

Databázový systém pro projekt přehledu vegetace České republiky

Database system for the project of the vegetation survey of the Czech Republic

Milan Chytrý

Katedra systematické botaniky a geobotaniky Přírodovědecké fakulty Masarykovy univerzity, Kotlářská 2, 611 37 Brno

Od roku 1989 je českými fytocenology pro editaci fytocenologických tabulek a uchovávání vegetačních dat široce využíván program FYT, resp. EDIFYT (E. Brabec, Botanický ústav AV ČR Průhonice, srovnej Kučera 1992). Tento program umožňuje uživatelsky velmi jednoduchým a přátelským způsobem provádět všechny základní operace jako zadávání dat včetně dat hlavičkových, úpravu tabulek manuálním i automatickým přeskupováním řádků a sloupců, práci se synoptickými tabulkami, přípravu výstupních souborů pro tisk nebo pro další zpracování v textovém editoru a řadu dalších velmi užitečných věcí. Výhodou tohoto programu je i hardwarová nenáročnost.

FYT je neocenitelným a naprosto vyhovujícím pomocníkem pro editaci tabulek a práci na menších fytocenologických projektech zahrnujících do cca 200–350 (–400) fytocenologických snímků (podle druhové bohatosti a heterogenity vegetace). Lze jej však jen velmi obtížně použít pro práci s velkými soubory, která začíná být nezbytným předpokladem vytvoření skutečně daty podloženého přehledu vegetace České republiky a jeho začlenění do mezinárodního projektu European Vegetation Survey. Pro tento projekt, který je zatím jen neformálně organizován zvláštní pracovní skupinou pod záštitou Mezinárodní společnosti pro studium vegetace (IAVS), byl vyvinut standardní databázový systém TURBO(VEG) (Hennekens 1995), který je v současné době distribuován do různých evropských zemí na ta botanická pracoviště, která budou spolupracovat na projektu European Vegetation Survey, resp. na projektech národních vegetačních přehledů. V únoru až březnu 1996 bylo na Univerzitě v Lancasteru (Velká Británie) zorganizováno dr. Johnem Rodwellem v rámci projektu Darwin Initiative rozsáhlé školení na tento databázový systém pro pracovníky z 5 evropských zemí (Albánie, Česká republika, Lotyšsko, Rusko a Slovensko). Z ČR se ho zúčastnili fytocenologové z Masarykovy univerzity v Brně, Botanického ústavu AV ČR v Průhonicích a Palackého univerzity v Olomouci. Na tato pracoviště byl systém TURBO(VEG) také instalován.

Popis databázového systému TURBO(VEG)

Tento popis se vztahuje na verzi distribuovanou v březnu 1996, se zohledněním vlastností výrazně vylepšené verze předběžně demonstrované na 39. symposiu IAVS v září 1996. Upozorňuji, že vzhledem k rychlému vývoji systému a prudkému rozvoji jeho využití v celoevropském měřítku se níže uvedené informace mohou měnit.

TURBO(VEG) je systém zahrnující více samostatných programů, z nichž pro uživatele jsou nejvýznamnější TURBOSHELL, vlastní TURBOVEG, MEGATAB a TWINSpan. Další programy jsou důležité pro funkci systému, vývoj však směřuje ke zjednodušení v tom, že uživatel s nimi nebude muset přímo komunikovat. Proto zde nebudou popisovány a případný zájemce nalezne jejich popis v manuálu (Hennekens 1995). Každý z těchto programů produkuje výstupní soubory, které lze použít jako vstup do programů ostatních nebo jako vstup do spreadsheetových programů, textových editorů a některých programů pro mnohorozměrnou statistickou analýzu, zejména těch, které pracují s daty v cornellském kondenzovaném formátu (např. CANOCO, ter Braak 1987), ale např. také pro SYNTAX-5 (Podani 1994) a MULVA (Wildi et Orlóci 1996).

TURBOSHELL zahrnuje všechny ostatní programy a umožňuje jejich aktivaci.

TURBOVEG je základní databáze. Lze do ní zadávat fytoocenologické snímky zapsané v několika různých standardních stupnicích pokryvnosti, buď přímo z klávesnice nebo automaticky ze souborů v cornellském kondenzovaném formátu. Rovněž umožňuje zadávání hlavičkových dat, včetně informací o lokalitě nebo bibliografických údajů v případě snímků převzatých z literatury. Výběr snímků z databáze se provádí na základě uživatelem definovaných údajů z hlavičky nebo podle přítomnosti vybraného druhu či kombinace několika druhů. Před analýzou lze srovnat všechny vybrané snímky pomocí koeficientu podobnosti a zjistit tak, jestli v databázi není jeden a tentýž snímek přítomen více než jedenkrát. Snímky vybrané z databáze mohou být exportovány do různých formátů (ASCII, spreadsheet, cornellský kondenzovaný formát aj.).

Program je strukturován jako relační databáze, tzn. pracuje s několika standardními dílčími databázemi, které jsou s vlastní databází fytoocenologických snímků propojeny číselnými kódy. Jedná se o tyto dílčí databáze: seznam druhů, bibliografie, stupnice pokryvnosti, seznam států, seznam autorů, seznam projektů, seznam syntonů a orientace svahů (N, NNE, ... NNW).

Nejdůležitější databází je seznam druhů. V současné verzi TURBOVEGu, upravené pro Českou republiku, je použit seznam druhů cévnatých rostlin Rakouska, České republiky a Slovenska, který připravili dr. Walter Gutermann a prof. Harald Niklfeld z Vídeňské univerzity a který bude základem publikace 3. vydání "Liste der Gefäßpflanzen Mitteleuropas" (2. vydání viz Ehrendorfer 1973). Výhodou tohoto seznamu je, že umožňuje použití různých taxonomických úrovní, např. agregátů pro skupiny druhů, druhů *sensu lato*, sekcí u rodu *Taraxacum* apod. Přestože je seznam vesměs založen na poněkud širší taxonomické koncepci rodu a druhu, než jaká je přijata v Dostálově Nové květeně ČSSR (Dostál 1989) nebo v Květeně České republiky (Hejný et Slavík 1988 a následující díly), maximálně respektuje podrobné taxonomické detaily střeoevropské flóry a umožňuje tak důsledně definovat taxony zaznamenané ve fytoocenologických snímcích z našeho území (což by nebylo patrně vždy možné při použití seznamu druhů z díla Flora Europaea – Tutin et al. 1964–1980). Výhodné rovněž je, že tento seznam bude použit i ve verzích TURBOVEGu pro Rakousko a Slovensko, čímž bude zajištěna bezproblémová výměna dat mezi pracovníky z těchto tří států.

Nomenklatura mechorostů je sjednocena podle příručky Frey et al. (1995) a nomenklatura lišejníků podle seznamu Pišút et al. (1993), doplněná o rakouské druhy podle práce Türk et Poelt (1995).

Další specifickou databází, která byla do české verze TURBOVEGu začleněna, je seznam syntaxonů na úrovni tříd až asociací, převzatý z publikace Moravec et al. (1995). Bibliografie fytocenologické literatury ani seznam autorů zatím vytvořeny nebyly a předpokládá se, že budou postupně doplňovány jednotlivými uživateli.

TWINSpan je program pro divizivní klasifikaci (Hill 1979). Verze TWINSpanu dodávaná s balíkem TURBO(VEG) umožňuje naráz analyzovat až 5000 fytocenologických snímků.

MEGATAB je editor fytocenologických tabulek importovaných z programu TURBOVEG nebo TWINSpan. Umožňuje rychlé manuální přeskupování polohy snímků a druhů v tabulce, automatické řazení druhů, vyhledávání druhů apod. Rovněž je možné provést klasifikaci vybrané části tabulky programem TWINSpan a následné přeskupení. Tabulka vytvořená v tomto programu může být exportována např. do textového editoru nebo do spreadsheetového programu pro finální úpravu.

Do systému může být dále začleněn např. program CANOCO (ter Braak 1987) pro přímou a nepřímou ordinaci (TURBOVEG může hlavičková data exportovat do souboru s daty o prostředí pro kanonickou korespondenční analýzu nebo analýzu redundance), dále program CanoDraw (Šmilauer 1993) pro grafickou prezentaci výsledků ordinací vytvořených programem CANOCO, programy pro síťové mapování apod. Vzhledem k tomu, že uvedené programy nejsou na rozdíl od programů popisovaných výše volně šiřitelné, nejsou v balíku TURBO(VEG) dodávány.

Minimální hardwarové nároky jsou procesor 80286 s 1–1,4 MB rozšířené paměti. Systém pracuje v prostředí MS-DOS.

Zhodnocení využitelnosti systému TURBO(VEG) v české fytocenologii

Výhody systému TURBO(VEG) jsou evidentní z výše uvedeného popisu. Jedná se o software, který je v současné době rozšířen na řadě pracovišť po celé Evropě a stane se nepochybně standardem pro projekt European Vegetation Survey. Jeho autor průběžně pracuje na zdokonalování podle připomínek uživatelů a podle požadavků z jednotlivých zemí. Neocenitelnou výhodou je možnost práce s prakticky neomezeně velkými soubory fytocenologických dat a propojení s celou řadou standardně používaných programů.

Pro usnadnění přechodu ze systému EDIFYT na databázi TURBOVEG byl v rámci projektu Darwin Initiative vyvinut program CODE96 (L. Tichý, Masarykova univerzita Brno, 1996), který zabezpečuje relativně snadný převod dat mezi oběma formáty. Je možné nejen uložit data vytvořená ve formátu EDIFYT do formátu TURBOVEG, ale také provést export z databáze TURBOVEGu do EDIFYTu. Uživatel tak může současně používat výhod obou systémů.

Správné zajištění funkce velké databáze, kterou současně buduje a využívá větší množství pracovníků, naráží na některé obtíže:

(1) Jako relační databáze nepracuje TURBOVEG s plnými jmény druhů, ale s čísly druhů podle standardního seznamu. Proto není možné, aby jednotliví uživatelé libovolně doplňovali druhy do seznamu, protože by pak při výměně dat mohly být k některým údajům přiřazeny jiné druhy podle seznamu použitého na jiném počítači. Na základě zkušeností z práce na nizozemském přehledu vegetace (Schaminée et al. 1995 a následující díly) bylo navrženo, aby v každé ze tří zemí, které budou používat společný druhový seznam, byla pouze jediná osoba pověřena péčí o databázi druhů. Ostatní uživatelé budou v případě potřeby doplnění druhového seznamu této osobě posílat, nejlépe prostřednictvím elektronické pošty, svoje požadavky a na jejich základě obdrží doplněnou verzi seznamu druhů v elektronické podobě. Mezi rakouským, českým a slovenským správcem databáze bude zajištěna koordinace. Pro Českou republiku bude uvedený servis zabezpečovat autor tohoto článku na adrese elektronické pošty chytry@sci.muni.cz.

(2) Podobným problémem je zadávání údajů do bibliografické databáze. Protože však vložení jedné citace do databáze dvakrát nebo vícekrát nemůže způsobit chybné přiřazení ke snímkům, bude nejlepším řešením tohoto problému přidělení určitého rozsahu čísel, pod kterými může být citace uložena, jednotlivým institucím, resp. uživatelům. Pro Českou republiku byl stanoven rozsah čísel 300.001 až 600.000. V rámci tohoto rozsahu bude čísla 300.001–400.000 používat Botanický ústav AV ČR, čísla 400.001–500.000 Masarykova univerzita v Brně a čísla 500.001–530.000 Univerzita Palackého v Olomouci. Další čísla budou přidělena jednotlivým pracovištím spolu s kopiemi programu. Při instalaci programu na každé nové pracoviště bude ustaven lokální správce, který bude přidělovat čísla jednotlivým pracovníkům.

Stejně tak by mělo být zajištěno rozdělení čísel fytoecologických snímků vkládaných do databáze, aby každý snímek zadaný do TURBOVEGu v Rakousku, České republice nebo na Slovensku měl svoje jedinečné číslo. Dohodnuté rozdělení čísel snímků pro státy a pro pracoviště je stejné jako rozdělení čísel pro bibliografické údaje.

(3) Relativní nevýhodou rakousko-česko-slovenské verze je použitá nomenklatura taxonů cévnatých rostlin, protože není dosud publikována a na některá jména v ní použitá nejsou čeští uživatelé zatím příliš zvyklí. Přes tyto nevýhody byl do TURBOVEGu, s ohledem na budoucí využití, zabudován právě tento seznam, protože lze očekávat, že po své publikaci se stane středoevropským standardem nomenklatury nejen pro fytoecologii, ale i pro fyto geografické síťové mapování. V současné době se intenzívně pracuje na doplnění synonym a autoři rovněž plánují před publikací do tohoto seznamu zahrnout flóry dalších středoevropských zemí (hotový už je např. doplněk druhů pro Maďarsko a zájem o rozšíření na své země projeví v poslední době také pracovníci z Německa a Slovinska). Po publikaci bude současný seznam nahrazen definitivní verzí, kompatibilní s verzí současnou, která bude distribuována uživatelům TURBOVEGu.

Návrh společných standardů pro databázi TURBOVEG v ČR

Vzhledem k usnadnění budoucí výměny dat mezi uživateli TURBOVEGu byly na základě diskuse s H. Hártelem, L. Tichým a M. Duchoslavem navrženy některé společné zásady, jak ukládat data v rámci databází používaných v České republice. Většina těchto návrhů byla akceptována i rakouskými a slovenskými fytoecology (L. Mucina, M. Valachovič, I. Jarolímek):

(1) Pro kódování češtiny v hlavičkových datech by měl být zásadně používán kód Latin 2.

(2) V TURBOVEGu je možnost zadávat u druhů nižší a vyšší, případně střední stromové, resp. keřové patro. Jestliže tato „podpatra“ nejsou rozlišována (což platí pro většinu našich fytoecologických snímků), mělo by být patro vždy zadáno do pole pro nejvyšší „podpatro“. Tzn. obvykle bude mít stromové patro kód 1, keřové patro 4, bylinné patro 6, juvenilní nebo zakrslé stromy a keře v bylinném patru (včetně semenáčů) 7 a mechorosty nebo lišejníky 9.

(3) Pokryvnosti pater v hlavičkových datech se v TURBOVEGu zadávají do numerického pole, ve kterém se automaticky, není-li nic jiného zadáno, objevuje nula. Ta však může znamenat buď, že patro má nulovou pokryvnost, nebo že pokryvnost je nenulová, ale autor snímku neudal hodnotu. Pro odlišení těchto dvou možností navrhuje v případě, že dané patro bylo ve snímku přítomno, ale jeho pokryvnost nebyla zaznamenána, vyplnit do daného pole -1.

(4) Ve standardním formuláři hlavičkových dat evropské verze TURBOVEGu je řada polí, které naši fytoecologové sotva budou používat, a na druhé straně pole pro některé významné údaje si každý bude muset definovat sám. Pro zajištění lepší výměny dat navrhuje zařadit tato pole (ve stejném pořadí):

Název pole	typ pole	délka pole	počet desetinných míst
LOCALITY	C	100	0
LONGITUDE	N	6	0
LATITUDE	N	6	0
CEBA_GRID	C	6	0
FIELD_NR	C	10	0
HABITAT	C	100	0
GEOLOGY	C	50	0
SOIL	C	50	0

První pole zahrnuje geografické určení lokality (viz bod 5), druhé a třetí pole zeměpisnou délku a šířku s přesností alespoň na desítky sekund (např. 160310 = 16° 3' 10" v. d.), čtvrté pole číslo kvadrantu (tj. čtvrtiny kvadrátu CEBA – Central European Basic Area, např. 6863c) sítě střeoevropského fyto geografického mapování (viz bod 6), páté pole obsahuje pracovní číslo každého autora, šesté pole stručný popis stanoviště (např. světlna v lese na

strmém svahu pod vrcholem hřebene), sedmé pole stručné označení geologického substrátu (např. ortorula) a osmé pole půdní druh podle nomenklatury Morfogenetického klasifikačního systému půd (viz Hraško et al. 1987, Novák 1989–1993) (např. kambizem).

(5) Lokality v textové formě by měly být uváděny tak, aby na prvním místě bylo vždy jméno obce, následované bližším určením lokality. Vzhledem k současnému udávání zeměpisných souřadnic nebo UTM čtverců by neměla hrozit záměna mezi dvěma obcemi stejného jména. V případě, že ale zadavatel má přesto pocit, že bližší lokalizace (např. orografický celek, politický okres apod.), by byla vhodná, měla by být uvedena v závorce za jménem obce. Např.:

Stojice (Železné hory), Svinčanský les 1 km SSV obce.

Lukov (NP Podyjí), svah nad levým břehem Dyje mezi ústím Klaperova potoka a Rambachu 2,1 km JZ obce.

(6) TURBOVEG standardně používá pro definování polohy síť UTM (Universal Transverse Mercator). Protože přímé zadávání polí UTM je prakticky výhodné pouze u nových snímků, zaměřených v terénu s použitím globálního systému určování polohy (GPS – Global Position System), je možné místo toho zadat pouze zeměpisné souřadnice (autor programu vyjádřil zájem o začlenění automatického převodníku mezi zeměpisnými souřadnicemi a poli UTM – Hennekens 1996, ústní sdělení). Buď pole UTM GRID SYSTEM CODE ze standardní hlavičky TURBOVEGu s přesností na čtverce 0,1 x 0,1 km nebo pole LONGITUDE a LATITUDE by však měly být vyplněny povinně a po jejich vyplnění už by nebylo nutné vyplňovat pole CEBA_GRID. V případě, kdy není možné určit zeměpisné souřadnice s přesností alespoň na desítky sekund (např. při přebírání údajů z literatury), měly by být místo souřadnic vyplněny, pokud je to možné, odkazy na čtverce středoevropského síťového mapování 5 x 3 nebo alespoň 10 x 6 a zároveň na UTM čtverce 1 x 1 km nebo větší.

(7) Literatura by měla být zadávána do jednotlivých polí podle následujícího vzoru:

Author: Kolbek J. & Petříček V.

Huml O. et al.

Year: 1995

Title: [Název publikace bez tečky na konci]

Published in: Folia Geobot. Phytotax., Praha, 7: 225–258.

230 pp., Academia, Praha.

In: Jeník J. (ed.), Vegetační problémy při budování vodních děl. NČSAV, Praha, pp. 21–37.

Jak TURBO(VEG) získat?

Autor programu Stephan Hennekens poskytl TURBO(VEG) zdarma botanické veřejnosti s cílem, aby sloužil pro společnou práci a výměnu dat mezi jednotlivými pracovníky na projektu European Vegetation Survey, respektive na projektech národních přehledů vegetace. TURBO(VEG) tedy může získat prakticky každý zájemce z řad botanické veřejnosti, který je ochoten na těchto projektech spolupracovat a svoje data zadaná do databáze (samozřejmě při dodržení zásad ochrany autorských práv) poskytovat zpracovatelům jednotlivých syntaxonů, případně některé typy vegetace sám zpracovávat. Přitom je nutné, aby každý uživatel dodržoval společné zásady práce s databází, zejména neměnil samostatně seznam druhů a dodržoval pro něj vymezený rozsah čísel pro označení snímků a pro bibliografickou databázi. Zájemci o získání programu z Botanického ústavu AV ČR mohou kontaktovat H. Härtela, z Univerzity Palackého M. Duchoslava a všichni ostatní M. Chytrého. Vlastní software může být zaslán elektronickou poštou (kódován jako UUENCODE nebo BinHex) nebo na disketách, v případě zájmu i s programem CODE96.

Poděkování

Za všechny (i budoucí) české uživatele TURBOVEGu bych chtěl vyjádřit dík dr. Stephanu Hennekensovi, který velkoryse umožnil bezplatné používání svého software a pohotově zapracoval řadu modifikací podle našich potřeb, a dr. Johnu Rodwellovi a Julianu Dringovi, kteří zorganizovali školení na Univerzitě v Lancasteru v rámci projektu „Safeguarding the Biodiversity of Eastern Europe“ (součást Darwin Initiative) a umožnili tak vlastní rozšíření tohoto software u nás. Velký dík patří rovněž prof. dr. Haraldovi Niklfeldovi, dr. Walteru Gutermannovi a prof. dr. Ladislavu Mucinovi za jejich obětavou práci na vytvoření databáze druhů pro použití u nás. Mgr. Lubomír Tichý vypracoval v rámci projektu Darwin Initiative převodní program a ochotně jej zdarma poskytl českým fytocenologům. Společné standardy práce s databází TURBOVEG pro naši fytocenologii byly navrženy ve spolupráci s prof. dr. Ladislavem Mucinou, dr. Milanem Valachovičem, ing. Handrijem Härtel, mgr. Lubomírem Tichým, mgr. Martinem Duchoslavem a dr. Ivanem Jarolímekem.

Summary

Since 1989 the user-friendly computer program FYT by Eduard Brabec (Průhonice) has been used for editing phytosociological tables and storing computerized data in Czech phytosociology. This program, however, was not designed as a database system. As storing a large number of computerized relevés is absolutely necessary for making large-scale vegetation surveys, database system TURBO(VEG) by Stephan Hennekens (Wageningen) was currently installed in several institutes in the Czech Republic, like in a number of other European countries. The installed version designed for the use in the Czech Republic was prepared in the Unit of Vegetation Science, Lancaster University, as a part of the Darwin Initiative project, and includes a common species list with Austria and Slovakia. The program CODE96 for the data conversion between FYT and TURBOVEG was prepared by Lubomír Tichý (Brno), also as a part of the Darwin Initiative project.

The present paper provides a brief description of the TURBO(VEG) system, evaluates its applicability with specific respect to the aims of Czech phytosociology and proposes common data standards for the database.

Literatura

- Dostál J. (1989): Nová květena ČSSR. – Academia, Praha.
- Ehrendorfer F. (ed.) (1973): Liste der Gefäßpflanzen Mitteleuropas. Ed. 2. – G. Fischer, Stuttgart.
- Frey W., Frahm J. P., Fischer E. et Lobin W. (1995): Die Moos- und Farnpflanzen Europas. – G. Fischer, Stuttgart.
- Hejný S. et Slavík B. (eds.) (1988): Květena České socialistické republiky. Vol. 1. – Academia, Praha.
- Hennekens S. M. (1995): TURBO(VEG). Software package for input, processing, and presentation of phytosociological data. User's guide. – IBN-DLO Wageningen et University of Lancaster.
- Hill M. O. (1979): TWINSpan. A Fortran program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. – Cornell Univ., Ithaca.
- Hraško J., Němeček J., Šály R. et Šurina B. (1987): Morfogenetický klasifikační systém pôd ČSSR. – Bratislava.
- Kučera T. (1992): [rec.] FYT. Editor fytoecologických tabulek. – Preslia, Praha, 64: 78.
- Moravec J. et al. (1995): Rostlinná společenstva České republiky a jejich ohrožení. Ed. 2. – Severočes. Přír., append. 1995: 1–206.
- Novák P. et al. (1989–1993): Syntetická půdní mapa České republiky 1 : 200 000. – Ministerstvo zemědělství et Ministerstvo životního prostředí České republiky, Praha.
- Pišút I., Lackovičová A. et Lisická E. (1993): Súpis lišajníkov Slovenska. – Biológia, Bratislava, 48, suppl. 1: 53–98.
- Podani J. (1994): Multivariate data analysis in ecology and systematics. A methodological guide to the SYN-TAX 5.0 package. – SPB Academic Publ., Amsterdam.
- Schaminée J. H. J., Stortelder A. H. F. et Westhoff V. (1995): De Vegetatie van Nederland. 1. Inleiding tot de plantensociologie – grondbeginselen, methoden en toepassingen. – OPULUS Press, Uppsala.
- Šmilauer P. (1993): CanoDraw 3.0. – Univ. College, London.
- ter Braak C. J. F. (1987): CANOCO – a Fortran program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis (version 2.1). – Agricultural Mathematics Group, Wageningen.
- Türk R. et Poelt J. (1993): Bibliographie der Flechten und flechtenbewohnenden Pilzen in Österreich. – Verlag der Österreichischen Akademie der Wissenschaften, Wien.
- Tutin T. G. et al. [eds.] (1964–1980): Flora europaea. Vol. 1–5. – Cambridge Univ. Press, Cambridge.
- Wildi O. et Orlóci L. (1996): Numerical exploration of community patterns. A guide to the use of MULVA-5. Ed. 2. – SPB Academic Publ., Amsterdam.