

RYCHLÁ KONTROLA KVALITY DRUHŮ JEDLÉHO HMYZU: MULTI-SPECIES FT-NIR KALIBRACE V ROZSAHU 13200–8700 CM⁻¹ PRO PREDIKCI OBSAHU NUTRIENTŮ

RAPID QUALITY CONTROL OF EDIBLE INSECTS: MULTI-SPECIES FT-NIR CALIBRATION IN THE 13200–8700 CM⁻¹ REGION FOR NUTRIENT PREDICTION

Matyáš Krτίčka¹

¹ Katedra mikrobiologie, výživy a dietetiky, Fakulta agrobiologie, potravinových a přírodních zdrojů,
Česká zemědělská univerzita v Praze, Kamýcká 129, 165 00 Praha – Suchbát, Česká republika

<https://doi.org/10.11118/978-80-7701-100-6-0211>



ABSTRAKT

Cílem této práce bylo vyvinout a ověřit multi-species FT-NIR kalibrační model pro simultánní predikci obsahu sušiny, hrubé bílkoviny, tuku a popelovin u vybraných druhů jedlého hmyzu (*Acheta domesticus*, *Tenebrio molitor* a *Alphitobius diaperinus*) v omezeném spektrálním rozsahu 13200–8700 cm⁻¹, typickém pro přenosné NIR spektrometry. Vzorby hmyzu byly lyofilizovány, homogenizovány a analyzovány Fourierovou transformační NIR spektroskopií v transmisním módu. Referenční hodnoty jednotlivých nutrientů byly stanoveny standardními laboratorními metodami. Kalibrační modely byly vytvořeny metodou parciálních nejmenších čtverců (PLS) s využitím vhodných preprocessing postupů a jejich kvalita byla hodnocena pomocí křížové validace. Nejlepších výsledků bylo dosaženo u kalibrace sušiny (R²CV = 0,93, RPD = 3,2), která vykazovala vynikající predikční schopnost. Kalibrace pro stanovení obsahu proteinu dosáhla dobré kvality (R²CV = 0,89, RPD = 2,8), zatímco modely pro tuky (R²CV = 0,86, RPD = 2,2) a popeloviny (R²CV = 0,75, RPD = 2,11) vykazovaly nižší predikční přesnost. Výsledky potvrzují, že multi-species FT-NIR kalibrace v omezeném spektrálním rozsahu představuje perspektivní nástroj pro rychlou a nedestruktivní kontrolu nutričního složení jedlého hmyzu.

Klíčová slova: FT-NIR, jedlý hmyz, kalibrační model, preprocessing

ABSTRACT

The aim of this study was to develop and validate a multi-species FT-NIR calibration model for the simultaneous prediction of dry matter, crude protein, fat, and ash content in selected edible insect species (*Acheta domesticus*, *Tenebrio molitor*, and *Alphitobius diaperinus*) using a limited spectral range of 13200–8700 cm⁻¹, characteristic of portable NIR spectrometers. Insect samples were freeze-dried, homogenized, and analyzed by Fourier transform near-infrared spectroscopy in transmission mode. Reference values of individual nutrients were determined using standard laboratory methods. Calibration models were developed using partial least squares (PLS) regression with appropriate spectral preprocessing, and their performance was evaluated by cross-validation. The best results were obtained for dry matter calibration (R²CV = 0.93, RPD = 3.2), demonstrating excellent predictive performance. The protein calibration achieved good accuracy (R²CV = 0.89, RPD = 2.8), while the models for fat (R²CV = 0.86, RPD = 2.2) and ash (R²CV = 0.75, RPD = 2.11) showed lower predictive ability. The results confirm that multi-species FT-NIR calibration within a limited spectral range represents a promising tool for rapid and non-destructive assessment of the nutritional composition of edible insects.

Keywords: FT-NIR, edible insect, calibration model, preprocessing

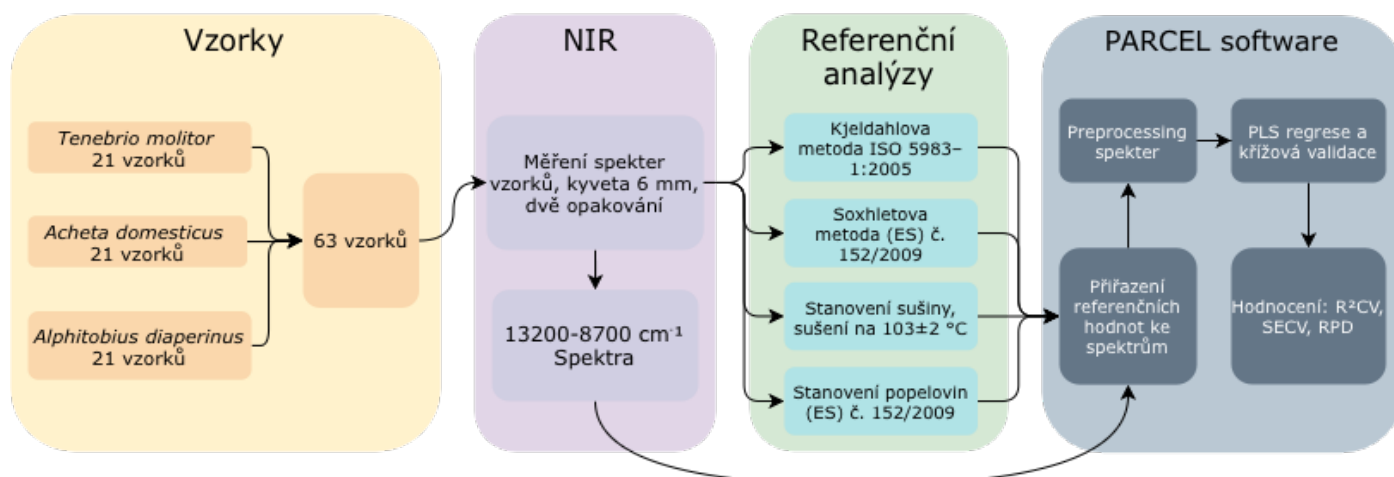
ÚVOD

Spektroskopie blízké červené oblasti (NIR) je nedestruktivní analytická metoda umožňující rychlou a relativně levnou predikci chemického složení materiálů na základě interakce elektromagnetického záření se vzorkem. Tato technologie se v posledních letech uplatňuje zejména v kvalitativní kontrole a analýze nutričního složení potravinářských produktů. Jedlý hmyz představuje udržitelný a environmentálně šetrný zdroj proteinu a esenciálních živin (Johnson, 2020). Na rozdíl od tradičního zemědělství vyžaduje hmyzí chov méně půdy, vody a energetických zdrojů, přičemž jeho produkce generuje nižší emise skleníkových plynů. Se zvyšujícím se zájmem o hmyzí proteiny roste poptávka po standardizovaných a efektivních metodách pro kontrolu kvality a určování obsahu živin v jedlém hmyzu. Současné aplikace NIR spektroskopie zaměřené na analýzu hmyzu se soustředí převážně na jednotlivé druhy nebo specifické parametry. Vývoj a údržba samostatných kalibrací pro každý druh je finančně i časově náročnější (Chadalavada et al., 2022). Multi-species kalibrační modely, které umožňují simultánní analýzu více druhů v jednom systému, představují zajímavé řešení. Navíc řada moderních a přenosných NIR přístrojů operuje v omezeném spektrálním rozsahu, což představuje technickou výzvu pro vývoj robustních predikčních modelů (de Lamo Castellvi et al., 2025). Cílem této práce bylo vyvinout spolehlivý, robustní FT-NIR kalibrační model pro simultánní predikci obsahu proteinu, tuku, popelovin a sušiny pro tři vybrané druhy jedlého hmyzu (*Acheta domesticus*, *Tenebrio molitor*, *Alphitobius diaperinus*) s využitím omezeného spektrálního rozsahu 13200–8700 cm⁻¹, typického pro nové generace přenosných spektrometrů. Tento přístup by měl poskytnout praktickou metodu pro efektivnější kvalitativní kontrolu ve výrobě a distribuci hmyzích produktů (Benes et al., 2022).

MATERIÁL A METODIKA

Příprava vzorků

Byly zkoumány tři druhy hmyzu: *Acheta domesticus* (AD), *Tenebrio molitor* (TM) a *Alphitobius diaperinus* (AID). Vzorky pocházely z chovných zařízení České zemědělské univerzity v Praze (ČZU) nebo byly získány od společnosti PAPEK s.r.o., která je zajišťovala od různých chovatelů. Hmyz byl chován v plastových nádobách za různých podmínek prostředí. Krmení se lišilo podle druhu a chovatele. Druh AD a AID byly krmeny obilovinou, pšeničnými klíčky nebo otrubami, doplněnými jablkem, mrkví nebo bramborem. Druh TM dostával upravené drůbežní krmivo se pšeničnými otrubami v poměru 1:4. Před usmrcením byl všechn hmyz 24 hodin bez potravy. Usmrcení probíhalo dvěma metodami. První metoda spočívala ve zmrazení při –20 °C. Druhá metoda zahrnovala vaření ve vodě v poměru vzorek ku vodě 1:10 po dobu 5 minut. Vzorky byly následně skladovány v mrazáku při –18 °C. Vzorky hmyzu podstoupily lyofilizaci, počínaje zmrazením při –80 °C po dobu jedné hodiny, následované lyofilizací pod nízkým tlakem (L10-55 PRO, GREGOR Instruments s.r.o., Česká republika). Lyofilizovaný hmyz byl homogenizován pomocí laboratorního mlýnku (IKA MultiDrive basic, Vekron, CZ) a uskladněn v mrazáku při –80 °C do doby analýzy. Na obrázku 1 je znázorněn experimentální postup zahrnující přípravu vzorků, paralelní NIR měření a referenční analýzy, následovaný preprocesingem spekter, tvorbou PLS kalibrace a validací modelů.



Obrázek 1: Schéma experimentálního postupu FT-NIR kalibrace pro stanovení nutričních složek v jedlém hmyzu

Referenční metody

Obsah sušiny byl stanoven gravimetrickou metodou. Vzorky byly sušeny v elektrické sušárně (Memmert, Schwabach, Německo) při teplotě 103 ± 2 °C po dobu 3,5 hodin. Doba sušení byla stanovena na základě předběžné analýzy, která sledovala dosažení konstantní hmotnosti vzorku, přičemž byl zohledněn vliv oxidace lipidů a případného zvýšení hmotnosti. Obsah popelovin byl analyzován spalováním v muflové peci (Vekron, Praha, Česká republika) při 550 °C v souladu s Nařízením Evropské komise. Stanovení dusíku proběhlo pomocí Kjeldahlovy metody dle normy ISO 5983–1:2005 za použití analyzátoru Kjelttec 2400 (FOSS, Hillerød, Dánsko). Obsah tuku byl určen Soxhletovou extrakcí v souladu s Nařízením (ES) č. 152/2009, přičemž jako rozpouštědlo byl použit petrolether (max. 0,01 % aromatických uhlovodíků, 40–60 °C, VWR International S.A.S., Francie). Obsah hrubých bílkovin byl vypočten z naměřeného obsahu dusíku pomocí konverzního faktoru 6,25. Zároveň byly aplikovány také druhově specifické přepočítávací faktory vycházející z dostupné vědecké literatury pro výpočet skutečného proteinu. V případech, kdy literatura obsahovala více faktorů, byla pro výpočty použita jejich průměrná hodnota.

$$Protein = N\% \times 6,25$$

NIR spektroskopické měření vzorků

Pro vytvoření kalibračního modelu byly vzorky jedlého hmyzu analyzovány na Fourierově transformačním NIR analyzátoru Lumex InfraLUM FT-12 v transmisním módu ve spektrální oblasti 13200–8700 cm^{-1} . Před samotným měřením byla v programu Parcel nastavena metoda pro přístroj, přičemž kalibrační kanál byl nastaven na rozlišení 16 cm^{-1} s Besselovou apodizační funkcí, referenční kanál byl nastaven na rozlišení 32 cm^{-1} rovněž s Besselovou apodizací. Před samotnou kalibrací byl přístroj podroben System suitability testu (SST) s referenční kyvetou pro stanovení základní linie. Celkem bylo analýze vybráno 63 vzorků a to AD – 21, TM – 21 a AID – 21. Vzorky vytemperované na pokojovou teplotu byly vloženy do 6 mm skleněné kyvety a proměřeny vždy se ve dvou opakováních. Pro každé spektrum bylo měřeno vždy 12 skenů. K získaným NIR spektrům byly přiřazeny referenční hodnoty obsahu proteinu, tuku, popelovin a sušiny stanovené standardními analytickými metodami. Naměřená spektra byla exportována z přístroje a zpracována v programu Parcel pro tvorbu kalibračních modelů metodou Partial Least Squares (PLS).

Analýza dat

Spektra byla před vývojem kalibračních modelů podrobena preprocessing postupům specifickým pro každý analyzovaný parametr. Pro predikci obsahu proteinů byla aplikována korekce baseline následovaná

normalizací směrodatnou odchylkou a výpočtem první derivace. Kalibrace pro stanovení tuku zahrnovala multiplikativní korekci, normalizaci směrodatnou odchylkou, vážený průměr a první derivaci. U popeloviny byla provedena korekce baseline, škálování odchylek (deviations scale), normalizace směrodatnou odchylkou, vážený průměr a druhá derivace. Pro predikci sušiny byl použit postup zahrnující korekci baseline, normalizaci směrodatnou odchylkou, vážený průměr a první derivaci. Pro každý nutrient (protein, tuk, popeloviny, sušina) byl vytvořen kalibrační model na základě vztahu mezi FT-NIR spektry a referenčními hodnotami stanovenými standardními laboratorními metodami. Kvalita modelů byla ověřena metodou křížové validace (cross-validation), která poskytuje realistický odhad predikční schopnosti modelu na nezávislých datech.

Kvalita kalibračních modelů byla hodnocena pomocí tří statistických parametrů vypočítaných pro kalibrační soubor i křížovou validaci. Koeficient determinace (R^2) vyjadřuje míru lineární závislosti mezi referenčními a predikovanými hodnotami, kde n je počet vzorků, y_i je referenční hodnota vzorku i , $\hat{y}_i$ je predikovaná hodnota vzorku i a $\bar{y}$ je průměr referenčních hodnot. Hodnoty R^2 blížící se 1 indikují silnou korelaci mezi modelem a referenčními daty. Pro hodnocení byly rozlišovány R^2_{cal} (kalibrační soubor) a R^2_{CV} (křížová validace).

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}$$

Střední kvadratická chyba (RMSE) kvantifikuje průměrnou odchylku predikovaných hodnot od referenčních hodnot vyjádřenou ve stejných jednotkách jako měřený parametr. Nižší hodnoty RMSE indikují vyšší přesnost predikce. Rozlišovaly se SEC (standard error of calibration) pro kalibrační soubor a SECV (standard error of cross-validation) pro křížovou validaci.

$$RMSE = \sqrt{\frac{\sum (\hat{y}_i - y_i)^2}{n}}$$

Poměr predikční schopnosti vůči odchylce (RPD) je definován jako poměr směrodatné odchylky (SD) referenčních hodnot k SECV. RPD představuje primární kritérium pro hodnocení použitelnosti NIR kalibrace. Hodnoty RPD byly interpretovány podle standardních kritérií: RPD <1,5 indikuje nepoužitelnou kalibraci, 1,5–2,0 umožňuje hrubé třídění vzorků, 2,0–2,5 je vhodné pro kvantitativní predikce, 2,5–3,0 představuje dobrou kalibraci a RPD >3,0 vynikající kalibraci s vysokou predikční přesností.

$$RPD = \frac{SD}{RMSE}$$

SECV a R^2_{CV} představovaly primární kritéria pro hodnocení predikční schopnosti modelů, neboť poskytují realistický odhad chyby predikce na nezávislých datech. Poměr SECV/SEC byl použit jako indikátor možného overfittingu modelu, přičemž hodnoty >1,5 signalizují nadměrné přizpůsobení modelu trénovacím datům.

VÝSLEDKY A DISKUZE

Vyvinuté kalibrační modely vykazovaly rozdílnou predikční schopnost v závislosti na analyzovaném nutrientu. Kalibrace pro stanovení obsahu proteinů dosáhla vysoké přesnosti s hodnotou $R^2_{CV} = 0,89$ a SECV = 0,8 %. Koeficient determinace kalibračního souboru $R^2_{cal} = 0,92$ byl mírně vyšší než R^2_{CV} , což indikuje dobrou generalizační schopnost modelu. Hodnota RPD = 2,8 klasifikuje tuto kalibraci jako

dobrou s dostatečnou přesností pro kvantitativní predikce obsahu proteinů v jedlém hmyzu. Poměr $SECV/SEC = 1,33$ potvrzuje, že model není nadměrně přizpůsoben trénovacím datům. Predikční chyba $SECV = 0,8 \%$ představuje přijatelnou úroveň nepřesnosti pro rutinní analýzu proteinů. Model pro predikci obsahu tuku vykazoval nižší predikční schopnost s $R^2CV = 0,86$ a $SECV = 1,0 \%$. Hodnota $RPD = 2,2$ klasifikuje tuto kalibraci jako vhodnou pro kvantitativní stanovení, avšak s nižší přesností ve srovnání s proteiny. Poměr $SECV/SEC = 1,43$ se blíží horní hranici doporučené hodnoty 1,5, což může indikovat počínající overfit modelu. Vyšší predikční chyba u tuků pravděpodobně odráží komplexnější spektrální odezvu lipidových složek a jejich vyšší variabilitu mezi jednotlivými druhy hmyzu. Kalibrace pro stanovení popelovin dosáhla nejnižší predikční schopnosti ze všech analyzovaných nutrientů s $R^2CV = 0,75$ a $SECV = 0,28 \%$. Přestože hodnota $RPD = 2,11$ stále umožňuje kvantitativní predikce, nižší koeficient determinace a vysoký poměr $SECV/SEC = 1,87$ signalizují overfit modelu. Nízká absolutní hodnota $SECV$ je způsobena nízkým obsahem a úzkou distribuční šířkou popelovin ve vzorcích, což se projevuje v nižší robustnosti kalibračního modelu. Tento model vyžaduje opatrnou interpretaci predikovaných hodnot a potenciálně další optimalizaci preprocessing postupů. Kalibrace pro predikci sušiny vykazovala nejlepší výsledky ze všech analyzovaných parametrů s $R^2CV = 0,93$, $SECV = 0,65 \%$ a $RPD = 3,2$. Hodnota RPD překračující 3,0 klasifikuje tento model jako vynikající kalibraci s vysokou predikční přesností. Poměr $SECV/SEC = 1,3$ je nejnižší ze všech modelů, což potvrzuje vynikající generalizační schopnost. Vysoká kvalita této kalibrace je pravděpodobně způsobena širokým distribučním rozsahem obsahu sušiny ve vzorcích a silným spektrálním signálem vody v blízké infračervené oblasti. Data jsou rovněž popsána v tabulce 1.

Porovnání poměrů $SECV/SEC$ napříč všemi modely (protein 1,33, tuk 1,43, popeloviny 1,87, sušina 1,3) ukazuje, že modely pro proteiny a sušinu vykazují nejlepší rovnováhu mezi kalibračním fitem a predikční schopností, zatímco model pro popeloviny trpí nadměrným přizpůsobením trénovacím datům. Vyvinuté multi-species FT-NIR kalibrace prokázaly použitelnost pro kvantitativní rychlou kontrolu nutrientů v jedlém hmyzu, přičemž modely pro sušinu a proteiny dosáhly vynikající až dobré kvality, model pro tuky přijatelné kvality a model pro popeloviny vyžaduje další optimalizaci.

Tabulka 1: Statistické parametry kalibračních modelů pro FT-NIR predikci obsahu nutrientů v jedlém hmyzu

Nutrient	R^2cal	R^2CV	SEC [%]	SECV [%]	RPD	SECV/SEC	SD [%]
Protein	0,92	0,89	0,6	0,8	2,8	1,33	2,24
Tuk	0,9	0,86	0,7	1	2,2	1,43	2,2
Popeloviny	0,82	0,75	0,15	0,28	2,11	1,87	0,59
Sušina	0,95	0,93	0,5	0,65	3,2	1,3	2,08

Legenda: R^2cal – koeficient determinace kalibračního souboru; R^2CV – koeficient determinace křížové validace; SEC – střední kvadratická chyba kalibrace; SECV – střední kvadratická chyba křížové validace; RPD – poměr predikční schopnosti vůči odchylce; SD – směrodatná odchylka referenčních hodnot

ZÁVĚR

V této práci byl úspěšně vyvinut multi-species FT-NIR kalibrační model pro rychlou predikci obsahu sušiny, hrubé bílkoviny, tuku a popelovin u tří druhů jedlého hmyzu. Použití omezeného spektrálního rozsahu $13200\text{--}8700\text{ cm}^{-1}$ umožnilo simulovat podmínky běžné pro moderní přenosné NIR spektrometry, aniž by došlo k zásadnímu snížení kvality kalibrací pro hlavní nutriční parametry. Nejvyšší predikční schopnosti bylo dosaženo u kalibrace sušiny, která vykazovala vynikající hodnoty koeficientu determinace a poměru RPD, což potvrzuje silný a stabilní spektrální signál vody v blízké infračervené oblasti. Kalibrace pro stanovení obsahu proteinů dosáhla dobré úrovně přesnosti a je vhodná pro kvantitativní použití v rutinní analýze. Model pro predikci tuků poskytoval přijatelné výsledky, avšak s nižší přesností, pravděpodobně v důsledku vyšší variability lipidového složení mezi jednotlivými druhy hmyzu. Nejnižší kvalitu vykazovala

kalibrace popelovin, u níž byla identifikována známka nadměrného přizpůsobení modelu, což omezuje její praktickou využitelnost bez další optimalizace.

Celkově lze konstatovat, že multi-species FT-NIR přístup představuje efektivní a perspektivní nástroj pro rychlou kontrolu kvality jedlého hmyzu, zejména pro stanovení sušiny a proteinů. Další zlepšení predikční schopnosti modelů, především u tuků a popelovin, lze očekávat rozšířením kalibračního souboru, optimalizací preprocessing postupů a zpřesněním referenčních dat.

PODĚKOVÁNÍ

Děkuji kolegům z Katedry mikrobiologie, výživy a dietetiky prof. Ing. Lence Kouřimské, Ph.D. a Ing. Pavlovi Kouřimskému za ochotu a pomoc při práci. Tato práce vznikla za podpory projektu MŠMT METROFOOD-CZ, č. LM2023064.

LITERATURA

Benes, E., Biró, B., Fodor, M., Gere, A. (2022): Analysis of wheat flour-insect powder mixtures based on their near infrared spectra. *Food Chemistry: X*, 13, 100266.

Dostupné z: <https://doi.org/10.1016/j.fochx.2022.100266>

De Lamo Castellvi, S., Méndez Sánchez, C., Matos, C., Wu, Y., Plans, M., Rodriguez-Saona, L. (2025): FT-NIR spectroscopy for protein and lipid analysis in insect powders and flour mixtures: A cross-instrument calibration approach. *Microchemical Journal*, 218, 115315. Dostupné z: <https://doi.org/10.1016/j.microc.2025.115315>

Chadalavada, K., Anbazhagan, K., Ndour, A., Choudhary, S., Palmer, W., Flynn, J. R., Mallayee, S., Pothu, S., Prasad, K. V. S. V., Varijakshapanikar, P., Jones, C. S., Kholová, J. (2022): NIR Instruments and Prediction Methods for Rapid Access to Grain Protein Content in Multiple Cereals. *Sensors*, 22(10), 3710. Dostupné z: <https://doi.org/10.3390/s22103710>

Johnson, J. B. (2020): An overview of near-infrared spectroscopy (NIRS) for the detection of insect pests in stored grains. *Journal of Stored Products Research*, 86, 101558. Dostupné z: <https://doi.org/10.1016/j.jspr.2019.101558>

Kontaktní adresa: Ing. Matyáš Krtička, Katedra mikrobiologie, výživy a dietetiky, Fakulta agrobiologie, potravinových a přírodních zdrojů, Česká zemědělská univerzita v Praze, Kamýcká 129, 165 00 Praha – Suchbátka, Česká republika, e-mail: krticka@af.czu.cz