

Inž. Vlastimil Konečný CSc, VÚHU :

POUŽITÍ METOD STATISTICKÉ INDUKCE KE ZPRACOVÁNÍ DAT FYZIKÁLNÍ ANALÝZY ZEMIN

Kvalita zemní hmoty projevuje se jako důležitý ekonomický činitel ve všech fázích výstavby a provozu velkých zemních děl, zejména povrchové těžby surovin. Jde hlavně o stabilitu pracovních řezů a bočních svahů, rozpojování zemin a s tím spojené konstrukce dobývacích strojů, únosnost podložky, dopravu hmot, zakládání výsypek atd. Ekonomický efekt této činnosti je podmíněn objektivností geotechnických podkladů, jejich průměrných hodnot a rozptylu od minimálních do maximálních hodnot.

Vzhledem ke spoustě materiálů ukazuje se vhodné jejich zpracování podle zákonů matematické statistiky. Rovněž předpoklad rozložení fyzikálních vlastností zemin podle normálního Gauss-Laplassova zákona zdál se velmi přijatelný.

Pod tímto dojmem byl proveden v oblasti Severočeského hnědouhelného revíru výběr asi tisíce neporušených vzorků terciérních jílu a učiněn pokus uvedené předpoklady aplikovat.

Již předem bylo jasné, že nebude možné použít všech získaných vzorků vzhledem k jejich zřejmé vzájemné odlišnosti. Jako první kritérium výběru bylo použito toho, že zemina podle křivky zrnitosti musí odpovídat jílu nebo jílovité zemině. Jako druhé kritérium byla vzata konzistence. Byly vybrány pouze ty vzorky, které na základě subjektivního srovnání odpovídají tuhé a pevné konzistenci podle ČSN, neboť těchto se vyskytovalo nejvíce.

Průzkumný materiál byl shromážděn asi z 8.500 odvrtných metrů. Průměrná hloubka vrtu byla cca 100 m. Jeden odběr vzorku připadal asi na 8 - 10 m.

Sledovány byly tyto fyzikální vlastnosti :

- 1/ Objemová váha γ [Mp/m^3]
- 2/ Měrná (specifická) váha γ_s [Mp/m^3]
- 3/ Vlhkost v % váhy suché substance W %
- 4/ Smyková soudržnost c [kp/cm^2]
- 5/ Součinitel tření φ [$\text{tg } \varphi$]

Ze statistických veličin pak :

- 1/ Aritmetický průměr
- 2/ Modus
- 3/ Medián
- 4/ Směrodatná odchylka
- 5/ Součinitel kososti

- 6/ Součinitel strmosti
- 7/ Nahrazení empirické čáry teoretickou
- 8/ Kontrola přesnosti nahrazení empirické čáry teoretickou, Pearsonovým a Bernštejnovým kritériem
- 9/ Posouzení homogenity souboru z významnosti rozdílů rozptylů mezi výběry a uvnitř výběrů - Z - test
- 10/ Posouzení homogenity souboru z významnosti rozdílu mezi průměry výběru - t - test.

Získané výsledky :

Nebudu uvádět analýzu všech výše uvedených pozorování, nýbrž jen těch, která v závěru budou předmětem diskuse.

Objemová váha

Aritmetický průměr : $1,98 \text{ Mp/m}^3$

Směrodatná odchylka : $0,129 \text{ Mp/m}^3$

Nahrazení empirické čáry četnosti teoretickou nevyhovuje ani podle Pearsonova testu. Rozdíly jsou větší než připouští test. Z posouzení homogenity souboru, to znamená v našem případě homogenity profilu terciérních jíílů do relativní cca 100 m hloubky, z významnosti rozdílů rozptylů mezi výběry a uvnitř výběrů - Z - testem, dostaneme tak velké rozdíly od požadovaných kritérií, že nelze soubor z hlediska objemové váhy považovat za homogenní.

Vezmeme-li však pro posouzení homogenity téhož souboru v úvahu významnost rozdílu mezi průměry výběrů - t - test, lze tento soubor považovat jako homogenní.

Měrná (specifická) váha

Aritmetický průměr : $2,68 \text{ Mp/m}^3$

Směrodatná odchylka : $0,064 \text{ Mp/m}^3$

Nahrazení empirické čáry četnosti teoretickou podle Bernštejnova kritéria nevyhovuje, avšak Pearsonův test vychází naopak příznivě.

Z posouzení homogenity souboru z významnosti rozdílů rozptylů mezi výběry a uvnitř výběrů (Z-test) vyplývá, že soubor není z hlediska specifických vah homogenní. Z posuzování homogenity souboru z významnosti rozdílů mezi průměry výběrů t-testem vychází naopak soubor homogenní.

Vlhkost v % váhy suché substance

Aritmetický průměr : 25,5 %

Směrodatná odchylka : 6,35 %

Nahrazení empirické čáry teoretickou nevyhovuje ani podle Bernštejnova kritéria, ani podle Pearsonova testu. Homogenita souboru již nebyla posuzována, neboť ze tvaru

histogramu je zřejmá značná difference.

Smyková soudržnost

Aritmetický průměr : $0,67 \text{ kp/cm}^2$

Směrodatná odchylka : $0,314 \text{ kp/cm}^2$

Nahrazení empirické čáry četnosti teoretickou nevyhovuje ani podle Bernštejnova kritéria ani podle Pearsonova testu. Z posouzení homogenity jak Z-testem tak t-testem vyplývá, že soubor nelze považovat z hlediska smykových soudržností za homogenní.

Součinitel tření

Aritmetický průměr : $\text{tg } \varphi = 0,35$

Směrodatná odchylka : $s = 0,165$

Z posouzení možnosti nahrazení empirické čáry teoretickou jak podle Bernštejnova kritéria tak i podle Pearsonova testu vyplývá neoprávněnost této možnosti.

Z podrobnější analýzy empirické čáry se však ukazuje, že hlavní váhu výchylky způsobuje pouze jeden interval, v němž je nahromaděno mimořádně velké množství pozorování (třídní znak $\text{tg } \varphi = 0,23$). Úpravou tohoto intervalu je nahrazení empirické čáry teoretickou podle Pearsonova testu vyhovující.

Posouzení homogenity t-testem prokazuje rozdíl mezi průměry výběrů za nevýznamný a lze tedy považovat soubor $\text{tg } \varphi$ za homogenní.

Shrneme-li zatím výsledky, ukazuje se, že u některých zkoušek nevyhovují všechna použitá kritéria statistické indukce (vlhkost, soudržnost), u některých část (objemová váha) a u dalších (specifická váha, součinitel tření) prakticky kritéria vyhovují.

Důvodů tohoto stavu může být několik :

- 1/ Nevhodná aplikace matematické statistiky na analýzu fyzikálních vlastností zemin v širším slova smyslu.
- 2/ Aplikace je sice vhodná, avšak použitá kritéria jsou v daných podmínkách příliš přísná.
- 3/ Aplikace je vhodná pouze v omezeném rozsahu nebo jen u některých zkoušek.

K objasnění těchto otázek bylo použito systematického průzkumu fyzikálních vlastností zemin v několika třisetmetrových příčných profilech terciárními sedimenty, z nichž jeden, graficky zpracovaný, je přiložen.

Zde můžeme pozorovat, že naměřené údaje jsou vždy rozloženy kolem jakési střednice v menším či větším rozptylu. U některých zkoušek tato střednice podléhá zákonu polohy, u některých ne. Tak např. bez vlivu hloubky uložení je specifická váha a úhel vnitřního tření a lze tedy jejich rozptyl považovat jako náhodný. Toto jsou

rovněž ty vlastnosti zemin, pro něž uvedená kritéria vyhovují.

U ostatních, tj. objemové váhy, vlhkosti a soudržnosti je proložená střednice odchýlena od svislice, tedy zřejmý vliv hloubky na změnu této vlastnosti. U těchto vlastností zemin také použitá kritéria statistické indukce nevyhovují. Výjimku tvoří jen posouzení homogenity souboru objemových vah t-testem, který zde vykazuje již homogenitu. Je to zřejmě tím, že přírůstek objemové váhy s hloubkou je mírnější (8 %) než vlhkosti (14 %) a soudržnosti cca (90 %) na 100 m vzhledem k výchozí hodnotě na 50 m hloubky.

Při sledovaném souboru pouze do 100 - 120 m hloubky je pak zřejmě benevolentní t-test schopen pojmout ještě jako homogenní menší rozdíly objemových vah.

Závěr

- 1/ Na základě této analýzy se domnívám, že principů matematické statistiky jsme při hromadných zpracováních fyzikálních vlastností zemin oprávněni používat, neboť pokud nejsou pod tendenčním vlivem jiného zákona, mají charakter náhodného rozptylu, jemuž vyhovuje zákon normálního Gauss-Laplassova rozložení.
- 2/ Pro účely geotechniky je dostatečné používat méně přísných kritérií, než jak je zvykem např. při kontrole strojní výroby. Pro nahrazení empirické čáry četnosti teoretickou doporučuji používat Pearsonova testu. K posouzení homogenity souboru doporučuji použít t-test, který pracuje s významností rozdílů mezi průměry.
- 3/ Podmínkou aplikace matematické statistiky při zpracování dat fyzikální analýzy zemin je, aby sledovaný soubor byl geologicky identický a stejné litologické facie.
- 4/ Za těchto podmínek pak soubor vzorků odebíraných z plochy lépe vyhovuje podmínkám náhodného rozptylu než vzorky z různých hloubek.
- 5/ I u geologicky identického útvaru nelze tento považovat za statistický soubor do libovolné hloubky, neboť řada zkoušek podléhá vlivu zákona polohy. Podle mého názoru je hranice hloubky 40 - 50 m maximální, kdy lze vliv hloubky ještě zanedbat.
- 6/ Nevyhoví-li sledovaný soubor ani kritériím uvedeným v bodě 2, pak je nutno zpravidla vždy příčinu hledat v účinku další zákonitosti, kterou jsme v řešení opomenuli, např. různé geologické útvary, různé litologické facie, příliš rozdílné hloubky apod.

Doplňky k použité matematicko-statistické terminologii :

- 1/ Aritmetický průměr je střední hodnota získaná jako n-tý podíl ze součtu u hodnot x , bez ohledu na jejich rozložení

$$M = \frac{1}{n} \sum_{i=1}^n x_i$$

- 2/ Směrodatná odchylka charakterizuje šíři neboli rozptyl intervalu znaku a je měřítkem jeho proměnnosti

$$\sigma = \pm \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$$

- 3/ Kolísání hodnot kolem znaku střední hodnoty je buď náhodné nebo způsobeno systematickými vlivy. Ty hromadné jevy, kde kolísání hodnot argumentu x kolem průměru M je zcela náhodné, tj. děje se z neznámých a nezjistitelných příčin, mají rozložení normální a lze je znázornit tzv. "normální křivkou četnosti" - zákonem Gauss-Laplaceovým

$$y = \frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{(x-M)^2}{2\sigma^2}}$$

y - počet prvků znaku x

N - rozsah souboru

M - aritmetický průměr

σ^2 - rozptyl

σ - směrodatná odchylka

e - přirozený logaritmus

π - Ludolfovo číslo

- 4/ Máme-li skutečný histogram, zajímá nás, do jaké míry je tento podobný teoretické normální křivce rozložení a zda tento histogram můžeme normální křivkou nahradit.

K posouzení, zda soubor je rozložen normálně, podle křivky Gauss-Laplaceovy, slouží různá kritéria a testy. Zde použito kritérium Bernštejnovo a Pearsonův test, jimiž sledujeme odchylky rozložení od ideálního a vyjadřujeme je přímo v procentech pravděpodobnosti výskytu.

Pearsonův test

$$\chi^2 = \sum \frac{(f_i - y_i)^2}{y_i}$$

- f_i - četnost třídy skutečná
 y_i - četnost třídy teoretická

Vypočtené hodnoty χ^2 srovnávají se pak s tabelovanými hodnotami pravděpodobnosti a nemají tuto zvolenou pravděpodobnost přestoupit.

Bernšteinovo kritérium je vlastně rozšířeným Pearsonovým testem a má tvar

$$\Delta = \frac{1}{l-1} \sum_i^l \frac{(f_i - y_i)^2}{y_i} < 1$$

- l - počet tříd
 f_i - četnost třídy skutečné
 y_i - četnost třídy teoretické

Udává přísnější hledisko než Pearsonův test.

- 5/ Z-test nebo t-test jsou kritéria analyzující statistický soubor z hlediska homogenity. To znamená, že posuzují, zda rozptyl uvnitř souboru je způsoben náhodně nebo systematicky.

Z-test posuzuje toto na základě analýzy rozptylu

$$Z = \frac{1}{2} \ln \frac{s_m^2}{s_n^2}$$

kde s_m^2 značí rozptyl mezi průměry výběrů a s_n^2 značí rozptyl uvnitř výběrů.

t-test posuzuje toto na základě významnosti rozdílů mezi průměry výběrů

$$t = \frac{m_1 - m_2}{\frac{(n_1 + n_2) (n_1 s_1^2 + n_2 s_2^2)}{n_1 n_2 (n_1 + n_2 - 2)}}$$

- $n_1; n_2$ - četnosti obou výběrů
 $m_1; m_2$ - aritmetické průměry těchto výběrů
 $s_1; s_2$ - směrodatné odchylky obou výběrů.

S h r n u t í

Na základě statistického zpracování výsledků rozborů několika set vzorků zemin z oblasti terciérních jíílů a geotechnického profilu napříč tímto útvarem dochází autor k názoru o možnostech a vhodnosti použití metod statistické indukce ke zpracování dat fyzikální analýzy zemin. Závěrem doporučuje kritéria a intervaly možností uplatnění statistické indukce v tomto oboru.

