

Decomposition Methods for Irregular Time Series

DMITS

UŽIVATELSKÝ MANUÁL

O B S A H

[O aplikaci DMITS](#)

[Uživatelské prostředí](#)

[Zadání časové řady](#)

[Zadání metody](#)

[Spuštění výpočtu](#)

[Výpis výsledků](#)

[Grafický výstup](#)

[Nastavení aplikace](#)

Tomáš Hanzák, 2005 - 2007

URL : www.sweb.cz/thanzak

e-mail : tomas.hanzak@post.cz

O aplikaci DMITS

Název DMITS je zkratkou pro *Decomposition Methods for Irregular Time Series* (Dekompoziční metody pro nepravidelné časové řady). Aplikace vznikla jako součást autorovy diplomové práce na Matematicko-fyzikální fakultě Univerzity Karlovy v Praze s názvem *Dekompoziční metody pro časové řady s nepravidelně pozorovanými hodnotami* (vedoucí Prof. RNDr. Tomáš Cipra, DrSc., rok odevzdání 2007).

Aplikace DMITS slouží výhradně k vyrovnávání a předpovídání v jednorozměrných nesezónních nepravidelných časových řadách pomocí metod typu exponenciálního vyrovnávání. Kromě metod popsaných v odborné literatuře nabízí aplikace DMITS několik metod navržených jejím autorem ve zmíněné diplomové práci. Pro detailní matematický popis všech zde dostupných metod a postupů odkazujeme tamtéž. Dodejme, že autorovi není znám jiný existující software pro vyrovnávání a předpovídání v nepravidelných časových řadách.

Seznam metod dostupných v aplikaci DMITS:

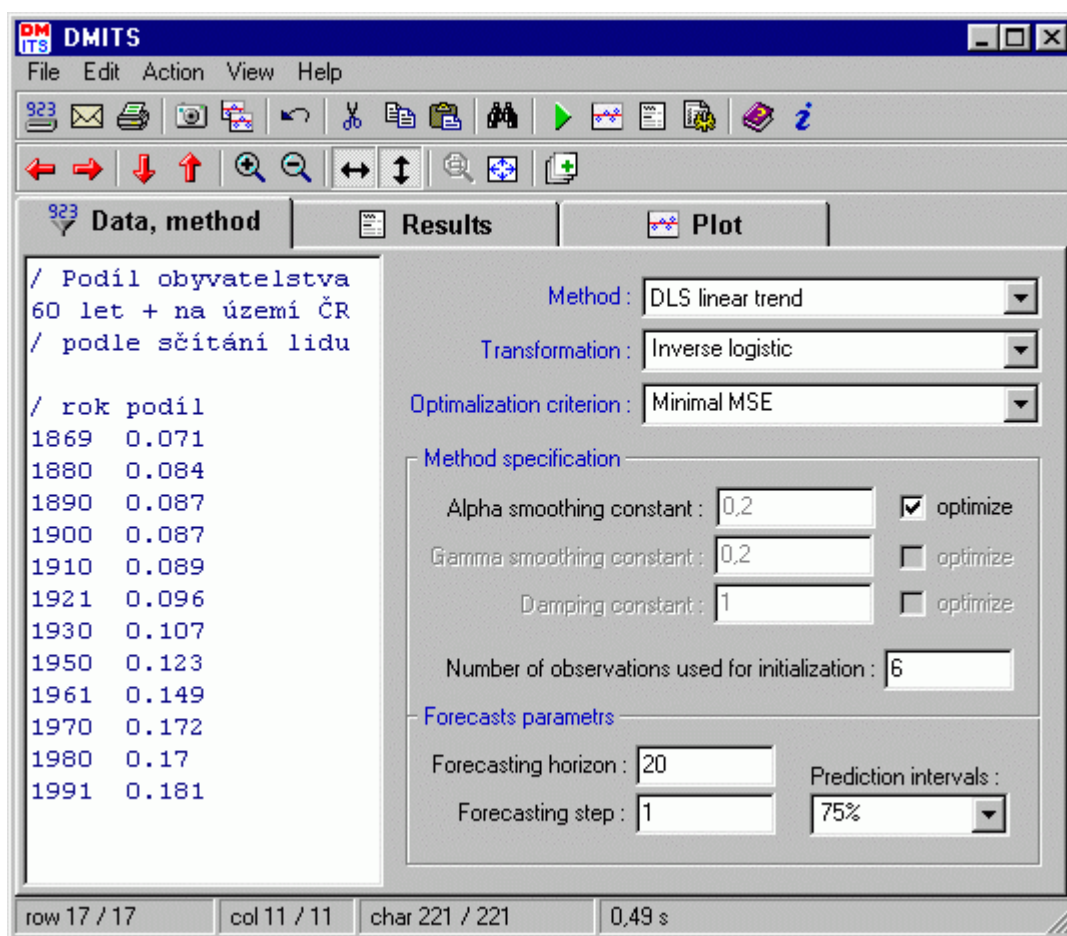
- Wrightova modifikace jednoduchého exponenciálního vyrovnávání
- Metoda založená na nepravidelně pozorovaném ARIMA(0, 1, 1) procesu
- Dvojitě exponenciální vyrovnávání
- DLS odhad lineárního trendu
- Wrightova modifikace Holtovy metody, včetně tlumeného trendu
- Trojitě exponenciální vyrovnávání
- DLS odhad kvadratického trendu

Stručný popis možností aplikace DMITS:

- použití některé ze čtyř transformací na zkoumanou časovou řadu
- automatický výpočet počátečních hodnot rekurentních metod
- manuální nebo optimální volba parametrů metody
- výpočet vyrovnaných hodnot, bodových předpovědí a mezí předpovědních intervalů se zadanou spolehlivostí
- vyhodnocení přesnosti prováděných předpovědí
- vyhodnocení adekvátnosti použití zvolené předpovědní metody
- prezentace výsledků v textové a grafické podobě a jejich snadný export nebo uložení

Uživatelské prostředí

Jazykem uživatelského prostředí aplikace DMITS je angličtina. Tvořeno je z velké části jediným **hlavním oknem** (obr. 1), kde se odehrává většina úkonů od zadání časové řady přes výběr a specifikaci použité metody až po výpis a vykreslení výsledků.



Obrázek 1: Hlavní okno aplikace DMITS

S hlavním oknem lze provádět všechny běžné úkony: přesouvat, měnit velikost, maximalizovat, minimalizovat atd. Aplikace se ukončí zavřením hlavního okna. Informace o poloze hlavního okna na obrazovce a jeho velikosti jsou uloženy a použity při příštím spuštění aplikace.

Na hlavním okně lze nalézt (shora dolů): **hlavní menu**, **panely tlačítek (Main Toolbar, Plot Toolbar)**, záložky pro přepínání mezi kartami **Data, method** (zadání časové řady a metody), **Results** (textový výpis výsledků) a **Plot** (grafický výstup) a stavový řádek (**Status Bar**).

Nachází-li se kurzor ve víceřádkovém textovém poli (pro zadání časové řady nebo



pro výpis výsledků), je na stavovém řádku vypisováno **číslo řádku, sloupce a znaku**, na kterém se kurzor aktuálně nachází. Za lomítkem je vypisován celkový počet řádků (sloupců, znaků). Další informace zobrazované na stavovém řádku budou popsány níže.

Víceřádková textová pole nabízejí (kliknutí pravým tlačítkem myši) kontextová menu obsahující příkazy z menu **Edit** v hlavním menu.



Na hlavním panelu tlačítek (**Main Toolbar**) se nacházejí tlačítka příkazů vesměs dostupných též z hlavního menu. Naopak tlačítka na grafickém panelu tlačítek (**Plot Toolbar**) poskytují funkce odjinud nedostupné.


Panely tlačítek nebo stavový řádek je možné zobrazovat / skrývat pomocí zaškrtnutých položek v menu **View**.

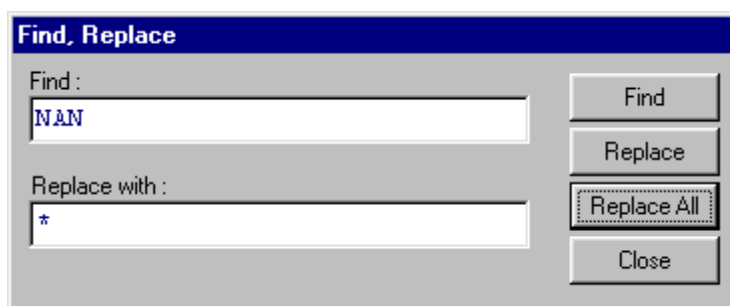
Kromě hlavního okna je součástí uživatelského prostředí ještě okno **Settings** sloužící k provádění změn v nastavení aplikace, okno **Find, Replace** pro vyhledávání a nahrazování textu a okno **About DMITS** s informacemi o aplikaci. Kromě toho aplikace využívá **systémových dialogů** pro otevírání a ukládání souborů, zadání barvy či fontu písma. Ve zvláštních oknech jsou zobrazována **chybová hlášení** a **potvrzující dotazy**.

Okno **About DMITS** se zobrazí stejnojmenným příkazem . Tento manuál je zobrazen příkazem  **User's manual**.

Zadání časové řady

Zkoumaná časová řada se zadává do textového pole v levé části karty  **Data, Method**. Je možné ji zde zadat ručně, vložit ze schránky či načíst ze souboru. K tomu slouží příkaz  **Load Data**, který zobrazí dialog pro zadání umístění a názvu souboru s časovou řadou. Načíst je takto možné vždy celý obsah libovolného souboru s prostým textem, nezávisle na jeho příponě.

Ve zmíněném textovém poli je možné provádět všechny běžné úpravy. Lze v něm také vyhledávat text, případně nahrazovat jeden textový řetězec jiným. K tomu slouží okno **Find, Replace** (viz. obr. 2), které se zobrazí stejnojmenným příkazem .




Obrázek 2: Okno *Find, Replace*

Zápis časové řady se řídí následujícími pravidly:

- Každé pozorování je zapsáno na jednom řádku.
- Na řádku jsou obecně čas a hodnota pozorování, oddělené tabulátorem.
- Tabulátory na začátku a na konci řádků a všechny mezery jsou ignorovány.
- Prázdné řádky a řádky začínající lomítkem (" / ") jsou ignorovány.
- Začíná-li řádek středníkem (" ; "), načítání časové řady je před ním ukončeno.
- Začíná-li řádek hvězdičkou (" * "), je považován za chybějící pozorování s časem o 1 větším než je čas předchozího pozorování.
- Je-li na řádku jen jedno číslo, je bráno jako hodnota pozorování. Čas pozorování je v tomto případě brán jako čas předchozího pozorování zvětšený o 1.
- Není-li uveden čas prvního pozorování (viz. předchozí bod), je brán roven 1.
- Používaný znak desetinného oddělovače musí odpovídat příslušnému nastavení aplikace DMITS (viz. odstavec *Nastavení aplikace*).

Výsledkem zadání časové řady je posloupnost časů pozorování t_1, t_2, \dots, t_n a posloupnost příslušných hodnot pozorování $y_{t_1}, y_{t_2}, \dots, y_{t_n}$. Počet pozorování n musí být roven alespoň 2. Platit musí $t_1 < t_2 < \dots < t_n$.

Zadání metody

Volba vyrovnávací a předpovědní metody, její specifikace a volba parametrů předpovědí budoucích hodnot zkoumané časové řady se provádí v pravé části karty  **Data**, **Method**, viz. obr. 1. Popíšme shora dolů jednotlivé volby.

V poli **Method** se volí jedna ze 7 vyrovnávacích a předpovědních metod. První dvě metody pracují s lokálně konstantním trendem, další tři s lokálně lineárním trendem a poslední dvě s lokálně kvadratickým trendem. Pro detailní matematický popis všech 7 metod odkazujeme na autorovu diplomovou práci.

V poli **Transformation** je možné zvolit jednu ze 4 dostupných transformací, která bude aplikována na zadanou časovou řadu ještě před použitím samotné vyrovnávací a předpovědní metody. Je-li zvolena *logaritmická* (**Logarithmic**) nebo *odmocninová* (**Square root**) transformace, musí být všechna pozorování zadané časové řady kladná. Je-li zvolena *inverzní logistická* (**Inverse logistic**) nebo *inverzní Gompertzova* (**Inverse Gompertz**) transformace, musí náležet všechna pozorování zadané časové řady do intervalu $(0, 1)$. Pokud uživatel zvolí možnost **None**, nebude použita žádná

z nabízených transformací.

V poli **Optimization criterion** je třeba vybrat jedno ze 4 kritérií pro optimální výběr hodnot parametrů metody. Je-li zvolena jako kritérium *minimální střední absolutní procentuální chyba* (**minimal MAPE**), musí být všechna pozorování zadané časové řady kladná. Pokud není u žádného parametru nastavena optimální volba jeho hodnoty (viz. níže), není toto pole aktivní a výběr kritéria se neprovádí.

Do pole **Alpha smoothing constant value** se zadává konkrétní hodnota vyrovnávací konstanty $\alpha \in (0, 1)$. Pokud je zaškrtnuta volba **optimize** napravo odtud, není toto pole aktivní a konkrétní hodnota se nezadává. Hodnota je v tom případě volena optimálně vzhledem k danému kritériu (viz. výše). Tento parametr používá všech 7 metod.

Podobně do pole **Gamma smoothing constant value** se v případě potřeby zadává konkrétní hodnota vyrovnávací konstanty $\gamma \in (0, 1)$. Tento parametr používá pouze Holtova metoda.

Podobně do pole **Damping constant value** se v případě potřeby zadává konkrétní hodnota tlumící konstanty $\varphi \in (0, 1]$. Tento parametr používá pouze Holtova metoda.

Do pole **Number of observations used for initialization** je třeba zadat počet pozorování ze začátku časové řady, ze kterých budou počítány počáteční hodnoty metody. U metod s konstantním, lineárním, resp. kvadratickým trendem to musí být alespoň 1, 2, resp. 3 pozorování. Nejvýše je možné inicializaci provést ze všech n pozorování.

Do pole **Forecasting horizon** je třeba zadat časový horizont určující, jak daleko do budoucnosti od času t_n posledního pozorování budou počítány předpovědi. Musí jít o nezáporné reálné číslo.



Do pole **Forecasting step** je třeba zadat časový krok pro výpis bodových a intervalových předpovědí budoucích neznámých hodnot zkoumané časové řady. Musí jít o kladné reálné číslo.


Nakonec v poli **Prediction intervals** je třeba zvolit spolehlivost předpovědních intervalů. Nabízeny jsou možnosti 50, 75, 90, 95 a 99 %. Pokud uživatel zvolí možnost **None**, nebudou předpovědní intervaly počítány.

Informace o zvolené metodě a její specifikaci jsou při ukončení aplikace uloženy a použity při jejím opětovném spuštění.

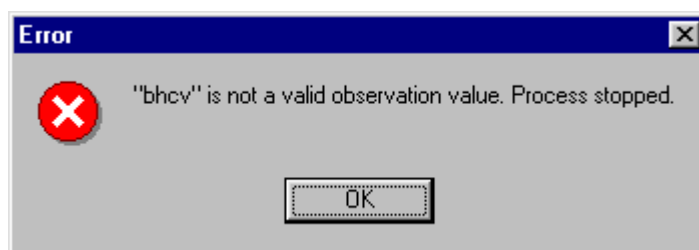
Spuštění výpočtu

Je-li zadána časová řada a zvolena vyrovnávací a předpovědní metoda včetně její specifikace, je možné dát aplikaci pokyn ke spuštění výpočtu příkazem

► **Run Computation** (F9). Pokud jsou všechny vstupy v pořádku, dojde k načtení zadané časové řady, načtení informací o výběru a specifikaci metody, načtení parametrů předpovědí budoucích hodnot řady, případnému výpočtu optimálních hodnot parametrů, vyčíslení všech vyrovnaných hodnot a předpovědí, vypsání výsledků do textového pole na kartě  **Results** a vykreslení výsledků na kartě  **Plot**.

V průběhu výpočtu má kurzor obvyklou podobu přesýpacích hodin . Po jeho skončení je **doba výpočtu** v sekundách vypsána na pravé části stavového řádku. Tato doba závisí na několika základních věcech. Jednak je to délka zadané časové řady, na které je doba výpočtu víceméně přímo úměrná. Dále hraje roli to, jestli je použita optimální volba parametrů metody. Platí, že čím více parametrů je voleno optimálně, tím déle bude výpočet trvat. Nakonec dobu výpočtu zvyšuje použití složitější metody. U dlouhých časových řad (tisíce pozorování) je zásadní ještě to, zda-li mají být vypisovány historické vyrovnané hodnoty a předpovědi (viz. odstavce *Výpis výsledků* a *Nastavení aplikace*). Většinou je výpočetní doba v řádu desetin sekundy, ve výjimečných případech v řádu desítek sekund.

Pokud je některá část vstupu neplatná nebo v rozporu s jinou částí vstupu, je o tom uživatel informován typickým "chybovým zvukem" a chybovým hlášením jako na obrázku 3. Výpočet je v tomto případě zastaven.







Obrázek 3: Příklad chybové hlášky



Z textu konkrétní chybové hlášky by mělo být patrné, v čem chyba spočívá. Možné příčiny chyb jsou tyto:


- Byl zadán neplatný čas pozorování nebo jeho hodnota.
- Posloupnost časů pozorování t_1, t_2, \dots, t_n není rostoucí.
- Počet pozorování je menší než 2.
- Pozorování neleží v intervalu, který je požadován zvolenou transformací.
- Pozorování nejsou všechna kladná a je zvolena MAPE jako kritérium optimální volby parametrů.
- Byla zadána neplatná hodnota parametru metody (jen při manuálním zadání).
- Byl zadán neplatný počet pozorování pro inicializaci.
- Byl zadán neplatný předpovědní horizont nebo krok.

Pokud dojde v průběhu výpočtu k neočekávané chybě, jejíž příčina nespočívá v ničem z výše uvedeného, nastane chybové hlášení **Unspecified error ocured**. Takovéto chyby by se v ideálním případě neměly vyskytovat vůbec.

Výpis výsledků

Je-li vstup v pořádku, dojde po zadání příkazu  **Run Computation** (F9) k vypsání výsledků do textového pole na kartě  **Results**. Pokud jsou poté změněny parametry předpovědí budoucích hodnot časové řady či změněna některá nastavení ovlivňující podobu dotyčného textového výstupu, je možné příkazem  **Write out Results** nechat tyto textové výsledky znovu vypsát. Při tom se projeví provedené změny ve výše uvedených nastaveních. Pokud od spuštění aplikace DMITS nebyl ještě spuštěn žádný výpočet, nemá příkaz  **Write out Results** žádný efekt.

Vypsané výsledky je možné uložit do souboru typu RTF nebo TXT příkazem  **Save Results**. Také je možné je vytisknout příkazem  **Print Results**.

Na kartě  **Results** bude postupně vypsáno:

- Zmínka o tom, že jde o výsledky aplikace DMITS, a aktuální datum a čas.
- Specifikace metody (**Method specification**).
- Vlastnosti zadané časové řady (**Data characteristics**).
- Míry přesnosti předpovědí (**Forecasting accuracy**).
- Míry efektivity předpovědní metody (**Forecasting efficiency**).
- Vyhodnocení adekvátnosti předpovědní metody (**Method adequacy**).
- Kontrola normality předpovědních chyb (**Forecasting errors normality check**).
- Empirické vlastnosti předpovědních intervalů (**Prediction intervals performance**).
- Historické vyrovnané hodnoty a předpovědi (**Historical forecasts and smoothed values**).
- Bodové a intervalové předpovědi budoucích hodnot (**Future forecasts and prediction intervals**).





Poznámky:

- V sekci **Method specification** se lze dozvědět konkrétní použité hodnoty parametrů metody. Pokud byla hodnota parametru zadána manuálně, je vedle ní v závorce slovo **fixed**. Pokud byla hodnota volena optimálně, je vedle ní v závorce

slovo **optimized**.

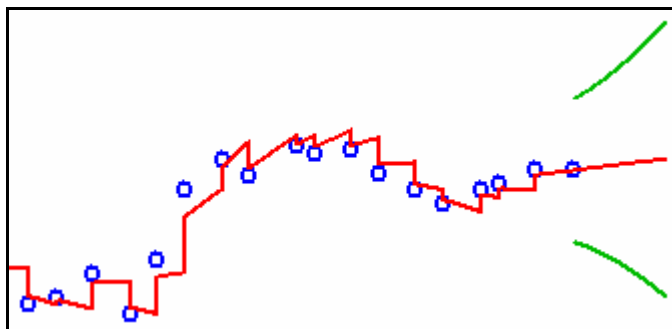
- Informace o zvoleném optimalizačním kritériu (**Optimized criterion**) je uváděna pouze tehdy, když byla prováděna optimální volba alespoň jednoho parametru.
- Informace o vyrovnávací konstantě γ a tlumící konstantě (**Damping constant**) je uváděna jen pro Holtovu metodu.
- **Average time spacing** je průměrná časová vzdálenost mezi dvěma sousedními pozorováními zadané časové řady, určená jako $(t_n - t_1)/(n - 1)$.
- Ukazatel **MAPE** v sekci **Forecasting accuracy** není počítán, pokud nejsou všechna pozorování zkoumané časové řady kladná.
- Vedle p -hodnot statistických testů jsou pro zvýraznění uváděny buď žádný, jeden, dva nebo tři vykřičníky (" ! "). To v případě, že tato p -hodnota je v intervalu $[0,1; 1]$, $(0,1; 0,05]$, $(0,05; 0,01]$ nebo $(0; 0,01)$.
- Ukazatele **R-Squared** v sekci **Forecasting efficiency** jsou analogií koeficientu determinace v lineární regresi. Pro podrobnosti odkazujeme do již zmíněné diplomové práce.
- Testy normality v sekci **Forecasting errors normality check** jsou prováděny pouze v případě, že počet pozorování n je nejméně 9.
- V sekci **Prediction intervals performance** je porovnání skutečných a teoretických procentuálních zastoupení vzájemné polohy jednokrokových předpovědních intervalů a příslušných pozorování.

Grafický výstup

Je-li vstup v pořádku, dojde po zadání příkazu  **Run Computation** (F9) na kartě  **Plot** k vykreslení časové řady spolu s vypočtenými předpověďmi. Pokud uživatel změnil některé parametry či nastavení, které ovlivňují podobu grafu, může ho příkazem  **Plot Data** kdykoli znovu vykreslit s použitím nových nastavení. Pokud od spuštění aplikace DMITS nebyl ještě spuštěn žádný výpočet, nemá příkaz  **Plot Data** žádný efekt.

Do jednoho grafu jsou vykresleny **značky pozorování** $y_{t_1}, y_{t_2}, \dots, y_{t_n}$, souvislá **křivka bodových předpovědí** pro časové rozmezí t_1 až $t_n + h$ (h je délka předpovědního horizontu) a souvislé **křivky mezi předpovědními intervaly** pro časové rozmezí t_n až $t_n + h$. Křivka bodových předpovědí je definována hodnotou $\hat{y}_{t_j+\tau}(t_j)$ (bodová předpověď pozorování $y_{t_j+\tau}$ z času t_j) v bodě $t_j + \tau$, kde $\tau \in [0, t_{j+1} - t_j]$. Dvě různé hodnoty $y_{t_j}(t_{j-1})$ a $\hat{y}_{t_j} = \hat{y}_{t_j}(t_j)$ (vyrovnaná hodnota v čase t_j) pro bod t_j ,

$j = 2, 3, \dots, n$, jsou v grafu spojeny vertikální úsečkou, viz. obr. 4.



Obrázek 4: Ukázka grafu časové řady a předpovědí



Příkazem **Save Plot** lze (celý) graf uložit do souboru typu BMP nebo JPG. Příkazem **Copy Plot to Clipboard** lze graf zkopírovat do schránky operačního systému. Pro uložení takto vyexportovaných obrázků doporučujeme formát GIF. Kurzorem myši (kliknutím levým tlačítkem a následným tažením) lze vybrat jen část grafu. Příkaz **Copy Plot to Clipboard** poté zkopíruje do schránky jen tuto vybranou část grafu.

Kliknutím pravým tlačítkem myši na obrázek se zobrazí kontextové menu obsahující tři výše zmíněné grafické příkazy. Je-li zobrazena karta **Plot**, zobrazuje se na stavovém řádku velikost grafu v pixelech a aktuální pozice kurzoru v grafu v časových jednotkách resp. v jednotkách pozorování zkoumané časové řady.


Velikost obrázku je možné měnit kliknutím na některý ze tří tmavomodrých čtverečků na jeho spodním a pravém okraji a následným tažením myši. Obsah obrázku je automaticky překreslen. V závislosti na nastavení aplikace (viz. následující odstavec) se při změně velikosti okna velikost obrázku automaticky přizpůsobí velikosti okna.

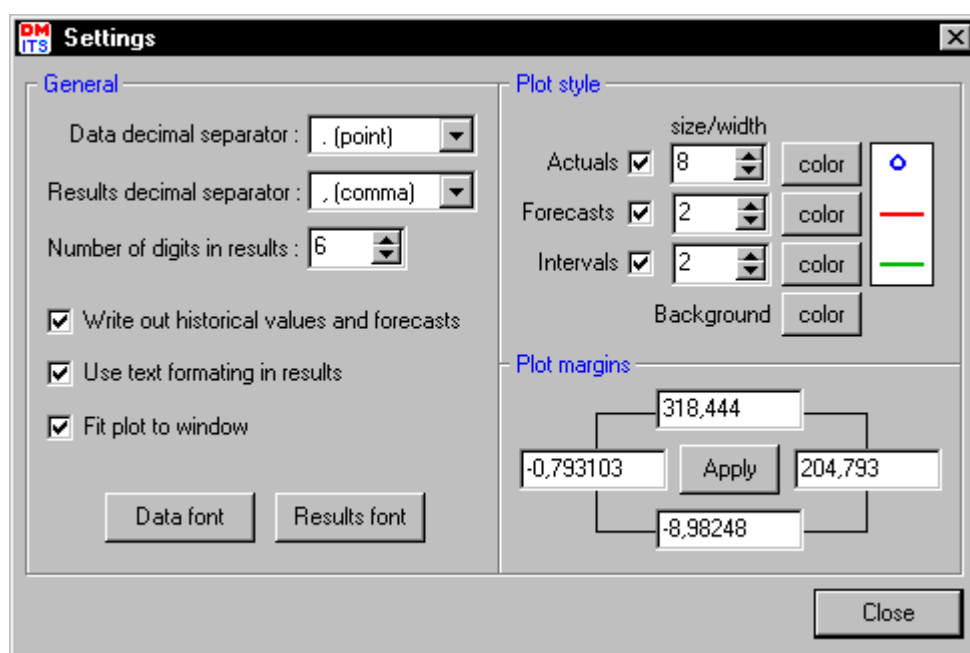
Na grafickém panelu tlačítek (**Plot Toolbar**) úplně vpravo se nachází „zaškrťovací“ tlačítko **Combine Plots**. Není-li stisknuto (výchozí stav), je vždy při novém vykreslení grafu smazán graf původní a pozice okrajů obrázku se přizpůsobí vykreslovaným datům. Je-li toto tlačítko stisknuto, není původní graf smazán a pozice okrajů obrázku je také zachována. To umožňuje kombinovat do jednoho obrázku více grafů, např. předpovědi ve stejné časové řadě, ale použitím různých metod. Při změně velikosti obrázku či jiné operaci vyžadující jeho znovuvykreslení je takto vykreslena pouze naposledy kreslená časová řada.

Zbýlá tlačítka na grafickém panelu tlačítek (**Plot Toolbar**) slouží k manipulaci s grafem ve smyslu změny pozice jeho okrajů. Šipky (**Move Plot Left**, **Move Plot Right**, **Move Plot Down** a **Move Plot Up**) slouží k „posunu“ obrázku požadovaným směrem. Tlačítka se symbolem lupy a (**Zoom In** a **Zoom Out**) slouží k přiblížení resp. oddálení obrázku. „Zaškrťovací“ tlačítka a (**Zoom in Horizontal Direction** a **Zoom in Vertical Direction**) určují, ve kterých směrech bude

přibližování / oddalování probíhat. Tlačítko  **Zoom to Selection** umožňuje přiblížit obrázek na úroveň vybrané oblasti. Tlačítko  **Adjust Plot Margins to Data** umožňuje přizpůsobit pozice okrajů obrázku naposledy vykreslené časové řadě.

Nastavení aplikace

Kromě již řečeného v předchozích odstavcích lze v aplikaci DMITS upravovat některá další nastavení. Ta jsou dostupná v okně **Settings** (viz. obr. 5), které se zobrazí stejnojmenným příkazem . Popíšme jednotlivá nastavení:



Obrázek 5: Okno **Settings**

Data decimal separator ... znak desetinného oddělovače používaný v zadání časové řady

Results decimal separator ... znak desetinného oddělovače používaný při výpisu výsledků

Number of digits in results ... počet číslic, který bude používán při výpisu číselných hodnot ve výsledcích. Netýká se hodnot parametrů, p -hodnot, ukazatelů R^2 a některých dalších číselných hodnot, která mají své počty cifer nezávislé na tomto nastavení.

Write out all historical values and forecasts ... určuje, zda mají být do textového výstupu vypsány hodnoty všech pozorování, vyrovnaných hodnot a bodových předpovědí za celou historii zadané časové řady.

Use text formatting in results ... určuje, zda má být při výpisu výsledků používáno

formátování textu (různé velikosti písma a nadpisy tučně).

Fit plot to window ... určuje, zda se při změně velikosti hlavního okna má automaticky přizpůsobit velikost grafu tak, aby vyplňoval celou disponibilní plochu okna.

V sekci **Plot style** (styl grafu) lze nastavit styl jednotlivých součástí grafu případně vypnout jejich vykreslování. Pro každou z položek **Actuals** (zadaná pozorování), **Forecasts** (bodové předpovědi a vyrovnané hodnoty) a **Intervals** (meze předpovědních intervalů) lze zaškrtnutím příslušného políčka zapnout jejich vykreslování do grafu. Dále lze nastavit parametr **size/width** určující velikost resp. tloušťku dané součásti grafu. Stisknutím tlačítka **Color** lze zadat barvu, s jakou bude položka vykreslována. Kromě toho lze ještě nastavit barvu pozadí grafu (**Background**). K dispozici je náhled na zvolený styl grafu.

V sekci **Plot margins** (okraje grafu) lze ručně zadat pozici okrajů grafu. Zadává se pozice levého a pravého resp. horního a dolního okraje grafu v časových jednotkách resp. v jednotkách pozorování zadané časové řady. Aby měla změna tohoto nastavení efekt, musí se stisknout tlačítko **Apply**.

Okno **Settings** se uzavře stisknutím tlačítka **Close**. Všechny změny nastavení (kromě pozic okrajů grafu, viz. výše) jsou tímto automaticky uloženy.