**Zavrnemo** H_0
H_0 pravilna	Pravilna odločitev: $P = 1 - \alpha$	**Napaka I. vrste:** stopnja tveganja = α
H_0 napačna (H_a pravilna)	**Napaka II. vrste:** stopnja tveganja = β	Pravilna odločitev: $P = 1 - \beta$ (moč testa)

Preglednica 1: Odločitve v postopku statističnega testiranja

S testiranjem hipotez se torej rešujeta dve osnovni težavi. Na eni strani gre za določitev začetne vrednosti stopnje tveganja za napako I. in II. vrste, na drugi strani pa za izračun skladne stopnje značilnosti ter kritične vrednosti testov. Na žalost ni mogoče doseči, da bi bila tako α kot β majhni vrednosti. Če zmanjšamo stopnjo tveganja za napako I. vrste, hkrati narašča možnost za napako II. vrste, saj se β povečuje, moč testa pa pada. Z zmanjševanjem stopnje značilnosti testa α se zmanjšuje verjetnost izločitve dobrih opazovanj, vendar se zmanjšuje tudi verjetnost odkrivanja grobih pogreškov v opazovanjih. Treba je najti pravilno ravnovesje med stopnjama tveganja za obe napaki. Glede na posledice, ki jih lahko povzroči napačna odločitev, ter glede na naravo obravnavane naloge se odločimo za numerične vrednosti.

3 TESTIRANJE HIPOTEZ V DEFORMACIJSKI ANALIZI

3.1 Globalni test modela

Za testiranje skladnosti opazovanj in matematičnega modela uporabimo a posteriori variančni test, v strokovni literaturi znan kot *globalni test modela*. Z njim testiramo skladnost referenčne variance a posteriori z referenčno varianco a priori σ_0^2. Globalni test modela izvedemo po izravnavi in je mogoč, če poznamo natančnost opazovanj in referenčno varianco a priori. Če njunih vrednosti ne poznamo dovolj zanesljivo, globalnega testa modela ne izvajamo.

Zapišemo ničelno hipotezo:

$H_0 : \hat{\sigma}_0^2 = \sigma_0^2$; referenčni varianci sta statistično skladni.

Tvorimo testno statistiko (Caspary, 2000):

$$T = \frac{\mathbf{v}^T \mathbf{P} \mathbf{v}}{\sigma_0^2} = \frac{r \cdot \hat{\sigma}_0^2}{\sigma_0^2} \sim \chi^2(r),$$

ki se porazdeljuje po porazdelitvi $\chi^2(r)$ z r prostostnimi stopnjami. Izračunamo kritično vrednost $\chi^2_{1-\alpha/2}(r)$ ob izbrani stopnji značilnosti testa α. Testiranje globalnega testa lahko izvedemo tudi s porazdelitveno funkcijo F.

Če velja $T \leq \chi^2_{1-\alpha/2}(r)$, ničelne hipoteze ne zavrnemo, saj test ne kaže na statistično neujemanje med opazovanji in matematičnim modelom. Če velja $T > \chi^2_{1-\alpha/2}(r)$, zavrnemo ničelno hipotezo bodisi zaradi slabo ocenjene natančnosti opazovanj bodisi zaradi prisotnosti grobih pogreškov v opazovanjih (Kuang, 1996). Obstaja tudi možnost, da zavrnemo ničelno hipotezo, čeprav testna statistika ne doseže kritične vrednosti. To se zgodi, če smo podcenili natančnost opazovanj, torej ima referenčna varianca a priori preveliko vrednost (Grigillo, Stopar, 2003).

3.2 Testiranje prisotnosti grobih pogreškov

Eden od razlogov za zavrnitev ničelne hipoteze pri globalnem testu modela je lahko tudi prisotnost grobih pogreškov v opazovanjih. Obstaja več metod iskanja in lociranja grobih pogreškov, vse pa predpostavljajo, da so opazovanja in popravki opazovanj normalno porazdeljeni. Če zanesljivo poznamo referenčno varianco a priori σ_0^2, je najučinkovitejša metoda iskanja grobih pogreškov metoda pregleda opazovanj (angl. data snooping), ki jo je leta 1968 predlagal Baarda. Po zavrnitvi globalnega testa grobi pogrešek v opazovanjih odkrijemo na podlagi izračuna standardiziranega popravka (Caspary, 2000). Ker v praksi redko dovolj zanesljivo poznamo referenčno varianco a priori σ_0^2, pogosto uporabljamo test τ, ki ga je leta 1976 predlagal Pope.

Zapišemo ničelno hipotezo:

$H_0 : E(v_i) = 0$; popravki opazovanj so statistično enaki nič.

Tvorimo testno statistiko (Caspary, 2000):

$$T = \frac{|v_i|}{\hat{\sigma}_{v_i}} = \frac{|v_i|}{\hat{\sigma}_0 \sqrt{q_{v_i v_i}}} \sim \tau(r),$$

ki se porazdeljuje po porazdelitvi $\tau(r)$ z r prostostnimi stopnjami. Ker so preglednice za porazdelitev τ težje dostopne od porazdelitve t, obstajajo enačbe za preračun kritičnih vrednosti (Kuang, 1996). Izračunamo kritično vrednost $\tau_{1-\alpha/2}(r)$ ob izbrani stopnji značilnosti testa α.

Če velja $T \leq \tau_{1-\alpha/2}(r)$, ničelne hipoteze ne zavrnemo, saj test ne kaže na prisotnost grobih pogreškov v opazovanjih. Če velja $T > \tau_{1-\alpha/2}(r)$, zavrnemo ničelno hipotezo in upravičeno sklepamo, da popravek v_i pripada verjetno grobo pogrešenemu opazovanju. Ker test temelji na predpostavki, da je v GMM-u le eno opazovanje obremenjeno z grobim pogreškom, je treba zelo previdno izločati domnevno grobo pogrešena opazovanja. Pomembno se je tudi zavedati, da pri testu τ neposredno uporabimo referenčno varianco a posteriori $\hat{\sigma}_0^2$, ki je ob prisotnosti grobega pogreška večja. To pomeni, da je vrednost testne statistike manjša, kar ima za posledico neodkrite manjše grobe pogreške.

3.3 Testiranje homogenosti natančnosti dveh terminskih izmer

Ocenjene vrednosti neznank po izravnavi (koordinate točk) so datumsko odvisne količine. Ker so premiki točk določeni kot razlika ocenjenih vrednosti koordinatnih neznank med obravnavanimi terminskimi izmerami, so tudi premiki in pripadajoča kovariančna matrika datumsko odvisne količine. V deformacijski analizi je najpomembnejša in hkrati najtežavnejša naloga določiti mirujoče točke med obravnavanimi terminskimi izmerami, ki jih lahko uporabimo kot referenčni koordinatni sistem za določitev premikov.

Navadno v praksi obravnavamo dve terminski izmeri, zato mora biti za testiranje homogenosti natančnosti izpolnjen pogoj, da sta referenčni varianci a posteriori, izračunani iz izravnave posamezne terminske izmere, statistično skladni. Natančneje ta pogoj zapišemo v obliki pričakovane vrednosti referenčne variance a posteriori za posamezno izmero. Zapišemo ničelno hipotezo:

$H_0 : E(\hat{\sigma}_{0i}^2) = E(\hat{\sigma}_{0j}^2)$; opazovanja dveh terminskih izmer i in j so statistično homogene natančnosti.

Hipotezo testiramo z zapisom testne statistike, ki se porazdeljuje po Fisherjevi porazdelitveni funkciji F:

$$T = \frac{\hat{\sigma}_{0i}^2}{\sigma_{0j}^2} \sim F(1-\alpha/2, f_i, f_j).$$

Če velja $T \leq F_{1-\alpha/2, f_i, f_j}$, ničelne hipoteze ne zavrnemo, saj test ne kaže na statistično neujemanje referenčnih varianc a posteriori med dvema izmerama. Ugotovimo lahko, da sta terminski izmeri homogene natančnosti, zato lahko izračunamo skupno a posteriori referenčno varianco obeh terminskih izmer:

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}_i^T \mathbf{P} \mathbf{v}_i + \mathbf{v}_j^T \mathbf{P} \mathbf{v}_j}{f_i + f_j} = \frac{f_i \hat{\sigma}_{0i}^2 + f_j \hat{\sigma}_{0j}^2}{f},$$ kjer je $f_i = n_i - u_i + d_i$ in $f_j = n_j - u_j + d_j$.

Če velja $T \leq F_{1-\alpha/2, f_i, f_j}$, zavrnemo ničelno hipotezo, kar kaže na neskladje referenčnih varianc a posteriori med terminskima izmerama. Ugotavljanje premikov točk na podlagi dveh terminskih izmer, ki ne izkazujeta homogenosti natančnosti opazovanj, lahko vodi k povsem napačnim sklepom glede stabilnosti obravnavanih geodetskih točk.

3.4 Testiranje značilnih premikov

Podlaga za ugotavljanje premikanja zgrajenega objekta ali naravnega dela zemeljskega površja je določitev spremembe položajev točk objekta. O premikih točk med dvema terminskima izmerama lahko sklepamo izključno za *identične točke*, izmerjene v dveh terminskih izmerah. Premike točk ugotavljamo s primerjavo izravnanih vrednosti koordinat točk v dveh terminskih izmerah. Predpostavimo, da obravnavamo koordinate točke P v ravnini v času t in $t + \Delta t$. Da bi lahko izračunali natančnost ocene premika točke, moramo poleg koordinat točke poznati tudi kovariančno matriko koordinat identičnih točk med terminskima izmerama:

$$\mathbf{\Sigma}_{P_tP_{t+\Delta t}} = \begin{bmatrix} \sigma^2_{y_t} & \sigma_{y_tx_t} & 0 & 0 \\ \sigma_{y_tx_t} & \sigma^2_{x_t} & 0 & 0 \\ 0 & 0 & \sigma^2_{y_{t+\Delta t}} & \sigma_{y_{t+\Delta t}x_{t+\Delta t}} \\ 0 & 0 & \sigma_{y_{t+\Delta t}x_{t+\Delta t}} & \sigma^2_{x_{t+\Delta t}} \end{bmatrix}.$$

Premik točke P v ravnini izračunamo po enačbi:

$$d = \sqrt{\Delta y^2 + \Delta x^2} = \sqrt{\left(y_{t+\Delta t} - y_t\right)^2 + \left(x_{t+\Delta t} - x_t\right)^2}.$$

Ob upoštevanju zakona o prenosu varianc in kovarianc zapišemo varianco premika $\sigma_d^2 = \mathbf{J}_d \Sigma_{P_tP_{\Delta t+t}} \mathbf{J}^T_d$, kjer je Jacobijeva matrika $\mathbf{J}_d$ enaka

$$\mathbf{J}_d = \begin{bmatrix} \frac{\partial d}{\partial y_t} & \frac{\partial d}{\partial x_t} & \frac{\partial d}{\partial y_{t+\Delta t}} & \frac{\partial d}{\partial x_{t+\Delta t}} \end{bmatrix} = \begin{bmatrix} -\frac{\Delta y}{d} & -\frac{\Delta x}{d} & \frac{\Delta y}{d} & \frac{\Delta x}{d} \end{bmatrix}.$$

Iz zapisanih zvez dobimo izraz za varianco premika točke P

$$\sigma_d^2 = \left(\frac{\Delta y}{d}\right)^2 \left(\sigma^2_{y_t} + \sigma^2_{y_{t+\Delta t}}\right) + 2\frac{\Delta y}{d}\frac{\Delta x}{d}\left(\sigma_{y_tx_t} + \sigma_{y_{t+\Delta t}x_{t+\Delta t}}\right) + \left(\frac{\Delta x}{d}\right)^2 \left(\sigma^2_{x_t} + \sigma^2_{x_{t+\Delta t}}\right),$$

ki ga uporabimo za testiranje značilnih premikov točk.

Zapišemo ničelno hipotezo:

$H_0 : d = 0$; točka miruje.

Hipotezo testiramo z zapisom testne statistike:

$$T = \frac{d}{\sigma_d},$$

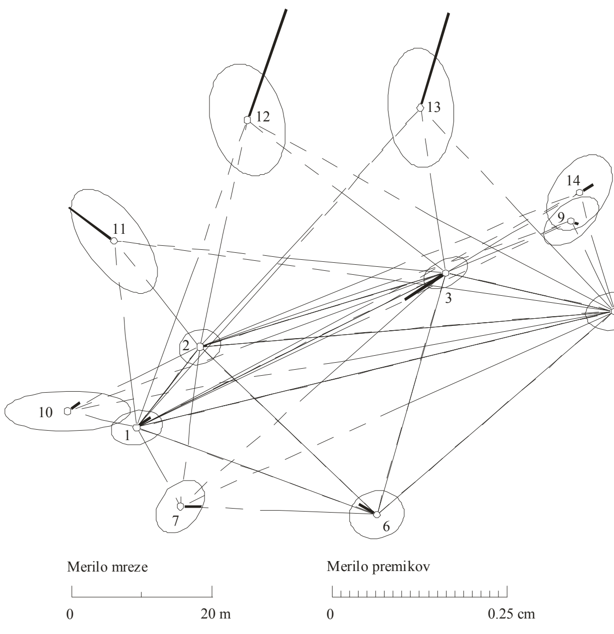
ki se ne porazdeljuje po nobeni od znanih porazdelitvenih funkcij. Porazdelitveno funkcijo za obravnavano testno statistiko zato določimo s simulacijami (Savšek-Safić et al., 2006). Osnovna ideja pri generiranju vzorca odvisnih normalno porazdeljenih slučajnih spremenljivk je, da najprej generiramo vzorec neodvisnih normalno porazdeljenih spremenljivk, potem pa uporabimo linearno transformacijo za pridobitev vzorca odvisnih slučajnih spremenljivk. Porazdelitvena funkcija testne statistike T je za vsako točko drugačne oblike, saj je standardna deviacija koordinat točk v posamezni terminski izmeri za različne točke različna. V n simulacijah nam postopek omogoča določitev empirične kumulativne verjetnostne porazdelitvene funkcije testne statistike za vsako posamezno točko, kar nam omogoča določitev kritične vrednosti T_α.

Testno statistiko T primerjamo glede na kritično vrednost T_α, ki jo pridobimo na podlagi empirične kumulativne porazdelitvene funkcije. Če velja $T \leq T_\alpha$, je tveganje za zavrnitev ničelne hipoteze preveliko. V tem primeru ugotovimo, da premik ni statistično značilen. Če velja $T > T_\alpha$, zavrnemo ničelno hipotezo, saj je tveganje za zavrnitev ničelne hipoteze manjše od izbrane stopnje značilnosti testa α. Tako potrdimo, da je obravnavani premik statistično značilen. Izbira stopnje značilnosti testa mora biti zelo premišljena, saj ima odločitev za posledico uvrstitev posamezne točke med mirujoče ali med točke, ki so se premaknile.

4 PRIMER TESTIRANJA PREMIKOV V MREŽI MONTSALVENS

V preteklosti je delovna skupina FIG uporabljala mrežo Montsalvens kot testni primer za primerjavo učinkovitosti različnih metod deformacijske analize (Delft, Fredericton, Hannover, Karlsruhe, München). Tudi mi jo uporabimo kot testni primer za ugotavljanje simuliranih premikov v mreži Montsalvens. Točke 10, 11, 12, 13 in 14 so na vrhu jezu in pomenijo kontrolne točke, za katere se ugotavlja značilnost premikov.

Opazovanja in premiki na točkah 3, 11, 12 in 13 so simulirani. Opazovanja so generirana kot neodvisna, normalno porazdeljena, s standardno deviacijo kotnih opazovanj $\sigma_\alpha = 0{,}3$ *mgon* = 1" in standardno deviacijo dolžinskih opazovanj $\sigma_s = 0{,}3$ *mm*. Za obe terminski izmeri velja identičen GMM. Obravnavana mreža vključuje n = 55 opazovanj in n = 29 neznank. Izvedeni sta dve terminski izmeri z identično vrsto in številom opazovanj, geodetski datum mreže je določen v okviru proste mreže. V postopku testiranja je postavljena ničelna hipoteza oblike $H_0 : d = 0$ in izbrana stopnja značilnosti testa α = 5 %. Opazovanja, približne koordinate in rezultati izravnave so navedeni v Caspary (2000). V obeh terminskih izmerah smo potrdili skladnost a priori in a posteriori referenčnih varianc. S testiranjem smo potrdili, da v opazovanjih ni prisotnih grobih pogreškov in da sta izmeri statistično homogene natančnosti.



Slika 2: Testna mreža Montsalvens

Slika 2 prikazuje opazovanja v mreži Montsalvens, izračunane vrednosti premikov in 95 % elipse zaupanja. Empirična kumulativna porazdelitvena funkcija je bila generirana s programom Premik

(Ambrožič et al., 2002). Empirična kumulativna porazdelitvena funkcija je določena za vsako točko posebej s 1.000.000 simulacijami. Na podlagi statističnega testiranja program Premik omogoča ugotavljanje statistično značilnih premikov posamezne točke. Izračunane vrednosti premikov primerjamo z znanimi vrednostmi.

Točka	Simulirani premiki [mm]	
	Δy	Δx
3	- 0,60	- 0,50
11	- 0,75	0,60
12	0,50	1,10
13	0,30	1,00

Preglednica 2: Simulirani premiki med dvema izmerama na točkah 3, 11, 12 in 13

Točka	dy [mm]	dx [mm]	d [mm]	σ_d [mm]	T	T_α	α_T (5 %)
1	0,2	0,2	0,2	0,1	1,8380	2,3566	16,75
2	0,1	0,0	0,1	0,1	0,9025	2,4418	66,46
3	**- 0,6**	**- 0,4**	**0,7**	**0,1**	**5,4660**	**2,2770**	**0,00**
4	0,1	0,1	0,2	0,1	1,1849	2,3562	46,81
6	- 0,3	0,1	0,3	0,1	2,1431	2,4472	10,16
7	0,3	0,0	0,3	0,1	2,3250	2,3878	5,71
9	0,1	0,0	0,1	0,1	0,9052	2,2849	61,46
10	0,2	0,1	0,2	0,3	0,7342	2,0720	64,40
11	**- 0,6**	**0,5**	**0,8**	**0,3**	**2,5340**	**2,2127**	**2,20**
12	**0,6**	**1,6**	**1,7**	**0,3**	**5,9921**	**2,3582**	**0,00**
13	**0,4**	**1,4**	**1,4**	**0,3**	**4,5304**	**2,2777**	**0,00**
14	0,2	0,1	0,2	0,2	1,0746	2,3590	52,81

Preglednica 3: Testiranje značilnosti premikov v mreži Montsalvens

Kot je razvidno iz preglednice 3, so premiki na točkah 3, 12 in 13 statistično značilni. Testna statistika je dvakrat večja od kritične vrednosti, $T > 4$. Dejansko tveganje α_T za zavrnitev ničelne hipoteze je minimalno. Predlagana testna statistika d/σ_d zazna tudi premik na točki 11, $T > T_\alpha$ = 2,2127 z dejanskim tveganjem za zavrnitev pravilne ničelne hipoteze α_T = 2,20 %, kar je manj od izbrane stopnje značilnosti testa α = 5 %. Dejanski premik domnevno stabilne točke 7 ni statistično značilen, saj velja $T < T_\alpha$ = 2,3878 z dejanskim tveganjem za zavrnitev pravilne ničelne hipoteze α_T = 5,71 %. Točke ne moremo obravnavati kot nestabilne. Dejansko tveganje za zavrnitev pravilne ničelne hipoteze na drugih, domnevno stabilnih točkah presega izbrano stopnjo značilnosti testa α = 5 %. Ker je dejansko tveganje za zavrnitev pravilne ničelne hipoteze na domnevno stabilnih točkah preveliko, zanje ni mogoče sklepati, da niso stabilne.

V preglednici 4 je prikazano, da kritične vrednosti za posamezne točke v mreži pri izbrani stopnji značilnosti testa niso enake. Zato je zelo pomembno, da pravilno določimo porazdelitveno funkcijo za testno statistiko za vsako točko v mreži. Ugotavljamo tudi, da s spremembo stopnje značilnosti testa ne vplivamo na vrednost premika in njegovo natančnost, kakor tudi ne na vrednost testne statistike in dejansko tveganje za zavrnitev pravilne ničelne hipoteze. Izbrana stopnja značilnosti testa α vpliva na izračun kritične vrednosti T_α, hkrati pa je neodvisna od dejanskega tveganja. Kolikor večje je dopustno tveganje za zavrnitev pravilne ničelne hipoteze α, toliko manjša je kritična vrednost T_α. To v praksi pomeni, da je pri večjem dopustnem tveganju za zavrnitev pravilne ničelne hipoteze hitreje izpolnjen pogoj $T > T_\alpha$, torej stabilno točko hitreje obravnavamo kot točko, ki se je značilno premaknila.

Točka	T	T_α (10 %)	T_α (5 %)	T_α (1 %)	T_α (0.1 %)
1	1,8380	2,0718	2,3566	2,9310	3,5542
2	0,9025	2,1444	2,4418	3,0386	3,6950
3	**5,4660**	**2,0056**	**2,2770**	**2,8518**	**3,4317**
4	1,1849	2,0713	2,3562	2,9337	3,5562
6	2,1431	2,1500	2,4472	3,0387	3,7105
7	2,3250	**2,0832**	2,3878	2,9597	3,6136
9	0,9052	1,9892	2,2849	2,8366	3,4253
10	0,7342	1,8132	2,0720	2,6549	3,3270
11	**2,5340**	**1,9384**	**2,2127**	2,8339	3,5184
12	**5,9921**	**2,0596**	**2,3582**	**2,9390**	**3,5748**
13	**4,5304**	**1,9987**	**2,2777**	**2,8691**	**3,4935**
14	1,0746	2,0459	2,3590	2,9156	3,5706

Preglednica 4: Odvisnost kritične vrednosti T_α od izbrane stopnje značilnosti testa α

Obratno za nestabilne točke velja, da je pri večjem dopustnem tveganju za zavrnitev pravilne ničelne hipoteze počasneje izpolnjen pogoj $T > T_\alpha$, torej je tveganje za neupravičeno uvrstitev nestabilne točke med stabilne najmanjše. Ker pri presoji o stabilnosti točk praviloma z ničelno hipotezo testiramo pogoj $d = 0$, torej ugotavljamo, ali se je točka statistično značilno premaknila, je izbira ustrezne stopnje značilnosti testa α zelo pomembna. Skladno s statistično teorijo lahko ugotovimo, da se stopnja zaupanja v rezultate oziroma verjetnost pravilne odločitve ustrezno manjša s povečevanjem dopustnega tveganja.

5 SKLEP

V deformacijski analizi imamo opravka z več opazovanji v več terminskih izmerah, zato je zanesljivo ocenjevanje opazovanj, neznank in premikov zelo pomembno. S statističnimi testi ocenjujemo parametre, ugotavljamo skladnost opazovanj in matematičnega modela, odkrivamo grobe pogreške v opazovanjih ter ugotavljamo skladnost domnevno stabilnih točk med terminskimi izmerami in določamo statistično značilne premike nestabilnih točk.

Statistično testiranje je zelo uporabno orodje, ki nam pomaga pri odločitvah. Na podlagi bogatih izkušenj je do podobnih odločitev mogoče priti tudi brez statističnega testiranja. Pomanjkanje izkušenj lahko le deloma nadomestimo z uporabo statističnih testov. V postopku testiranja hipotez si postavimo ničelno in alternativno hipotezo. V primeru zavrnitve ničelne hipoteze je treba z nadaljnjimi statističnimi testi ugotoviti vzrok za neujemanje med opazovanji in matematičnim modelom. Mogoči vzroki za napake so *funkcionalni model* (koordinatni sistem, instrument, gravitacijsko polje, model refrakcije idr.), *stohastični model* (a priori variance, korelacije idr.), *opazovanja* (grobi pogreški, napake pri identifikaciji točk, nestabilna stabilizacija, napake centriranja idr.), *izračuni* (napake pri vnosu podatkov, numerična nestabilnost, pogreški zaokroževanja idr.).

Na testni mreži Montsalvens, v kateri so na štirih točkah simulirani premiki v velikostnem redu 1 mm, smo praktično uporabili statistične teste pri presoji značilnih premikov. Ugotavljamo, da je mogoče s predlagano testno statistiko d/σ_d zelo uspešno odkriti vse domnevno nestabilne točke, in sicer s tveganjem, manjšim od izbrane stopnje značilnosti testa $\alpha = 5$ % (preglednica 3). Iz preglednice 4 lahko sklepamo glede korelacije med izbrano stopnjo značilnosti testa in izračunano kritično vrednostjo porazdelitvene funkcije. Kolikor večje je dopustno tveganje, toliko manjša je kritična vrednost porazdelitvene funkcije, hitreje neko stabilno točko obravnavamo kot nestabilno. Iz preglednice 4 je razvidno, da pri izbrani stopnji značilnosti testa $\alpha = 10$ % točke 7 ne moremo obravnavati kot stabilne. Iz preglednice 4 je hkrati razvidno, da pri izbrani stopnji značilnosti testa $\alpha = 1$ % in manj točke 11 ne moremo več obravnavati kot nestabilne. Upravičeno torej sklepamo, da najbolj verodostojne rezultate dobimo pri izbrani stopnji značilnosti testa $\alpha = 5$ %, kar je tudi v praksi najpogosteje izbrana vrednost.

Kakovost opazovanj in verodostojnost pridobljenih rezultatov izkazujemo z ocenjevanjem posameznih parametrov, pri čemer so nam v veliko pomoč postopki statističnega testiranja hipotez. Kolikor večja je verjetnost pravilnega rezultata, toliko manjše je za naročnika tveganje ob morebitni napačni odločitvi. Pomembno je, da tako naročniki kot izvajalci razumemo testiranje hipotez le kot uporabno orodje, ki nam pogosto olajša nekatere odločitve, nikakor pa ne more nadomestiti merskih in računskih postopkov v geodeziji.

Literatura in viri:

Ambrožič, T., Turk, G., Stopar, B. (2002). Premik – računalniški program za izračun in ugotavljanje statistično značilnih premikov v položajni geodetski mreži. Interna izdaja, Ljubljana.

Caspary, W. F. (2000). Concepts of Network and Deformation Analysis. School of Surveying. Kensington: University of New South Wales.

Grigillo, D., Stopar, B. (2003). Metode odkrivanja grobih pogreškov v geodetskih opazovanjih. Geodetski vestnik, 47(4), 387–403.

Harvey, B. R. (1991). Practical Least Squares and Statistics For Surveyors, University of New Outh Wales, School of Surveying, Kensington.

Kuang, S. (1969). Geodetic Network Analysis and Optimal Design: Concepts and Applications, Ann Arbor Press, Inc., Chelsea, Michigan.

Savšek-Safić, S. (2002). Optimalna metoda določanja stabilnih točk v deformacijski analizi. Doktorska disertacija. Ljubljana: FGG OGG.

Savšek-Safić, S., Ambrožič T., Stopar B., Turk G. (2006). Determinantion of Point Displacements in the Geodetic Network. Journal of Surveying Engineering, 132(2), 58–63.

Prispelo v objavo: 24. januar 2013
Sprejeto: 3. april 2013

doc. dr. Simona Savšek, univ. dipl. inž. geod.
UL, FGG – Oddelek za geodezijo, Jamova 2, SI-1000 Ljubljana
e-pošta: simona.savsek@fgg.uni-lj.si