

Geografiese klassifikasie van Suid-Afrikaanse heuning met naby-infrarooi spektroskopie

M Vermeulen,¹ M Allsopp,² F Marini,^{1,3} P Williams,¹ M Manley¹

¹Departement Voedselwetenskap, Universiteit Stellenbosch, Suid-Afrika

²Plantbeskermingsnavorsingsinstituut, Landbounavorsingsraad, Suid-Afrika

³Departement Chemie, Universiteit van Rome, Italië

Korresponderende outeur: M Manley **E-pos:** mman@sun.ac.za

Geografiese en bloemoorsprong bepaal die samestelling van heuning. Melissopalynologie is die tradisionele metode om die bron van die stuifmeel in heuning te bepaal. Hierdie metode van analise is moeisam, tydrowend en moet deur 'n deskundige uitgevoer word. Metodes wat vinnig en maklik uitvoerbaar is om die geografiese oorsprong van heuning te bepaal is dus belangrik en nodig. Hierdie studie het daarom ten doel gehad om naby-infrarooi (NIR) spektroskopie, tesame met partiële kleinste kwadrate-diskriminantanalise (PLS-DA) en verskillende voorverwerkingstegnieke te evalueer vir die vinnige klassifikasie van Suid-Afrikaanse heuning van drie verskillende geografiese oorspronge in die Wes-Kaap. Twee verskillende NIR-instrumente, naamlik 'n laboratoriuminstrument en 'n handapparaat (glasflesie en Teflonkoppie) is onderskeidelik in transfleksie- en reflektansiemodus gebruik. Hoofkomponentanalise-tellinggrafieke het skeiding getoon tussen die Hopefield- en die Stanford-klas. Die Stellenbosch-klas het oorvleuel met beide die Hopefield- en Stanford-klas as gevolg van Stellenbosch se plantegroei wat as nektarbron vir die bye gedien het, wat 'n kombinasie van fynbos en *Eukaliptus* was, terwyl dit in die geval van Hopefield en Stanford onderskeidelik fynbos en *Eukaliptus* was. Die beste voorspellingsakkuraatheid (84,21%) om tussen die drie geografiese liggings te klassifiseer is met die handapparaat (glasflesie) tesame met PLS-DA en eksterneparameter-ortogonalisering verkry. Vir die datastelle wat in hierdie studie gebruik is, het dit geblyk dat die nektarbron grootliks bygedra het tot die geografiese klassifikasie. NIR spektroskopie sou dus ook as 'n ondersteunende siftingsmetode vir tradisionele melissopalynologie oorweeg kon word.

Sleutelwoorde: chemometrika, diffusierefleksie, DUPLEX, fynbos, Kennard-Stone, mobiele handapparaat, NIR, partiële kleinste kwadrate-diskriminantanalise, PLS-DA, tafellaboratoriuminstrument, transfleksie

Geographical classification of South African honey with near-infrared spectroscopy: Geographical and floral origin determines the composition of honey. Melissopalynology analysis is the traditional method for determining the source of the pollen in honey. This method of analysis is difficult, time-consuming and needs to be executed by an expert. Methods that are rapid and easily performed to determine the geographical origin of honey are therefore important and necessary. This study aimed to evaluate NIR spectroscopy, together with partial least squares discriminant analysis (PLS-DA) and different pre-processing methods to classify South African honey from three different geographical locations in the Western Cape. Two different NIR instruments, i.e. a laboratory instrument and a handheld device (glass tube and Teflon cup), were used in transreflectance and reflectance mode respectively. Principal component analysis score plots showed separation between the Hopefield and Stanford classes. The Stellenbosch class overlapped with both the Hopefield and Stanford classes due to the Stellenbosch vegetation, which served as a nectar source for the bees, being a combination of fynbos and *Eucalyptus*, while for Hopefield and Stanford, it was predominantly fynbos and *Eucalyptus* respectively.

The best prediction accuracy (84,21%) to classify between the three geographical origins was achieved with the handheld device (glass vial) together with PLS-DA and external parameter orthogonalisation as pre-processing method. From the data sets used in this study, the nectar source contributed greatly to the geographical classification. NIR spectroscopy could thus be considered as a supportive screening method for traditional melissopalynology.

Keywords: benchtop device, chemometrics, DUPLEX, diffuse reflection, fynbos, Kennard-Stone, mobile handheld device, NIR, partial least squares discriminant analysis, PLS-DA, transreflectance

Inleiding

Heuning, waarna dikwels verwys word as 'vloeiende goud', word reeds vir duisende jare as voedsel sowel as vir die gesondheids- en medisinale voordele daarvan gebruik (Ball, 2007; Francis & Gontier, 1982; White, 1976). Die eienskappe en samestelling van heuning word beïnvloed deur die bloemoorsprong en geografiese ligging (Latorre, et al., 2013; Salonen & Julkunen-Tiitto, 2012). Die tradisionele metode om die bloemoorsprong van heuning te bepaal is melissopalinalogie. Melissopalinalogie is die mikroskopiese ontleding van stuifmeel teenwoordig in heuning, spesifiek om die nektarbron te identifiseer. Deur die stuifmeel in 'n monster heuning te bestudeer, kan bewyse verkry word van die spesifieke plante wat as nektarbronne gedien het. Die metode is egter tydrowend en moet deur 'n deskundige uitgevoer word (Louveaux, et al., 1978). As gevolg hiervan en weens die koste van die metode is daar nie tans in Suid-Afrika geakkrediteerde laboratoriums wat gestandaardiseerde metodes implementeer om hierdie toets uit te voer nie. Dit is dus nie moontlik om die nektarbronne van heuning in Suid-Afrika op 'n gereelde basis te bevestig nie.

Etikette wat in die Europese Unie (EU) gebruik word om voedselprodukte met spesifieke geografiese oorsprong te erken, is beskermde geografiese aanduiding (BGA), beskermde oorsprongbenaming (BOB) en tradisionele spesialiteit gewaarsborg (TSG). BGA geld in die geval van voedsel- en landbouprodukte wat gedurende ten minste een stadium van produksie gekoppel is aan 'n geografiese gebied. BOB verwys na voedsel- en landbouprodukte wat tydens elke stap van die produksie verband hou met 'n geografiese gebied. Laastens word TSG gebruik om voedsel- en landbouprodukte wat tradisionele grondstowwe bevat, tradisionele produksiemetodes volg of 'n tradisionele chemiese samestelling het wat nie met die geografiese oorsprong van die produk verband hou nie, aan te dui (Dias & Mendes, 2018). Wanneer voedselprodukte as TSG gesertifiseer en met die gepaardgaande etikette gemerk is, bevoordeel dit die persentasie wins daarop asook kompetisie op die internasionale mark. Om BGA-etikette te kan aanbring, moet die oorsprong van die produkte bepaal kan word om sertifisering te kan bekom. Weens die koste van analitiese ontledings word die sertifisering in die reël aan 'n noukeurige naspourbaarheidstelsel gekoppel. Met 'n aanvaarbare akkuraatheid kan naby-infrarooi spektroskopie gebruik word as ondersteunende siftingsmetode vir die duur, tydrowende analitiese tegnieke wat dikwels ook deur hoog opgeleide tegnici uitgevoer moet word.

Verskeie internasionale studies, wat naby-infrarooi (NIR) spektroskopie in kombinasie met meerveranderlike data-analise as toetsmetode ondersoek het, het gevind dat dit wel moontlik is om die stuifmeelinhoud (Escuredo, et al., 2015), botaniese oorsprong (Bisutti, et al., 2019; Chen, et al., 2012; Tahir, et al., 2016) en geografiese ligging (Latorre, et al., 2013; Woodcock, et al., 2007) van heuning te voorspel. Deur gebruik te maak van sagte onafhanklike modellering van klasanalogie (SIMCA) en parsieël kleinste kwadrate-diskriminantanalise (PLS-DA) met verskeie voorverwerkings van die spektra (geskandeer in transflektansiemodus) het Woodcock et al. (2007) klassifi-

kasiekoerse van tussen 94 en 96% met SIMCA vir ongefiltreerde monsters (n=167) verkry uit drie verskillende lande (Ierland, Mexiko en Spanje), terwyl 100% korrekte klassifikasies met PLS-DA vir elk van die heuningklasse verkry is. In die geval van gefiltreerde heuning (n=125) (van Argentinië, die Tsjeggiese Republiek, Hongarye en Ierland) het die beste SIMCA-modelle korrekte klassifikasiekoerse van tussen 92 en 100% gelewer, terwyl PLS-DA korrekte klassifikasies van tussen 96 en 100% vir dieselfde monsters gelewer het. In 'n soortgelyke studie (Latorre et al., 2013) is SIMCA en PLS-DA toegepas op heuningspektra wat vooraf behandel is met standaardnormaalvariaat (SNV) om heuning met BGA van kommersiële heuningmonsters van ander oorsprong te onderskei. Die beste BGA-model is bereik met SIMCA met 'n sensitiwiteit van 93,3% en spesifisiteit van 100%.

Aangesien heuningtipes verskil, sal die reeds ontwikkelde NIR spektroskopie-voorspellingsmodelle nie op Suid-Afrikaanse heuning van toepassing wees nie. 'n Studie is in Suid-Afrika uitgevoer waartydens NIR spektroskopie ondersoek is as 'n moontlike metode om die egtheid van Suid-Afrikaanse heuning te bepaal (Guelpa, et al., 2017). 'n Opvolgstudie deur Rust et al. (2021) het enkele moontlike tekortkominge, byvoorbeeld opberging van monsters soos beïnvloed deur tyd en temperatuur, waargeneem en daarop verbeter. Aangesien suiwer heuningmonsters geneig is om te kristalliseer, is dit dikwels nodig om die monsters vir tot 24 uur by 'n verhoogde temperatuur te inkubeer om die kristalle op te los voordat dit met die spektrofotometer geskandeer word. Chemiese veranderinge wat die spektrale data beïnvloed, kon moontlik in die heuningmonsters plaasgevind het by die inkubasiestemperatuur (55 °C) wat deur Guelpa et al. (2017) gebruik is. In teenstelling met Guelpa et al. (2017) het Rust et al. (2021) slegs vars geoesde monsters gebruik wat by die aanvang van die studie minder as ses maande oud was. Rust et al. (2021) het monsters uit heuningkorwe versamel en vir twee tot drie uur voor ontleding by 'n maksimum temperatuur van 40 °C geïnkubeer, wat voldoende was om die suikerkristalle op te los en die heuning vloeibaar te maak. Laasgenoemde outeurs, sowel as Guelpa et al. (2021), het die totale oplosbare vaste stowwe (TSS) gestandaardiseer tot 70 °Brix; dit was 'n verbetering op 'n aantal studies wat geen °Brix-standaardisering toegepas het nie (Bázár, et al., 2016; Bisutti, et al., 2019; Chen, et al., 2012; Escuredo, et al., 2013, 2015; Latorre, et al., 2013; Leme, et al., 2018; Tahir, et al., 2016). Hierdie standaardisering sluit vogvariasie tussen monsters uit wat voorkom dat klassifikasie nie op grond van hoofsaaklik vogverskille gedoen word nie.

Die studies deur Guelpa et al. (2017) en Rust et al. (2021) het nie die gebruik van NIR spektroskopie om geografiese ligging van Suid-Afrikaanse heunings te voorspel, ondersoek nie. Die doel van die hierdie studie was om NIR spektroskopie, tesame met parsieël kleinste kwadrate diskriminantanalise (PLS-DA) en toepassing van verskillende voorverwerkingstegnieke op die rou spektrale data te evalueer vir die klassifikasie van geografiese oorsprong van Suid-Afrikaanse heuning van drie geografiese omgewings.

Materiale en metodes

Heuningmonsters

Egte, ongeprosesseerde heuningmonsters is direk van Suid-Afrikaanse byeboere oor twee seisoene (November 2018 en Maart 2019) uit dieselfde korwe ingesamel. Bye kan tot 10 km van hulle korwe af vlieg, na gelang van die beskikbare nektarbronne rondom die korf. Dit is ag geneem, is heuningmonsters van drie Wes-Kaapse liggings (meer as 130 km uit mekaar en met verskille in beskikbare nektarbronne) geoes, naamlik Hopefield (Weskus; 8 persele), Stanford (Overberg; 3 persele) en Stellenbosch (Kaapse Wynlande; 7 persele). By elke perseel is monsters uit verskeie korwe geoes. Die plantegroei wat in die drie omgewings voorgekom het, is aangeteken (Tabel 1). Nadat die heuning ingesamel is, is die waslaag van die heuningkoek verwyder en laasgenoemde teen 'n skuinste gelaat sodat die heuning daaruit kon drup. Daarna is die heuning deur 'n 420 µm-maas gefiltreer om was, bye en onsuiverhede te verwyder. Alle monsters is so gou moontlik nadat dit geoes is, ontleed, en geen monsters was by die aanvang van die analise vir langer as 6 maande opgeberg nie. Die monsters is in 'n yskas (4 °C) geberg tot analise om enige irrelevante bronne van variasie te minimaliseer en veranderinge wat nie verantwoord kan word nie te voorkom. Die heuning is 12 uur voor analise in 'n inkubator (40 °C) geplaas om die heuning vloeibaar te maak en enige suikerkristalle wat teenwoordig was, op te los.

Aanpassing van totale oplosbare vastestowwe

Die totale oplosbare vaste stowwe (TOVS) is gemeet as °Brix met 'n Atago-refraktometer (Atago, Japan) en tot 70 °Brix aangepas. Die hoeveelheid dH₂O wat bygevoeg is, is bereken na aanleiding van vergelykings 1 en 2.

$$M_{\text{heuning}} = \frac{B_{\text{einde}} M_{\text{einde}}}{B_{\text{heuning}}} \quad (1)$$

$$M_{\text{water}} = M_{\text{einde}} \cdot M_{\text{heuning}} \quad (2)$$

waar:

M = massa in g;

B = TOVS in °Brix;

M_{heuning} = massa heuning bepaal vir elke monster wat aangepas moet word;

B_{einde} = voorafbepaalde TOVS (70 °Brix) tot waar heuningmonsters aangepas moet word;

M_{einde} = voorafbepaalde massa (20 g) ná aanpassing van TOVS;

B_{heuning} = TOVS (°Brix) van oorspronklike heuningmonster; en

M_{water} = massa water wat by M_{heuning} , soos bepaal vir elke monster bygevoeg moet word.

Naby-infrarooi spektrale analise

Die heuningmonsters is met twee verskillende NIR spektrofotometers geskandeer. Transfleksiespektra (1 mm-padlengte) is met 'n BÜCHI NIRFlex-500 Fourier-transformasie-naby-infrarooi (FTNIR) laboratoriumspektrofotometer (BÜCHI Labortechnik GmbH, Flawil, Switserland) in die golflengtegebied 1 100 tot 2 500 nm verkry (Figuur 1a). Vir elke meting is gemiddeld 32 skanderings geneem. 'n Spektrale interval van 0,4 nm en resolusie van 16 cm⁻¹ (halwe resolusie) is gebruik. Elke monster is drie keer in 'n roterende glas petribakkie (Duran Group, Mainz, Tsjeggiese Republiek) sonder enige herverpakking tussen skanderings) geskandeer. Slegs genoeg heuning om die helfte van die oppervlakte van die petribakkie te bedek, is gebruik. Die transfleksieplaat is op so 'n wyse (diagonaal) op die petribakkie met heuning geplaas dat dit die vorming van lugborrels voorkom het. Difussierefleksiespektra is ook met 'n mobiele MicroNIR 1700-handspektrofotometer (JDSU Corporation, Santa Rosa, CA, USA) met 'n geïntegreerde vakuumwolframlamp verkry. Die monsters is in 'n wit Teflonhouer (10 mm-radius x 10 mm-diepte) (Figuur 1b) sowel as in 'n glasflesie (1 mm-padlengte) (Figuur 1c) in die golflengtegebied 950 tot 1 650 nm (golflengtespesifieke resolusie van 1% van die golflengte) ontleed. Daarby het die instrument 100 skanderings per analise uitgevoer. Die heuningmonster is in

Tabel 1: Geografiese oorsprong van heuningmonsters, tyd van insameling, tipes plantegroei in die omgewings en getal monsters tydens elke seisoen ingesamel

Plek van insameling	Tyd van insameling	Plantegroei	Getal monsters versamel
Hopefield	November 2018	Sandveld-fynbos (wasbossies, penbossies en proteas), etlike grys- en swartysterbas-bloekoms	8
	Maart 2019	Sandveld-fynbos (erikas), etlike suikerbloekoms	8
Stanford	November 2018	Elim-ferrikrete*- en Strandveld-fynbos (wasbossies, penbossies en proteas), Australiese mirt en baie suikerbloekoms en bloekoms	5
	Maart 2019	Etlike Elim-ferrikretefynbos (erikas) en baie suikerbloekoms en bloekoms	9
Stellenbosch	November 2018	Boland-graniet-, Swartland-graniet- en Swartland-alluviumfynbos (wasbossies, penbossies en proteas), swartysterbas-bloekoms, salignabloekoms, rooiblombloekoms, ramenas, blou-echium, wilde-ertjie (<i>vetch</i>) en gousblom	13
	Maart 2019	Boland-graniet-, Swartland-graniet- en Swartland-alluviumfynbos (erikas), suiker- en rooibloekoms	13

*Hoewel Elim-ferrikreetfynbos as kritiek bedreig geklassifiseer is, is dit steeds in die Stanford geografiese gebied teenwoordig.

die Teflonkoppie gegiet, liggies teen die blad van die bank getik en vir 5 min ongesteurd gelaat terwyl die borrels na die oppervlak van die heuning beweeg het. Daarna is die borrels met 'n spatel afgeskep. Die monster is drie maal geskander met 'n rotasie van die Teflonkoppie tussen skanderings. Die glasflessie is kwartvol met heuning gevul, waarna 'n Teflonstaaf in die flessie geplaas is om 'n padlengte van 1 mm te verkry. Die glasflessie is in die flessiehouer geplaas en 90° tussen skanderings gedraai. Tot 10 herskanderings met die glasflessie was nodig weens beweging van die Teflonstaaf tydens herposisionering.

Spektrale en meerveranderlike data-analise

Spektrale en meerveranderlike data-analise is met The Unscrambler X 10.5 (Camo Software AS, Oslo, Noorweë), MATLAB (R2015b, The Mathworks, Natick, MA) en PLS_Toolbox (Solo, Eigenvector Research Inc., Wenatchee, WA) gedoen. Klassifikasie Modelle met PLS-DA (Wold, et al., 1984) is ontwikkel om tussen heunings van Stellenbosch, Hopefield en Stanford te onderskei. Spektre is met verskeie kombinasies van voorverwerkingstegnieke geprosesseer, naamlik gladstryking met die Savitzky-Golay (SG)-metode (Savitzky & Golay, 1964), afgeleides, multiplikatiewe spreidingskorreksie (MSC) (Geladi, et al., 1985), standardnormaalvariaat (SNV) en tendensverwydering (DT) (Barnes, et al., 1989), asook eksterneparameter-ortogonalisering (EPO) (Roger, et al., 2003; Sharma, et al., 2014). Die tegnieke is in die volgende kombinasies getoets: SNV; SNV + DT; DT; SG (2^e afgeleide, 3^e polinoom, 11 gladstrykingspunte); EPO (5 PC); EPO (5 PC) + SNV; EPO (5 PC) + SNV + DT; EPO (5 PC) + DT; EPO (5 PC) + SG (2^e afgeleide, 3^e polinoom, 11 gladstrykingspunte).

Nadat die data voorverwerk is, is dit met óf die Kennard-Stone (KS) óf die DUPLEX-algoritme in afsonderlike kalibrasie- (70%)

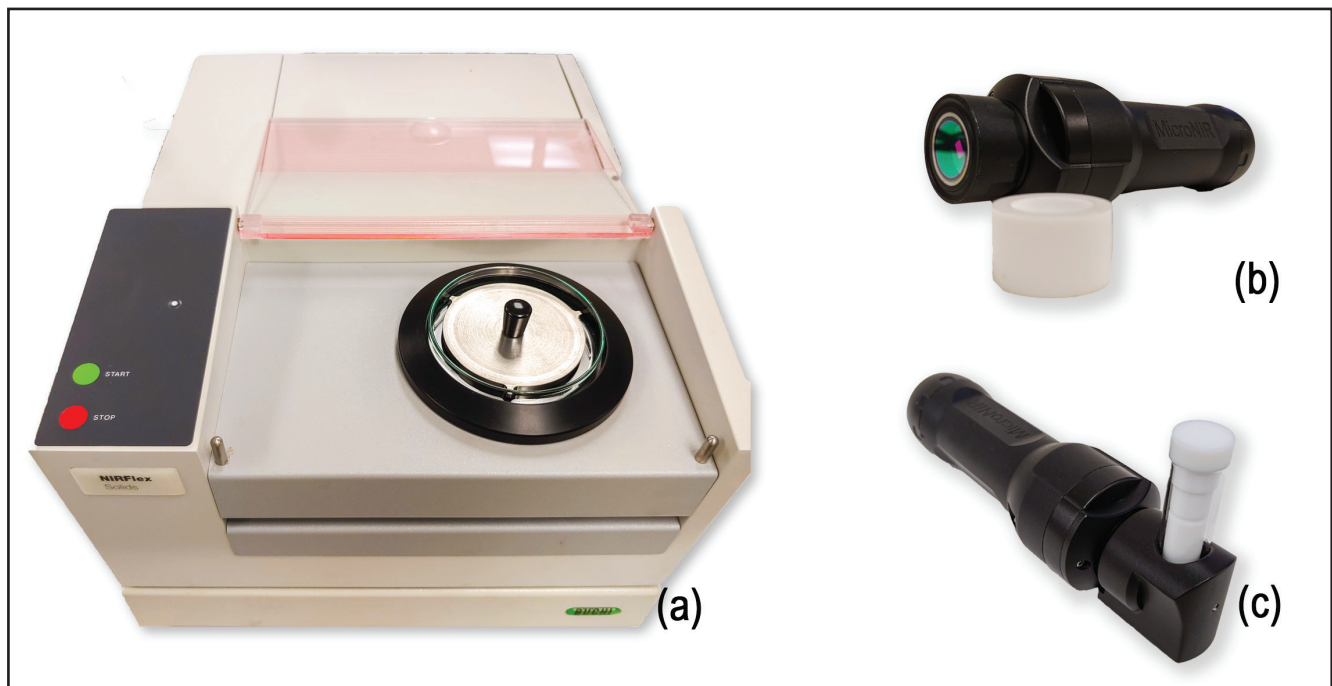
en validasiestelle (30%) verdeel. Die KS-algoritme laat toe om monsters te kies met 'n eenvormige verspreiding oor die voorspellerruimte heen (Kennard & Stone, 1969). Die proses begin deur twee punte te kies wat die verste van mekaar is. Hierdie punte word aan die kalibrasiestel toegewys en van die lys punte verwyder. Daarna word die oorblywende punte aan die kalibrasiestel toegewys deur die afstand tussen die punte wat nog nie toegeken is nie en die reeds gekose punte te maksimeer. Hierdie proses duur voort tot die voorafbepaalde getal punte verkry is. Die afstandmetriek wat deur die prosedure gebruik word, is die Euklidiese afstand (sien Gower, 1985).

DUPLEX is soortgelyk aan KS, maar verskil aangesien dit nie net vir die kalibrasiestel punte kies nie, maar ook vir die validasiestel (Snee, 1977). Dit begin soortgelyk aan die KS-algoritme deur twee punte wat die verste van mekaar is te kies, en dit aan die kalibrasiestel toe te wys. Hierdie twee punte word dan van die lys punte verwyder. Daarna kies dit die volgende twee punte wat die verste van mekaar is en ken dit aan die validasiestel toe – weereens word die twee punte van die lys punte verwyder. Die prosedure gaan voort om die oorblywende punte alternatiewelik aan die kalibrasie- en validasiestelle toe te ken, gebaseer op die afstand na die punte wat reeds gekies is.

Klassifikasie-akkuraatheid, sensitiwiteit en spesifisiteit is volgens vergelykings 3, 4 en 5 onderskeidelik bereken. Laastens is die wanklassifikasie volgens vergelyking 6 bereken om te bepaal hoe gereeld die klassifikasie verkeerd was.

$$\text{Klassifikasie-akkuraatheid (\%)} = \frac{TP+TN}{(TP+TN+FP+FN)} \times 100\% \quad (3)$$

$$\text{Sensiwiteit (\%)} = \frac{TP}{(TP+FN)} \times 100\% \quad (4)$$



Figuur 1: (a) Die BÜCHI NIRFlex-500 Fourier-transformasie-naby-infrarooi laboratoriumspektrofotometer met glaspetribakkie en transleksieplaat. MicroNIR 1700-handspektrofotometer met (b) Teflonkoppie en (c) glasflessie met Teflonstaaf.

$$\text{Spesifisiteit (\%)} = \frac{\text{TN}}{(\text{TN} + \text{FP})} \times 100\% \quad (5)$$

$$\text{Wanklassifikasie (\%)} = \frac{\text{FP} + \text{FN}}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \times 100\% \quad (6)$$

waar:

ware positiewe (TP) = positiewe uitkoms korrek as positief geklassifiseer;

ware negatiewe (TN) = negatiewe uitkoms korrek as negatief geklassifiseer;

vals positiewe (FP) = negatiewe uitkoms verkeerdlik as positief geklassifiseer; en

vals negatiewe (FN) = positiewe uitkoms verkeerdlik as negatief geklassifiseer.

Sensitiwiteit staan bekend as die ware positiewe koers wat aandui hoe sensitief die model is om elke keer die ware positiewe uitkoms korrek as positief te klassifiseer. Dit beskryf die doeltreffendheid van die model (Parikh, et al., 2008). Eweneens staan spesifisiteit bekend as die ware negatiewe koers. **Spesifisiteit** dui aan hoe gereeld die model die negatiewe uitkoms korrek as negatief klassifiseer het en hoe gereeld die voorspelling korrek was (Parikh, et al., 2008).

Resultate en bespreking

Voorverwerkingstegnieke

NIR spektra is ryk aan inligting met betrekking tot die fisies-chemiese samestelling van monsters asook die verskille tussen monsters. Hierdie fisies-chemiese samestelling kan egter nie visueel uit die spektra waargeneem word nie (Bokobza, 1998; Manley & Baeten, 2018). Dit is as gevolg van die breë NIR absorpsiebande wat gevorm word wanneer individuele pieke oorvleuel. Effektiewe ontleding van sodanige data is egter moontlik met behulp van meerveranderlike data-analisetegnieke. Die eerste stap behels toepassing van voorverwerkingstegnieke op die onverwerkte data om oorbodige inligting in die spektra wat nie tot die samestelling van die heuningmonsters bydra nie en wat dus geen chemiese inligting byvoeg nie te verwyder. Hierdie proses berei die data voor vir die volgende stap, naamlik ontwikkeling van klassifikasie modelle. Voorverwerkingstegnieke wat meestal gebruik en in die literatuur beskryf word, is gladstryking (Savitzky & Golay, 1964), eerste of tweede afgeleides, multiplikatiewe spreidingskorreksie (MSC) (Geladi, et al., 1985), standaardnormaalvariaat (SNV) (Barnes, et al., 1989) en tendensverwydering (DT). Gladstryking word gebruik om spektrale geraas te verminder (Heise & Winzen, 2006). Die algemeenste vorm van gladstryking is die Savitzky-Golay-metode (Savitzky & Golay, 1964). Multiplikatiewe spreiding as gevolg van ligverstrooiing vind dikwels plaas tydens diffusierefleksie en transfleksie wanneer nie net die geabsorbeerde lig nie maar ook die verspreide lig deur die detektor waargeneem word (Manley & Baeten, 2018). MSC kan die effek van ligverstrooiing verwyder (Geladi, et al., 1985). SNV word dikwels in kombinasie met DT gebruik (Barnes, et al., 1989) vir die regstelling van basislynverskuiwings wanneer diffusierefleksie gebruik word (Duckworth, 2004). Soms is daar eksterne invloede wat effektief met behulp van eksterne-parameter-ortogonalisering (EPO) verwyder kan word (Roger, et

al., 2003; Sharma, et al., 2014). Die meeste voorverwerkingstegnieke streef daarna om die residuele matriks (R) van die X-matriks (vergelyking 7) te verwyder.

$$X = XP + XQ + R \quad (7)$$

waar:

X = matriks;

P = projeksie van die komponent van belang;

Q = projeksie van die eksterne parameter; en

R = residuele matriks.

Met EPO, 'n voorverwerkingstegniek wat onlangs ontwikkel is, word die spektra in twee dele verdeel, naamlik die nuttige komponent wat 'n direkte verhouding met die reaksieveranderlike (P) het, en die komponent wat van 'n eksterne invloed afkomstig is, naamlik Q. Eksterne invloede sal byvoorbeeld in die geval van heuning die volgende insluit: temperatuurkommelings van die monster self, verskille in TOVS, die oorsprong van die monster en die oeseisoen. Deur die spektra op die ortogonaal van die spasie te projekteer kan veranderinge wat deur eksterne invloede geïnduseer is, waargeneem word. Omdat daar verskeie faktore is wat Q kan beïnvloed, is die invloed daarvan nie altyd bekend nie, maar met EPO is die inligting nie nodig om hierdie voorverwerkingstegniek toe te pas nie. Om EPO prakties uit te voer, kan hoofkomponentanalise (HKA) toegepas word om Q te beraam. Deur 'n HKA te doen kan die getal EPO-komponente bepaal word en daarna kan die getal latente veranderlikes (LV's) bepaal word. Voordele van EPO is dat die eksterne faktore nie konstant hoef te wees nie; slegs 'n klein dataset word benodig, en dit kan die model meer robuus maak. Met die verskillende getalle hoofkomponente (HK's) wat met behulp van EPO verkry is, is daar gevind dat die beste resultate met vyf HK's verkry is. Dieselfde geld vir tweede afgeleides met die SG-metode (derde veelterm; 11 gladstrykingspunte). EPO is gebruik om 'n Q-matriks uit die kalibrasiedata te ontwikkel – in dié geval om die verskille wat in die spektra teenwoordig is weens die twee seisoene (November 2018 en Maart 2019) te verwyder.

Hoofkomponentanalise

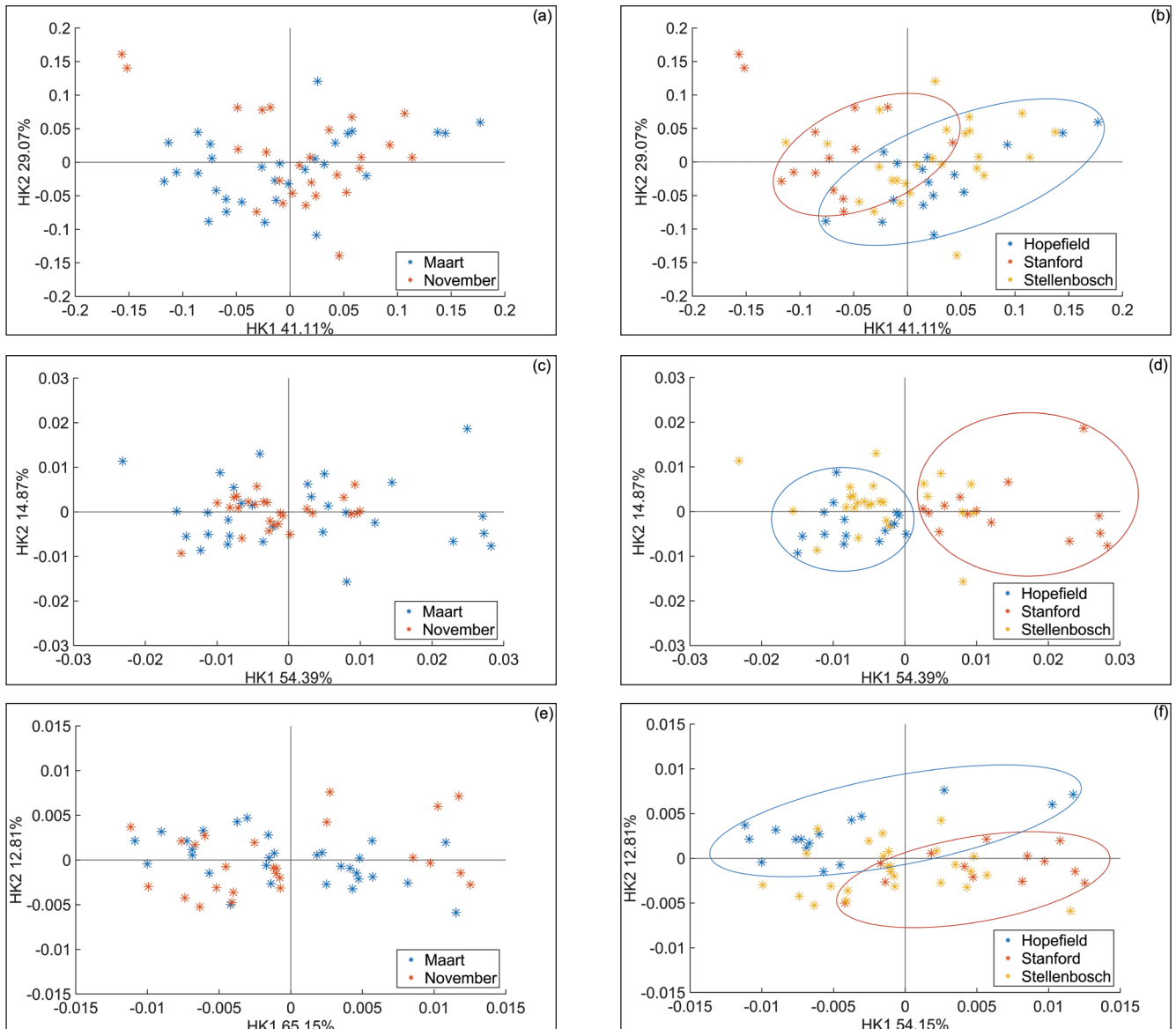
Om die dimensionaliteit van NIR spektroskopiedata te verminder, word hoofsaaklik meerveranderlike kwalitatiewe onbegeleide data-analises, soos HKA en trosanalise (Cowe & McNicol, 1985), gebruik. Dit word gedoen deur die aantal veranderlikes te verminder terwyl die meeste van die variansie in die dataset behou word. HKA gebruik die inligting van die oorspronklike dataset om hoofkomponente (HK) te skep wat met 'n tweedimensionele of driedimensionele tellinggrafiek gevisualiseer kan word (Manley, 2014). Die eerste HK is die veranderlike wat die grootste deel van die variansie beskryf, die tweede HK sal die tweede grootste deel van die variansie beskryf, en so aan. Die resultate van HKA word dikwels gebruik om verdere analises te lei (Szymańska, et al., 2015). Trosanalise komplementeer HKA (Callao & Ruisánchez, 2018). Groepe word gevorm wanneer daar ooreenkomste is tussen die spektra van die monsters wat ondersoek word. Dit laat onder andere ook toe om uitskieters te identifiseer (Szymańska, et al., 2015).

Metings met NIR spektroskopie en die daaropvolgende voorspellingsmodelontwikkeling behels die analise van 'n monster waarop geen voorafbereiding toegepas is nie. Alle variasie wat moontlik kan voorkom in monsters wat in die toekoms met die model voorspel moet word, moet dus in die datastel ingesluit word. Daarom is monsters oor twee seisoene vanuit dieselfde korwe versamel. Sodoende is variasie as gevolg van omgewingstoestande in die datastel ingesluit. So ook is verdere variasie ingesluit deur monsters uit korwe op verskillende persele binne elke geografiese omgewing te versamel.

Uit die tipes plantegroei wat vir elk van die geografiese omgewings (Tabel 1) waargeneem is, kon die blomplante wat deur die bye as nektarbronne gebruik is vir elke omgewing geïdentifiseer word. Dit was duidelik dat Sandveld-fynbos en *Eukalptus*-bome hoofsaaklik onderskeidelik in die Hopefield- en Stanford-omgewings waargeneem is, terwyl beide in die Stellenbosch-omgewing voorgekom het.

Die HKA-tellinggrafieke van die datastelle, nadat EPO toegepas is, soos verkry met die laboratoriuminstrument (Figuur 2a en b), handapparaat met Teflonkoppie (Figuur 2c en d) en handapparaat met glasflesie (Figuur 2e en f) dui die kwalitatiewe onderskeid van die heuningmonsters aan wat oor die twee seisoene uit die drie geografiese omgewings Stellenbosch, Hopefield en Stanford verkry is.

Geen onderskeid is waargeneem tussen monsters van die twee verskillende seisoene nie vir al drie instrumente (Figuur 2a, 2c en 2e). Wat geografiese omgewing betref, is daar met die laboratoriuminstrument (Figuur 2b) 'n mate van onderskeid tussen die heuningmonsters van die Hopefield- en Stanford-omgewings waargeneem, maar dié van Stellenbosch toon geen onderskeid van die monsters van die Hopefield- en Stanford-omgewing nie. Die onderskeid van die heunings van Hopefield en Stanford is in die rigting van beide HK1 en HK2. Metings met die handapparaat met Teflonkoppie (Figuur 2d) het gelei tot 'n



Figuur 2: HKA-tellinggrafieke van HK1 vs. HK2 van die datastelle, nadat EPO toegepas is, soos verkry met die (a) en (b) laboratoriuminstrument, (c) en (d) handapparaat met Teflonkoppie en (e) en (f) handapparaat met glasflesie. HKA-tellinggrafiek (a), (c) en (e) dui die afwesigheid van onderskeid tussen seisoene aan, terwyl grafiek (b), (d) en (f) onderskeid tussen die drie geografiese oorspronge aandui.

duidelike onderskeid tussen die heunings van Hopefield en Stanford, naamlik in die rigting van HK1. Ná die toepassing van EPO was een hoofkomponent met laasgenoemde apparaat voldoende om tussen die klasse te onderskei. Weereens kon geen onderskeid waargeneem word tussen die heunings van Stellenbosch en dié van Hopefield- en Stanford-omgewing nie. Duidelike onderskeid is waargeneem vir die metings wat met die handapparaat met glasflessie (Figuur 2f) gedoen is, maar in dié geval was dit in die rigting van HK2. In dié geval moet die tweede HK ook in die model ingesluit word om skeiding tussen die klasse te bewerkstellig.

Die voorkoms van spesifieke nektarbronne, naamlik beide fynbos- en *Eukalptus*-spesies in die omgewing van Stellenbosch, fynbos in die Hopefield-omgewing en *Eukalptus* in die Stanford-omgewing, is weerspieël in die onderskeid van die heuningmonsters soos waargeneem in die HKA-tellinggrafieke. Die voorkoms van die twee verskillende nektarbronne in die

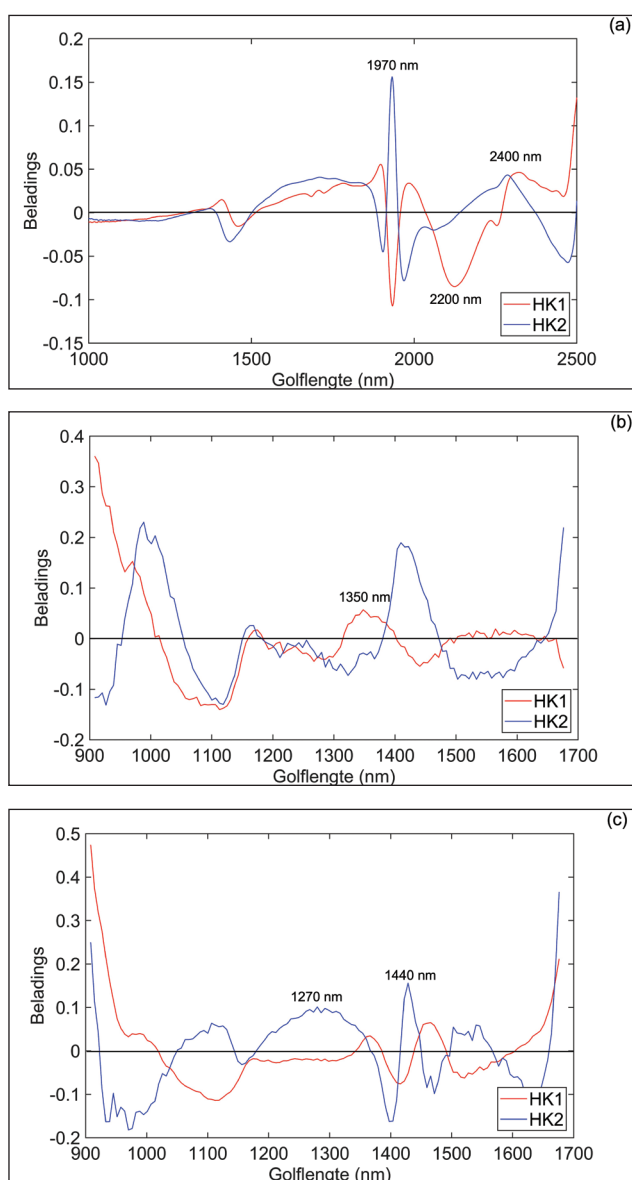
Hopefield- en Stanford-omgewing onderskeidelik het tot gevolg gehad dat 'n duidelike onderskeid tussen heunings van dié twee omgewings waargeneem is. Die voorkoms van beide nektarbronne in die Stellenbosch-omgewing verklaar waarom heunings uit die Stellenbosch-omgewing met dié van beide Hopefield (fynbos) en Stanford (*Eukalptus*) in die tellinggrafieke oorvleuel het. Die ondersoek van hoërde-hoofkomponente het nie tot enige verdere onderskeid van die heuningmonsters van die Stellenbosch-omgewing van dié van die ander twee geografiese omgewings bygedra nie.

Interpretasie van HKA beladingswaardes is belangrik om die intraklasvariasie, met ander woorde die variasie wat aanleiding gee tot die onderskeid tussen die geografiese omgewings, van die heuningmonsters te beskryf. 'n Positiewe HKA-telling gaan gepaard met positiewe beladings. Dus sal die tellings in die positiewe kwadrante van die tellinggrafiek 'n hoë spektrale reaksie vir die veranderlikes (golflengtes) aan die positiewe kant van die beladingsgrafiek toon – en omgekeerd vir die negatiewe tellings en beladings. Die onderskeid tussen die verskillende geografiese omgewings (intraklasvariasie), soos waargeneem met die drie verskillende instrumente, kan dus beskryf word (Figuur 3). In die geval van die laboratoriuminstrument (Figuur 3a) dra beide HK1 en HK2 tot die onderskeid by. Die golflengtebande 1 970 nm, 2 200 nm en 2 400 nm word onderskeidelik geassosieer met OH-kombinasies (polisakkariede), CH- en CO-kombinasies (koolhidrate) en CH-aromatiese bindings (Williams, et al., 2019). Vir die handapparaat (Teflonkoppie) word die golflengtes 1 300 nm tot 1 400 nm (HK1) met CH-aromatiese bindings (Figuur 3b) geassosieer. Hoofkomponent 2 het in dié geval nie bygedra tot die onderskeid tussen die geografiese omgewings nie. In die geval van die handapparaat (glasflessie) (Figuur 3c) word 1 270 nm van HK2 geassosieer met CH-kombinasies soos ook waargeneem deur Leme, et al. (2018) en 1 440 nm met OH-kombinasies (1e oortoon; suikers) (Mouazen & Al-Walaan, 2014). Die onderskeid tussen die geografiese oorspronge kan dus nie aan 'n enkele komponent gekoppel word nie, maar dui op verskille in al die makrokomponente van die heuningmonsters.

Klassifikasievoorspellingsmodelle

Tydens die ontwikkeling van klassifikasievoorspellingsmodelle is inligting rakende die klasse of groeperings bekend. Lineêre diskriminantanalise (LDA) beperk die variansie in die klasse tot 'n minimum, maar maksimeer die variansie tussen die klasse. LDA word gebruik om tussen klasse te onderskei, en dit word gedoen deur die afstand van byvoorbeeld die heuningmonster na die middel van 'n klas te bereken en die monster aan die naaste klas toe te ken. LDA het soms kolineariteitsprobleme, maar PLS-DA kan hierdie probleem oorkom. In teenstelling hiermee laat PLS-DA toe, deur die voorspeller matriks se dimensionaliteit met projeksie op latente veranderlikes te verminder, om die beperkings wat vir LDA bespreek is te oorkom.

'n Oorsig van die resultate van die PLS-DA-voorspellingsmodel verkry deur die laboratoriuminstrument (model 1 tot 4) sowel as die handapparaat met beide monster-aanbiedingsmetodes (Teflonkoppie – model 5 tot 8; glasflessie – model 9 tot 12) word onderskeidelik in Tabel 2 tot 4 weergegee. Belowende resultate



Figuur 3: Belading-lyngrafieke van HK1 en HK2 vir die datastelle soos verkry met (a) die laboratoriuminstrument, (b) die handapparaat (Teflonkoppie) en (c) die handapparaat (glasflessie).

is met sommige van die modelle verkry. Met die eerste oogopslag blyk model 2d (DUPLEX (EPO + SG)) (Tabel 2) die beste te wees met 'n voorspellingsakkuraatheid van 75%, maar wanneer die kalibrasie-akkuraatheid van 92,5% in ag geneem word, is dit duidelik dat die model 'n oormatige passing is. Model 4 (KS (EPO)) en 4c (EPO + DT) blyk meer robuuste modelle te wees, hoewel daar ook 'n mate van oormatige passing is (soos aangedui deur die groot verskil tussen kalibrasie- en voorspellingsakkuraatheid).

Die toepassing van EPO as voorverwerkingstegniek het slegs sommige van die modelle van die laboratoriuminstrument verbeter (Tabel 2). In die geval van model 2b (DUPLEX (EPO+SNV+DT)) het die insluiting van EPO die kalibrasie-akkuraatheid van 65% tot 87,5% verbeter, maar die voorspellingsakkuraatheid onveranderd gelaat. Ten spyte van die goeie klassifikasieresultate was daar steeds 'n groot verskil in die kalibrasie- en voorspellingsakkuraatheid en was die model dus 'n oormatige passing. In die geval van die modelle waar KS gebruik is, is die effek van EPO ook duidelik. In die geval van model 4b (KS (EPO+SNV+DT)), het die byvoeging van EPO egter tot 'n model met verlaagde kalibrasie- en voorspellingsakkuraatheid gelei. Die swakste klassifikasie-model was 3d (KS (SG)) met 'n kalibrasie-akkuraatheid van 70,27% en 'n baie lae voorspellingsakkuraatheid van 26,32%, wat daarop dui dat die model weereens 'n oormatige passing was.

In Tabel 3, wat die klassifikasie-akkuraatheid van die handapparaat (Teflonkoppie) weergee, is die effek van die gebruik van óf KS óf DUPLEX duidelik wanneer dieselfde voorverwerkingstegnieke gebruik is. Die gebruik van DT het tot die swakste model (7c) gelei met die data wat met KS geskei is. Dit het die beste model (5c) gelever met die data wat met DUPLEX geskei is. Model 7c (KS (DT)) het 'n lae voorspellingsakkuraatheid van 42,11% gehad, in teenstelling met model 5c (DUPLEX (DT)), wat 'n voorspellingsakkuraatheid van 87,5% gehad het. Hoewel laasgenoemde die beste model was, dui die swak kalibrasie-akkuraatheid van 67,50% op 'n mate van oormatige passing. Model 6d (DUPLEX (EPO+SG)) is 'n oormatige passing, maar het wel die beste model vir hierdie apparaat gelever.

By die handapparaat (glasflessie) het KS beter modelle gelever, veral met die byvoeging van EPO (Tabel 4). Oor die algemeen het die meeste van die modelle met DUPLEX en KS goeie kalibrasie-akkuraatheid behaal, met die uitsondering van model 9c (DUPLEX (DT)) wat oor die algemeen, saam met model 1d (DUPLEX (SG)), die swakste modelle gelever het. Weereens was daar 'n tendens, soos in die geval van die handapparaat (Teflonkoppie) waar die swakste modelle van DUPLEX (model 9c (DUPLEX (DT)) en 1d (DUPLEX (SG))) ook swak modelle met KS (model 11c (KS (DT)) en 11d (KS (SG))) gelever het – met dieselfde voorverwerkingstegnieke. Die beste model (12a (KS (EPO+SNV))) het EPO en SNV as voorverwerkingstegniek gehad, wat 'n uitstekende kalibrasie-akkuraatheid van 94,59% gelever het, en 'n voorspellingsakkuraatheid van 84,21%. Dit was die beste model van die drie datastelle soos vervat in Tabel 2, 3 en 4. Die model het drasties verbeter (van model 11a (KS (SNV))) deur EPO by te voeg, wat die suksesvolle bydrae van EPO, soos vroeër gewys is (Preys, et al., 2008), bevestig. In 'n vorige studie wat heuning suksesvol volgens geografiese oorsprong geklassifiseer

het, is heuning van verskillende lande gebruik (Woodcock, et al., 2007). Behalwe die chemiese samestelling word heuning van verskillende liggings ook deur klimaat, weerpatrone, grond, soort plant en seisoen (wanneer heuning uit die korf gehaal word) beïnvloed. Die invloed van die nektarbron is in bostaande studie, wat ook 'n groter datastel gehad het (n=292), deur die gebruik van multiblom-heuningmonsters beperk.

'n Algemene patroon is met die toepassing van KS opgemerk, aangesien dit 'n beter dataskedingsmetode vir die laboratoriuminstrument en die handapparaat (glasflessie) was, terwyl DUPLEX vir die handapparaat (Teflonkoppie) die beste was. Hoewel hierdie spesifieke vergelyking nie in die literatuur van heuningnavorsing voorkom nie, kan daar steeds afleidings gemaak word. Die ooreenkoms tussen die laboratoriuminstrument en die handapparaat (glasflessie) is dat die NIR lig eers deur 'n laag glas beweeg voordat dit deur die heuningmonster beweeg, met 'n padlengte van 1 mm vir beide apparate. Dit is teenoor die handapparaat met Teflonkoppie, wat 10 mm diep is, waar die NIR lig dadelik die heuningmonster binnedring met geen glas tussenin nie. Die glans van die glaslaag kan ligverstrooiing laat plaasvind (Workman & Shenk, 2004). Volgens Beer-Lambert se wet is daar 'n direkte en liniêre verhouding tussen die monsterkonsentrasie, padlengte en die absorpsie van lig by 'n spesifieke golflengte (Duckworth, 2004). Dit is dus moontlik dat die teenwoordigheid van die glaslaag die spektrale data op so 'n wyse beïnvloed dat data verkry met die laboratoriuminstrument en handapparaat (Teflonkoppie) uit dieselfde metode (DUPLEX) voordeel trek.

Hoewel die geografiese oorsprong volgens die resultate voorspel kon word, is die klassifikasie moontlik nie onafhanklik van die nektarbron nie. Die oorvlueling van die Stellenbosch-klas met dié van beide Hopefield en Stanford het voorgekom aangesien die nektarbronne in Stellenbosch 'n mengsel van fynbos en *Eukalptus* was, terwyl dit in die geval van Hopefield fynbos was, en *Eukalptus* in die geval van Stanford. Dit is dus duidelik dat die spektrale data van heuningmonsters van verskillende geografiese liggings, soos ondersoek in die studie, ook variasie as gevolg van die nektarbron insluit. Aangesien al drie geografiese liggings in die Wes-Kaap is, was daar moontlik ook minder omgewingsinvloede. Die goeie resultate van 'n vorige studie (Woodcock, et al., 2007) was moontlik danksy die geografiese liggings van verskillende lande van herkoms. Variasie in die spektrale data van laasgenoemde studie, wat tot die goeie voorspellingsakkuraatheid gelei het, was dus ook moontlik as gevolg van omgewingsfaktore soos grondtipe en weersomstandighede (Alvarez-Suarez, et al., 2014; Francis & Gontier, 1982).

Voorverwerkingstegnieke het 'n groot invloed op die klassifikasie-akkuraatheid van 'n model. Met die ontwikkeling van modelle in hierdie studie is bevind dat EPO die beste klassifikasie-modelle gelever het. Die EPO-tegniek is kompleks, wat moontlik die rede is waarom dit nog nie wyd in navorsing gebruik word nie. Dit dra egter in 'n groot mate tot verbetering van modelle by.

Tydens hierdie studie is klem gelê op die insameling van egte, ongeprosesseerde heuningmonsters. Die koste en tyd-

rowendheid om byeboere te besoek, het die hoeveelheid monsters wat ingesamel kon word, beperk. Die akkuraatheid en robuustheid van die modelle kan dus verbeter word deur die monsterstel te vergroot met verdere monsters uit dieselfde geografiese liggings en persele, asook uit verdere seisoene. Dit sal die variasie, gebaseer op die faktore wat die samestelling van heuning kan beïnvloed, binne 'n geografiese ligging vergroot. Dit sal ook die robuustheid van die model vir die voorspelling van toekomstige monsters verbeter (Alvarez-Suarez, et al., 2014;

Francis & Gontier, 1982). 'n Volledige chemiese analise van die monsters sal bydra tot 'n beter verduideliking van die basis waarop die heuningmonsters op die HKA-tellinggrafiek groepeer en kan ook as verwysingswaardes dien wanneer kwantitatiewe modelle ontwikkel word. Wanneer 'n voldoende aantal monsters ingesamel is, kan ander masjienleertegniese soos terugpropagasië-kunsmatige neurale netwerke (BP-ANN) (Rumelhart, et al., 1986) moontlik tot beter klassifikasie modelle lei (Chen, et al., 2012; Marini, 2009).

Tabel 2: Oorsig van die resultate van PLS-DA-modelle se kalibrasie- en voorspellingsakkuraatheid van die laboratoriuminstrument om die vertoning van verskillende voorverwerkingstegniese te evalueer

Modelnommer	Voorverwerkingstegniese	LV	Kalibrasie-akkuraatheid (%)	Voorspellingsakkuraatheid (%)
DUPLEX				
1a	SNV	5	65,00	50,00
1b	SNV + DT	5	65,00	50,00
1c	DT	4	65,00	50,00
1d	SG	7	80,00	62,50
2	EPO	5	92,50	62,50
2a	EPO + SNV	4	87,50	56,25
2b	EPO + SNV + DT	4	87,50	50,00
2c	EPO + DT	4	95,00	62,50
2d	EPO + SG	7	92,50	75,00
KS				
3a	SNV	5	75,68	47,37
3b	SNV + DT	5	75,68	52,63
3c	DT	2	64,86	36,84
3d	SG	3	70,27	26,32
4	EPO	5	86,49	73,68
4a	EPO + SNV	5	89,19	68,42
4b	EPO + SNV + DT	1	64,86	47,37
4c	EPO + DT	5	86,49	73,68
4d	EPO + SG	4	81,08	42,11

DT (tendensverwydering), EPO (eksterneparameter-ortogonalisering), KS (Kennard-Stone), LV (latente veranderlikes), P (voorspelling), SG (Savitzky-Golay), SNV (standaardnormaalvariaat)

Tabel 3: Oorsig van die resultate van die PLS-DA-modelle se kalibrasie- en voorspellingsakkuraatheid van die handapparaat met die Teflonkoppie om die vertoning van verskillende voorverwerkingstegniese te evalueer

Modelnommer	Voorverwerkingstegniese	LV	Kalibrasie-akkuraatheid (%)	Voorspellingsakkuraatheid (%)
DUPLEX				
5a	SNV	6	80,00	62,50
5b	SNV + DT	6	80,00	62,50
5c	DT	6	67,50	87,50
5d	SG	5	72,50	50,00
6	EPO	4	85,00	56,25
6a	EPO + SNV	4	92,50	68,75
6b	EPO + SNV + DT	4	95,00	68,75
6c	EPO + DT	4	85,00	56,25
6d	EPO + SG	4	97,50	75,00
KS				
7a	SNV	5	89,19	57,89
7b	SNV + DT	5	81,08	57,89
7c	DT	2	54,05	42,11
7d	SG	3	70,27	36,84
8	EPO	2	83,78	47,37
8a	EPO + SNV	2	91,89	52,63
8b	EPO + SNV + DT	2	91,89	52,63
8c	EPO + DT	5	94,59	57,89
8d	EPO + SG	3	89,19	52,63

DT (tendensverwydering), EPO (eksterneparameter-ortogonalisering), KS (Kennard-Stone), LV (latente veranderlikes), P (voorspelling), SG (Savitzky-Golay), SNV (standaardnormaalvariaat)

Tabel 4: Oorsig van die resultate van die PLS-DA-modelle se kalibrasie- en voorspellingsakkuraatheid van die handapparaat met glasflessie om die vertoning van verskillende voorverwerkingstegnieke te assesser

Modelnommer	Voorverwerkingstegnieke	LV	Kalibrasie-akkuraatheid (%)	Voorspellingsakkuraatheid (%)
DUPLEX				
9a	SNV	4	75,00	62,50
9b	SNV + DT	4	75,00	62,50
9c	DT	3	62,50	43,75
9d	SG	4	70,00	43,75
10	EPO	4	87,50	56,25
10a	EPO + SNV	4	82,50	62,50
10b	EPO + SNV + DT	4	82,50	62,50
10c	EPO + DT	4	87,50	56,25
10d	EPO + SG	4	85,00	50,00
KS				
11a	SNV	4	75,68	63,16
11b	SNV + DT	4	75,68	63,16
11c	DT	4	70,27	52,63
11d	SG	6	86,49	47,37
12	EPO	7	94,59	73,68
12a	EPO + SNV	7	94,59	84,21
12b	EPO + SNV + DT	7	94,59	84,21
12c	EPO + DT	7	94,59	73,68
12d	EPO + SG	4	91,89	52,63

DT (tendensverwydering), EPO (eksterneparameter-ortogonalisering), KS (Kennard-Stone), LV (latente veranderlikes), P (voorspelling), SG (Savitzky-Golay), SNV (standaardnormaalvariaat)

Gevolgtrekking

Ten spyte van die oorvleueling in die HK-spasie van die Stellenbosch-klas met beide die Hopefield- en Stanford-klas was dit moontlik om PLS-DA-modelle te ontwikkel om die heuningmonsters volgens hulle geografiese oorsprong te klassifiseer met 'n voorspellingsakkuraatheid van 75% tot 84%, na gelang van die instrument wat gebruik is. Vir die datastelle wat in hierdie studie gebruik is, het dit geblyk dat die verskille in die samestelling van die heuning as gevolg van die nektarbron grootliks bygedra het tot die onderskeid wat betref geografiese oorsprong. Daar was geen onderskeid tussen die monsters op grond van die twee seisoene nie.

Ten spyte van die kompleksiteit van die toepassing van EPO het dié tegniek, vergeleke met ander voorverwerkingstegnieke van rou spektrale data, opmerklik tot die verbetering van akkuraatheid van klassifikasie-modelle bygedra. Die gebruik van DUPLEX as dataskeidingsmetode het gelei tot hoër voorspellingsakkuraatheid vir beide die laboratoriumapparaat en handapparaat (Teflonkoppie). In teenstelling hiermee het die gebruik van KS as dataskeidingsmetode beter resultate gelewer vir die handapparaat (glasflessie). Hoewel laasgenoemde die akkuraatste modelle gelewer het, moet die feit dat die monster meer kere geskandeer is, as gevolg van die padlengte wat verander het tydens herposisionering van die Teflonstaaf, in ag geneem word. Op grond hiervan blyk die handapparaat (Teflonkoppie) die voorstellingsmetode en apparaat te wees wat potensieel die betroubaarste en beste modelle sal lewer. Met die laboratoriuminstrument en die handapparaat (glasflessie), wat beide dieselfde padlengte gehad het, het moontlik meer ligverstrooiing plaasgevind, aangesien die NIR lig eers deur 'n glaslaag moes dring voordat dit die monsters kon bereik. Waar die laboratoriuminstrument outomaties tientalle metings

neem, moes die skanderings met die handapparaat (glasflessie) een vir een met die hand gedoen word.

NIR spektroskopie bied 'n vinnige, goedkoop uitvoerbare metode wat gebruik kan word sonder die hulp van 'n deskundige. Met die handapparaat kan die monsters op enige plek getoets word, aangesien die toestel mobiel is, in teenstelling met die laboratoriuminstrument.

Deur die geografiese oorsprong van Suid-Afrikaanse heuning te bepaal kan verseker word dat heuning korrek gemerk word. Die moontlike implementering van sodanige sertifisering kan bydra om heuning van oorsprong van goedkoop ingevoerde produkte te onderskei en produsente te beskerm. Hierdie proses is egter gekompliseerd en tydrowend en sal die ontwikkeling van 'n institusionele raamwerk vereis met, op sy beurt, unieke uitdagings. Intussen, met 'n klassifikasie-akkuraatheid van 84%, kan NIR spektroskopie as 'n ondersteunende siftingsmetode vir die tradisionele melissopalnologie gebruik word.

Erkenning

Die outeurs bedank Liaan van der Merwe, Pierre van der Westhuizen, Alex Chouler, Danie Vorster en Simply Bee vir verskaffing van heuningmonsters, Sebastian Orth vir die opstel van die tellings- en beladingslyngrafieke asook Eduard Zehrt en Zikrullah Bolukkaya vir die foto's van die instrumente. Finansiële ondersteuning (MV) deur Die Suid-Afrikaanse Akademie vir Wetenskap en Kuns word opreg waardeer.

Botsende belange

Die outeurs verklaar dat daar geen botsing van belange bestaan nie.

ORCID

M Allsopp <https://orcid.org/0000-0003-0620-0612>
 M Manley <https://orcid.org/0000-0001-7581-7208>
 F Marini <https://orcid.org/0000-0001-8266-1117>
 PJ Williams <https://orcid.org/0000-0002-6014-2049>

Datums:

Ontvang: 31/03/2022
 Aanvaar: 21/06/2022
 Gepubliseer: 03/08/2022

Bronnelys

- Alvarez-Suarez, J.M., Gasparrini, M., Forbes-Hernández, T.Y., 2014, The composition and biological activity of honey: A focus on manuka honey. *Foods* 3, 420-432. <https://doi.org/10.3390/foods3030420>.
- Ball, D.W., 2007, The chemical composition of honey. *Journal of Chemical Education* 84(10), 1643-1646. <https://doi.org/10.1021/ed084p1643>.
- Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989, Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy* 43(5), 772-777. <https://doi.org/10.1366/0003702894202201>.
- Bázár, G., Romvári, R., Szabó, A., 2016, NIR detection of honey adulteration reveals differences in water spectral pattern. *Food Chemistry* 194, 873-880. <https://doi.org/10.1016/j.foodchem.2015.08.092>
- Bisutti, V., Merlanti, R., Serva, L., 2019, Multivariate and machine learning approaches for honey botanical origin authentication using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy* 27(1), 65-74. <https://doi.org/10.1177/0967033518824765>.
- Bokobza, L., 1998, Near infrared spectroscopy. *Journal of Near Infrared Spectroscopy* 6, 3-17. <https://doi.org/10.1255/jnirs.116>.
- Callao, M.P., Ruisánchez, I., 2018, An overview of multivariate qualitative methods for food fraud detection. *Food Control* 86, 283-293. <https://doi.org/10.1016/j.foodcont.2017.11.034>.
- Chen, L., Wang, J., Ye, Z., 2012, Classification of Chinese honeys according to their floral origin by near infrared spectroscopy. *Food Chemistry* 135(2), 338-342. <https://doi.org/10.1016/j.foodchem.2012.02.156>.
- Cowe, I.A., McNicol, J.W., 1985, Use of principal components in the analysis of near-infrared spectra. *Applied Spectroscopy* 39(2), 257-266. <https://doi.org/10.1366/0003702854248944>.
- Dias, C., Mendes, L., 2018, Protected designation of origin (PDO), protected geographical indication (PGI) and traditional speciality guaranteed (TSG): A bibliometric analysis. *Food Research International* 103, 492-508. <https://doi.org/10.1016/j.foodres.2017.09.059>.
- Duckworth, J., 2004, Mathematical data preprocessing, in: Workman, J. Jr, Reeves, J.B. III (eds.). *Near-infrared Spectroscopy in Agriculture*. Roberts, C.A., American Society of Agronomy, Crop Science of America, Soil Science of America, Madison, WI, pp. 115-132.
- Escuredo, O., Seijo, M.C., Salvador, J., 2013, Near infrared spectroscopy for prediction of antioxidant compounds in the honey. *Food Chemistry* 141(4), 3409-3414. <https://doi.org/10.1016/j.foodchem.2013.06.066>.
- Escuredo, O., González-Martín, M.I., Rodríguez-Flores, M.S., 2015, Near infrared spectroscopy applied to the rapid prediction of the floral origin and mineral content of honeys. *Food Chemistry* 170, 47-54. <https://doi.org/10.1016/j.foodchem.2014.08.061>.
- Francis, C., Gontier, F., 1982, *The Book of Honey*. Howard Timmins, Kaapstad, pp. 9-12; 26-29
- Geladi, P., MacDougall, D., Martens, H., 1985, Correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy* 39(3), 491-500. <https://doi.org/10.1366/0003702854248656>.
- Gower, J.C., 1985, Properties of euclidean and non-euclidean distance matrices. *Linear Algebra and its Applications* 67, 81-97. [https://doi.org/10.1016/0024-3795\(85\)90187-9](https://doi.org/10.1016/0024-3795(85)90187-9).
- Guelpa, A., Marini, F., Plessis, A., 2017, Verification of authenticity and fraud detection in South African honey using NIR spectroscopy. *Food Control* 73, 1388-1396. <https://doi.org/10.1016/j.foodcont.2016.11.002>.
- Heise, H.M., Winzen, R., 2006, Fundamental chemometric methods, in: Siesler, H.W., Ozaki, Kawata, Y.S., Heise, H.M. (eds.). *Near-infrared spectroscopy*. Wiley-VCH, Weinheim, pp. 125-162. <https://doi.org/10.1002/9783527612666.ch07>.
- Kennard, R.W., Stone, L.A., 1969, Computer aided design of experiments. *Technometrics* 11(1), 137-148. <https://doi.org/10.1080/00401706.1969.10490666>.
- Latorre, C.H., Crecente, R.M.P., Martín, S.G., 2013, A fast chemometric procedure based on NIR data for authentication of honey with protected geographical indication. *Food Chemistry* 141(4), 3559-3565. <https://doi.org/10.1016/j.foodchem.2013.06.022>.
- Leme, L.M., Montenegro, H.R., Da Rocha dos Santos, L., 2018, Relation between near-infrared spectroscopy and physicochemical parameters for discrimination of honey samples from Jatai weyrauchi and Jatai angustula bees. *Food Analytical Methods* 11(7), 1944-1950. <https://doi.org/10.1007/s12161-018-1148-7>.
- Louveau, J., Maurizio, A., Vorwohl, G., 1978, Methods of mellissopalynology. *Bee World* 59(4), 139-157. <https://doi.org/10.1080/0005772X.1978.11097714>.
- Manley, M., 2014, Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews* 24, 8200-8214. <https://doi.org/10.1039/C4CS00062E>.
- Manley, M., Baeten, V., 2018, Spectroscopic technique: Near-infrared (NIR) spectroscopy, in: Sun, D.-W. (red.). *Modern Techniques for Food Authentication*. Elsevier, Oxford, pp. 51-102. <https://doi.org/10.1016/B978-0-12-814264-6.00003-7>.
- Marini, F., 2009, Artificial neural networks in foodstuff analyses: Trends and perspectives - a review. *Analytica Chimica Acta* 635(2), 121-131. <https://doi.org/10.1016/j.aca.2009.01.009>.
- Mouazen, A.M., Al-Walaan, N., 2014, Glucose adulteration in Saudi honey with visible and near infrared spectroscopy. *International Journal of Food Properties* 17, 2263-2273. <https://doi.org/10.1080/10942912.2013.791837>.
- Parikh, R., Mathai, A., Parikh, S., 2008, Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmol* 56, 45-50. <https://doi.org/10.4103/0301-4738.37595>
- Preys, S., Roger, J.M., Boulet, J.C., 2008, Robust calibration using orthogonal projection and experimental design: Application to the correction of the light scattering effect on turbid NIR spectra. *Chemometrics and Intelligent Laboratory Systems* 91(1), 28-33. <https://doi.org/10.1016/j.chemolab.2007.10.007>.
- Roger, J.M., Chauchard, F., Bellon-Maurel, V., 2003, EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. *Chemometrics and Intelligent Laboratory Systems* 66(2), 191-204. [https://doi.org/10.1016/S0169-7439\(03\)00051-0](https://doi.org/10.1016/S0169-7439(03)00051-0).
- Rust, A., Marini, F., Allsopp, M., 2021, Application of ANOVA-simultaneous component analysis to quantify and characterise effects of age, temperature, syrup adulteration and irradiation on near-infrared (NIR) spectral data of honey. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* 253, 119543. <https://doi.org/10.1016/j.saa.2021.119546>.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986, Learning representations by back-propagating errors. *Nature* 323, 533-536. <https://doi.org/10.1038/323533a0>.
- Salonen, A., Julkunen-Tiitto, R., 2012, Characterisation of two unique unifloral honeys from the boreal coniferous zone: Lingonberry and mire honeys. *Agricultural and Food Science* 21(2), 159-170. <https://doi.org/10.23986/afsci.4931>.
- Savitzky, A., Golay, M.J.E., 1964, Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 36(8), 1627-1639. <https://doi.org/10.1021/ac60214a047>.
- Sharma, S., Goodarzi, M., Ramon, H., 2014, Performance evaluation of preprocessing techniques utilizing expert information in multivariate calibration. *Talanta* 121, 105-112. <https://doi.org/10.1016/j.talanta.2013.12.053>.
- Snee, R.D., 1977, American society for quality validation of regression models: methods and examples validation of regression models: methods and examples. *Technometrics* 19(4), 415-428. <https://doi.org/10.1080/00401706.1977.10489581>.
- Szymańska, E., Gerretzen, J., Engel, J., 2015, Chemometrics and qualitative analysis have a vibrant relationship. *TrAC Trends in Analytical Chemistry* 69, 34-51. <https://doi.org/10.1016/j.trac.2015.02.015>.
- Tahir, H.E., Xiaobo, Z., Tinting, S., et al., 2016, Near-infrared (NIR) spectroscopy for rapid measurement of antioxidant properties and discrimination of Sudanese honeys from different botanical origin. *Food Analytical Methods* 9(9), 2631-2641. <https://doi.org/10.1007/s12161-016-0453-2>.
- White, J.W., 1976, Physical characterisations of honey, in: Crane, E. (red.). *Honey: A comprehensive survey*. Heinemann, London, pp. 207-239.
- Williams, P., Antoniszyn, J., Manley, M., 2019, Near-infrared Technology: Getting the best out of light. Stellenbosch: African Sun Media, p. 304. <https://doi.org/10.18820/9781928480310>.
- Wold, S., Ruhe, A., Wold, H., 1984, The collinearity problem in linear regression: The partial least squares (PLS) approach to generalized inverses. *Journal of Scientific Computing* 5(3), 735-743. <https://doi.org/10.1137/0905052>.
- Woodcock, T., Downey, G., Kelly, J.D., 2007, Geographical classification of honey samples by near-infrared spectroscopy: A feasibility study. *Journal of Agricultural and Food Chemistry* 55(22), 9128-9134. <https://doi.org/10.1021/jf072010q>.
- Workman, J. Jr, Shenk, J., 2004, Understanding and using the near-infrared spectrum as an analytical method, in: Roberts, C.A., Workman, J. Jr, Reeves, J.B. III (eds.). *Near-infrared Spectroscopy in Agriculture*. American Society of Agronomy, Crop Science of America, Soil Science of America, Madison, WI, pp. 3-10.