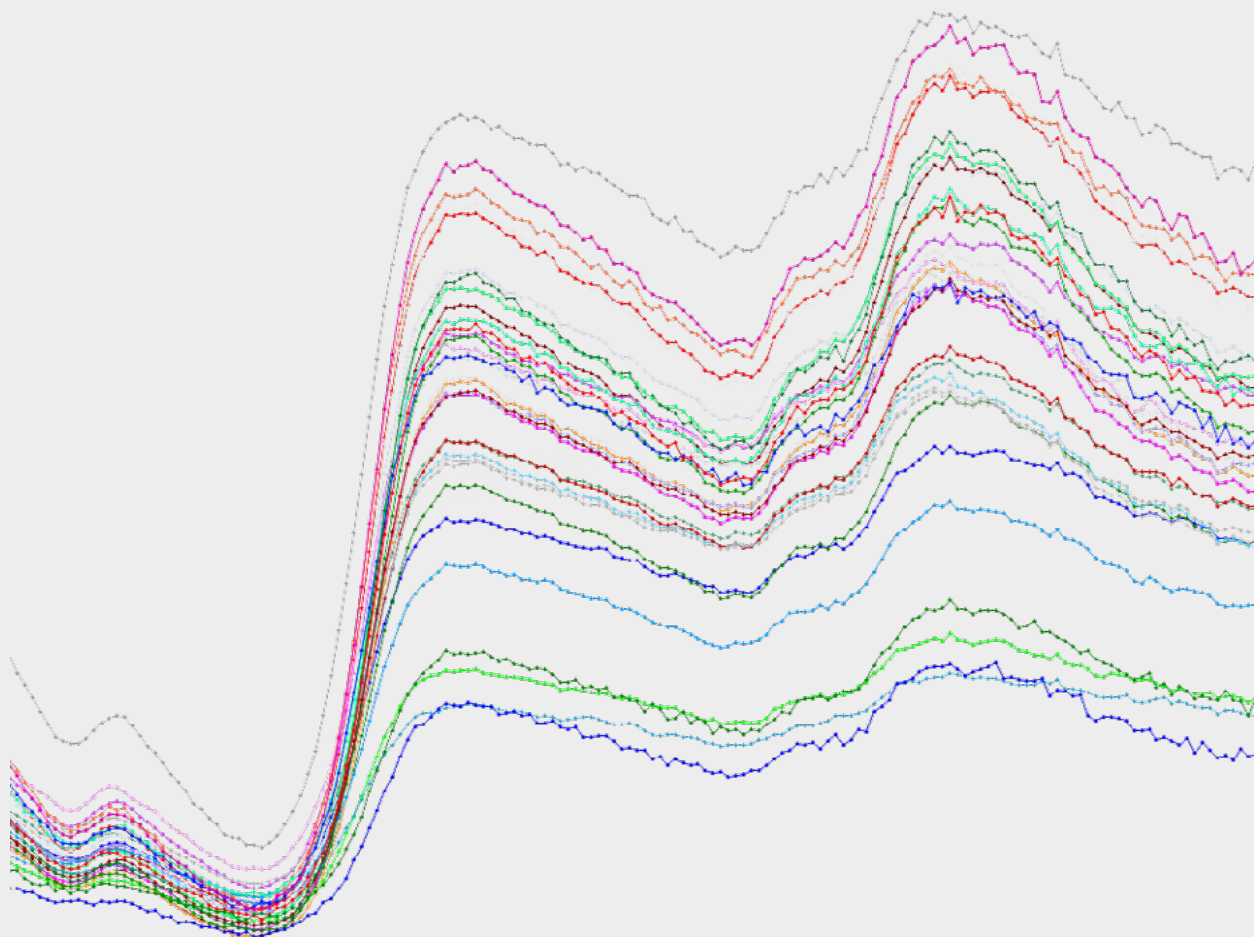


Heinz W. Zwanziger

**Spektroskopie historischer Pergamente im
Domstiftsarchiv Merseburg**



Impressum

Information zur Reihe

Die monografische Reihe „[Manuskripte zur Regionalgeschichte](#)“ (Mitteldeutschlands) erscheint im Hochschulverlag Merseburg.

Kontakt zum Herausgeber:

Regionalgeschichte@hs-merseburg.de

Hochschule Merseburg, Hochschulbibliothek

Eberhard-Leibniz-Str. 2, 06217 Merseburg

Information zur Publikation

2021 Heinz W. Zwanziger

Autor: Heinz W. Zwanziger¹

Titel: Spektroskopie historischer Pergamente im Domstiftsarchiv Merseburg

DOI: <http://dx.doi.org/10.25673/38461>

43 Seiten, 16 Abbildungen, 5 Tabellen (Bearbeitungsende: 07.06.2021)

Bildrechte: Heinz W. Zwanziger

Layout: Yann Labry

Dieses Werk ist ein eBook im PDF-Format und eine freie online-Ausgabe.

Das Werk darf heruntergeladen und als Privatkopie ausgedruckt werden.

Es ist nicht gestattet, Kopien oder gedruckte Fassungen kommerziell zu nutzen.

Werden Auszüge aus dem Werk verwendet, ist es wie folgt zu zitieren.

Zwanziger, Heinz W. (2021): Spektroskopie historischer Pergamente im Domstiftsarchiv Merseburg. In: *Manuskripte zur Regionalgeschichte* 1/2021. Hochschulverlag Merseburg. DOI: <http://dx.doi.org/10.25673/38461>

¹ Hochschule Merseburg, Forschungssemester April bis September 2012, Vorläufiger Abschlussbericht zum Teilprojekt „Anwendungen der Nah-Infrarot-Spektroskopie“, Abschluss der messtechnischen Arbeiten: 29.05.2012. Zur online-Publikation wurde die Fassung vom 24.07.2012 aufbereitet. Herr **Markus Cottin**, Leiter der Domstiftsbibliothek und des Domstiftsarchivs Merseburg, (Vereinigte Domstifter zu Merseburg und Naumburg und des Kollegiatstifts Zeitz) assistierte bei der Spektrenaufnahme und gab Hinweise zum Forschungsbericht.

Zusammenfassung

Mittels zerstörungsfreier Nah-Infrarot-Reflexionsspektroskopie wurden 32 Pergamente (Diplome/ Urkunden) vom 13. bis 18. Jahrhundert aus dem Merseburger Domstiftsarchiv untersucht. Die spektroskopischen Daten wurden mit mathematisch-statistischen Methoden ausgewertet. Mittels explorativer Methoden (Mustererkennung) konnten Ausreißer erkannt werden. Bei den konfirmativen Methoden wurde mittels Diskriminanzanalyse eine 100% richtige Re-Klassifikation der Jahrhundert-Klassen erreicht. Mittels multipler Regression konnte die Zielgröße, das Ausfertigungsjahr von Urkunden, statistisch sicher modelliert werden.

Stichwörter

Pergament, historisches, Spektroskopie, NIR, diffuse Reflexion, explorative Datenanalyse, Clusteranalyse, Faktoranalyse, Diskriminanzfunktionen, Regressionsmodell, Datierung

Abstract

A set of 32 parchments dated from 13th to 18th century has been investigated by near infrared spectroscopy using a reflectance probe connected to a miniature spectrometer. Spectral data have been explored by both explorative and confirmative data analytical methods. Cluster and factor analysis allowed outlier detection. Discriminant functions performed without classification error with respect to century categories. Multivariate regression allowed modelling of given dates of issue.

Keywords

parchment, ancient, historical, spectroscopy, NIR, diffuse reflection, explorative data analysis, cluster analysis, factor analysis, discriminant functions, regression modelling, dating

Inhaltsverzeichnis

1	Was ist Pergament?	5
2	Was wissen wir über unsere Pergamente?	7
3	Was wollen wir wissen?	8
3.1	Motivation	8
3.2	Konkrete Fragestellung	8
4	Welche analytische Technik passt?	9
5	Wie sehen die Spektren aus?	10
6	Sind die Spektren plausibel?	13
6.1	Was sagen Literaturquellen?	13
6.2	Welche Spektralbereiche wurden erfasst?	14
6.3	Fazit Spektrenbewertung	15
7	Welche Kofaktoren sind möglich?	16
7.1	Kofaktor Tierart	16
7.2	Kofaktor Fett	16
7.3	Kofaktor Feuchte	17
7.4	Kofaktor Tinte	17
7.5	Kofaktor Palimpsest	17
8	Welche Datenanalysen passen?	18
8.1	Kurzbeschreibung Clusteranalyse	18
8.2	Kurzbeschreibung Faktoranalyse	19
8.3	Kurzbeschreibung Diskriminanzanalyse	20
8.4	Kurzbeschreibung Regressionsanalyse	20
9	Was entdeckten die explorativen Methoden?	22
9.1	Clusteranalyse	22
9.1.1	Algorithmus nach Ward	22
9.1.2	Single-Linkage-Algorithmus	25
9.2	Faktoranalyse	26
9.2.1	Faktoranalyse aller Wellenlängendaten	26
9.2.2	Faktoranalyse der Obertonschwingungen	27
9.2.3	Faktoranalyse der Kombinationsschwingungen	28

10 Fazit der explorativen Methoden	29
11 Wie erfolgreich waren die modellierenden Methoden?	30
11.1 Diskriminanzanalyse	30
11.1.1 Visuelle Darstellung	30
11.1.2 3D-Darstellung	31
11.1.3 Klassifikationsergebnisse	32
11.2 Multiple Regression (MR)	33
11.2.1 Schrittweiser Modellaufbau	33
12 Fazit der modellierenden Methoden	36
13 Zusammenfassung	37
Quellen	38
Anhang 1: Spektrometer, Software, Rohdaten	38
Anhang 2: Ergebnis-Reproduktion	40
Anhang 3: Liste der Abbildungen und Tabellen	43

1 Was ist Pergament?

Das Wort Pergament geht etymologisch auf die griechische (türkische) Stadt Pergamum zurück. Es wird sprachlich gerne mit dem Begriff Vellum oder Velin gleichgesetzt, der aber speziell Kalbshaut meint. Später hat man besonders glattes, zunächst nur handgeschöpftes Papier, bei dem die Rip-pung des Drahtsiebes nicht mehr erkennbar war, als Velinpapier bezeichnet. (Weiß 1962)

Pergament ist ein beschreibbarer Werkstoff, der wohl schon ca. 2700 v. Chr. dafür benutzt wurde. Tatsächlich sollen die Schriftrollen, die am Toten Meer gefunden wurden, mit Tannin imprägnier-tes Pergament sein.¹ Im Mittelalter hat man es mit Leinsamenöl transparent gemacht und Fenster bespannt. Heute wird Pergament fast ausschließlich für Trommel-Bespannungen u. ä. verwendet.² Pergament hatte zunächst die Massenware Papyrus abgelöst, bis es selbst vom Massenprodukt Papier überholt wurde. Allerdings wurde es nie zum Allgemeingut wie Papyrus und Papier.

Pergament wurde und wird aus Tierhäuten hergestellt. Das wertvollste Pergament erhält man von besonders jungen Tieren oder Föten. Als Tierarten wurden in nördlichen Ländern vor allem Kalb, in südlichen Ländern Ziege und Schaf genutzt. (Rück 1991, S. 20)

Rück (1991, S. 15) nennt anhand einer historischen Quelle die konkreten Arbeitsgänge, nachdem der Fleischer (Metzger) das Rohprodukt Tierhaut hergestellt hatte. Danach wurde es vom Weiß-gerber zum Halbprodukt Pergament verarbeitet. Anschließend oblag früher dem Pergamenter die Feinverarbeitung. Auf die einzelnen technologischen Schritte wie Weichen/Waschen, Äschern mit Kalk (Kälken), Enthaaren und Ausfleischen, Fettentfernung, Entkalken, Trocknen unter Spannung etc. gehen wir nicht näher ein. Im Gegensatz zum Leder wird die Tierhaut nicht gegerbt.

Jedenfalls läuft die Prozedur darauf hinaus, dass von den drei wesentlichen Schichten der Tierhaut zunächst die beiden äußeren Hautschichten (Epidermis und Hypodermis) entfernt wurden. Die verbleibende Dermis weist ihrerseits zwei Schichten bzw. Seiten auf: haarseitig die Papillarschicht (Haarwurzeln) und fleischseitig das Netzgewebe (Chorion). Anhand bestimmter Merkmale (Poren-bild der Haarseite) lassen sich Tierhäute haptisch und optisch unterscheiden, sofern sie nicht zu stark bearbeitet worden sind.³

Man vermutet, dass Pergament über Jahrtausende nach einer mehr oder weniger gleichbleibenden Technologie hergestellt wurde, was im Wesentlichen konstant bleibende Eigenschaften bedingt.⁴

Die Dermis besteht aus unterschiedlich stark miteinander verflochtenen Kollagenfasern. Im Gegen-satz zum Papier, dessen Hauptbestandteil Cellulose vorwiegend aus Cellobiose-Molekülen besteht, ist das Faserprotein Kollagen aus verschiedenen Aminosäuren aufgebaut. 2009 kannte man mindes-tens 20 verschiedene Typen von Kollagenmolekülen. Für das Hautkollagen sind besonders wichtig

1 vgl. Michael L. Ryder (1991): The Biology and History of Parchment, in Rück (1991, S. 13ff.)

2 vgl. Edgar Müller (1991): Echte Hautpergamente und Naturtrommelfelle, in Rück (1991, S. 120ff.)

3 vgl. Erika Eisenlohr (1991): Die Pergamente der St. Galler Urkunden (8.-10. Jahrhundert), in Rück (1991, S. 31ff.)

4 Damit entfielen ein unerkannter Kofaktor, den Datenanalytiker fürchten müssen.

die sog. α -Ketten des Kollagen-Typs I, die als Tripel-Helix vorliegen.⁵ Diese Struktur und damit die Festigkeit des Pergaments kann durch Gelatinisierung zerstört werden, die durch Oxidations- und Hydrolyseprozesse ausgelöst werden kann.^{6,7}

Pergamente können durch viele Faktoren geschädigt werden. Dazu gehören Auswirkungen von Farbstoffen, Tinten, Mikroorganismen, sowie Umweltfaktoren wie Wärme, Licht, Luftfeuchte, staubgetragene und andere Schadstoffe aus der Luft. Somit können insbesondere Lagerungsbedingungen, jedoch auch unsachgemäße Kopierversuche Schäden verursachen.⁸

Unter **Pergamentpapier** versteht man übrigens ein mittels Chemikalien (vor allem Schwefelsäure) fettdicht, nassfest und kochfest gemachtes Papier aus Zellstoff. **Pergamentersatz** („Butterbrotpapier“) erhält dagegen seine Fettdichtigkeit durch eine besondere Mahlung der Zellstofffasern; es ist im Gegensatz zum echten Pergamentpapier nicht wasser- und kochfest.⁹

5 vgl. Herbert Stachelberger, Gerhard Banik und Anna Haberditzl (1991): Naturwissenschaftliche Untersuchungen zum Pergament: Methoden und Probleme, in: Rück (1991, S. 88ff.)

6 Verhornung scheint ein anderes Schadensphänomen (in Verbindung mit Pigmenten) zu sein. siehe Fußnote Faubel et al.

7 vgl. die Untersuchungen von Lee Gonzalez und Tim Wess: Use of Attenuated Total Reflection-Fourier Transform Infrared Spectroscopy to Measure Collagen Degradation in Historical Parchments. Appl. Spectroscopy 62(2008)no.10, 1108-1114

8 Auch diese Schädigungen können zu unerkannten Kofaktoren bei der Datenanalyse führen.

9 Bibliographisches Institut & F. A. Brockhaus AG, 2004

2 Was wissen wir über unsere Pergamente?

Für unsere Untersuchungen haben wir 32 Pergamente aus dem 13. bis 18. Jahrhundert ausgewählt. Die Pergamente sind der Beschreibstoff für Urkunden und lassen sich aufgrund der Datierung der Urkunden relativ genau zeitlich einordnen. Das Merseburger Domstiftsarchiv verfügt insgesamt über 2000 Originalurkunden vom 10. bis zum 18. Jahrhundert, wobei die meisten Urkunden aus dem 15. und 16. Jahrhundert überliefert sind. Während bis um 1450 Pergament als Schriftträger dominiert, tritt nach 1450 Papier immer stärker hervor. Pergament blieb jedoch bis ins 18. Jahrhundert für besondere und wichtige Urkunden die erste Wahl als Schriftträger.

Die ersten fünf Pergamente stammen aus dem 13. Jahrhundert. Die entsprechenden Urkunden sind von verschiedenen Bischöfen ausgestellt worden: Dietrich von Meißen (1204-1215), Ekkehard (1216-1240), Rudolf von Webau (1240-1244) sowie Heinrich von Wahren (1244-1265).

Die nächsten fünf Pergamente stammen aus der Amtszeit von Bischof Gerhard von Schraplau (1320-1340).

Weitere fünf Pergamente sind der Amtszeit von Bischof Johannes II. Bose (1431-1463) zuzuordnen.

Aus der Zeit des Bischofs Sigismund von Lindenau (1535-1544) stammen nochmals fünf Pergamente.

Die fünfte Gruppe enthält Dokumente aus der Zeit, als der sächsische Kurfürst Johann Georg I. Administrator des Hochstifts Merseburg (weltlicher Verwaltungsbereich) war (1603-1656).

Die letzte Gruppe vereint Schriftstücke aus Moritz Wilhelms Sachsen-Merseburger Herzogszeit (1694-1731) und der Endphase der Regierungszeit des sächsischen Kurfürsten Friedrich August II., der 1738 das Sekundogeniturfürstentum übernahm und 1763 starb.

3 Was wollen wir wissen?

3.1 Motivation

Oftmals bestehen Zweifel an der „Authentizität“ eines Dokuments (v.a. Urkunden, aber auch Handschriften, d. h. Texten in Buchform), also der „Stimmigkeit“ von Information und Informationsträger, z. B. dem Inhalt, der Aussage oder der Datierung eines Texts und dem Trägermaterial Papier, Pergament etc. Damit verbundene Fragen sind in der Regel nicht allein „materialkundlich“, sondern nur in Verbindung mit einer fundierten Textkritik zu klären, die stilistische, etymologische und andere Aspekte einbezieht.

Innerhalb bestimmter zeitlicher Grenzen könnte aber infolge der Wiederverwendung bereits beschriebener Pergamente (Palimpseste) eine deutliche Unstimmigkeit auftreten; allerdings nur „unidirektional“: sehr neuzeitlicher Text auf altem Pergament, aber nie „historischer“ Text auf neuzeitlichem Pergament.

Im letztgenannten Fall oder aus anderen Gründen interessiert manchmal einfach nur eine plausibel präzise historische Einordnung eines Pergaments.

Dies sollte gelingen, wenn man einerseits über ausreichend viele und gesichert datierte Pergamente verfügt und andererseits mindestens einen Parameter kennt, der diese Bezugsobjekte sowohl eindeutig beschreibt als auch kausal vom Alter eines Pergaments abhängt.

Bei vielen gleichartigen Materialien lassen sich aus spektralen Eigenschaften solche Parameter gewinnen. Eigenschaften wie Glanz, Vergilbungsgrad (veränderte Lichtabsorption infolge alterungsbedingter Strukturveränderung auf molekularer Ebene) sind vergleichsweise einfach messbar. Spektroskopische Untersuchungen bieten sich hierzu an.

3.2 Konkrete Fragestellung

Hat man genügend Messdaten gewonnen – je mehr, desto besser – kann mittels mathematisch-datenanalytischer Methoden versucht werden, zu klären

- (1) ob den reinen Messdaten „eine Struktur innewohnt“,
- (2) ob die bekannte (vorgegebene) Jahrhundert-Zuordnung modelliert werden kann und
- (3) ob eine Berechnung des Ausstellungsdatums der Urkunden anhand der Messdaten möglich ist.

4 Welche analytische Technik passt?

Bei einzigartigen historischen Objekten wird man zerstörungsfreie und möglichst kontaktlose Methoden bevorzugen. Dazu zählen die photoakustische Spektroskopie und die photothermische Ablenkungs-Spektroskopie¹⁰ ebenso wie andere optische Verfahren.¹¹ Vor allem die Nah-Infrarot-Spektroskopie in Verbindung mit Messtechniken für diffuse Reflexion (NIRS) hat sich bei eigenen Untersuchungen bewährt.

In Vorversuchen an Schreibpapieren,¹² Papyri,¹³ Zeitungs- und Druckerpapier¹⁴ sowie kommerziellen Pergamenten¹⁵ konnte die prinzipielle Eignung der FT-IR-Spektroskopie (ATR-Technik) und der Nah-Infrarot-Reflexionsspektroskopie zur Materialcharakterisierung nachgewiesen werden.

Wir haben hier erstmals versuchsweise die Nah-Infrarot-Reflexions-Spektroskopie auf eine vergleichsweise kleine Stichprobe von Pergamenten angewendet.

Die Nah-Infrarot-Spektroskopie nutzt die Information, die aus der Wechselwirkung von Wärmestrahlung (Infrarot-Licht) mit Molekülen im Spektralbereich der Wellenlängen von 780 bis 2500 nm entsteht.¹⁶ Mit NIR-Frequenzen werden Ober- und Kombinationsschwingungen¹⁷ von funktionellen Gruppen angeregt, wobei insbesondere OH-, CH- und NH-Gruppen erfasst werden.

In der Infrarotspektroskopie arbeitet man vorzugsweise mit Wellenzahlen-Angaben. Der oben genannte Bereich entspricht Wellenzahlen von ca. 12500 bis 4000 cm⁻¹. Aus messtechnischen Gründen haben wir nur den Bereich von 1501 bis 2304 nm (Wellenzahlen von 6662 bis 4340 cm⁻¹) erfasst.

Eine Abbildung zur Technik kann entfallen, zumal wir nicht auf theoretische und konstruktive Details eingehen wollen; man vergleiche Anhang 1. Im Zeitalter der Miniaturisierung ist die **Versuchsanord-**

10 vgl. W. Faubel, S. Heißler, H. Klewe-Nebenius, E. Willin: Zerstörungsfreie Analytik zur Überwachung von metallischen Kunstwerken und historischen Handschriften. Nachrichten Forschungszentrum Karlsruhe 35(2003) No. 4, S. 2002-2006

11 vgl. Bella Dolgin, Valery Bulatov, Israel Schechter: Non-destructive assessment of deterioration by optical methods. Anal. Bioanal. Chem. 388(2007)1885-1896

12 BKA-Forschungsprojekt, „Möglichkeiten zur kriminaltechnischen Differenzierung und Identifizierung von **Schreibpapieren** mittels NIR-Spektroskopie“, (Bericht 23.11.1999, Abschlussbericht 06.01.2000); Spektrometer: NIRSystems SY-1711 (MC 6500), Steuersoftware VISION (1998), Sonde: „Remote Reflectance Probe“ mit interner Referenz, Probenunterlage: Material mit definiertem Reflexionsgrad (99 %); Datentableau: nach Export aus VISION unter Excel 97 (Microsoft, SR-2) durch Organisationsvariablen ergänzt; Statistische Auswertungen: STATISTICA (StatSoft Europe, 1997), ergänzende Auswertungen (Illustrationen): UNSCRAMBLER 6.1 (Camo, 1996)

13 private Mitteilungen an A. Scharsky (UB Leipzig) vom Januar und April 2004; Messungen an **Papyrus-Fragment**: Technik: ATI Unicam UV/VIS-Spektrometer UV2-300; Lichtquelle: 190-1100 nm; Spaltbreite: 4 mm; Abstand 2 mm; Probe zwischen Quarzküvette; Detektor: Photodiode; Referenz: Luft; [AG K. Schlothauer, Hochschule Merseburg]; Messungen an **Papyrus-Fragmenten**: Spektrometer: NIRSystems 6500 von Foss/Perstorp; Lichtquelle: via Sonde: IRP; Technik: Diffuse Reflexion; Sonde aufgesetzt auf Glasscheiben; Raumtemperatur

14 Messungen von **Zeitungs- und Druckerpapier**: Spektrometer: Omega 20 der Fa. Bruins; Lichtquelle: 600-2000 nm; Spalt 3,0; Technik: Diffuse Reflexion; Ulbrichtkugel; Probe im "rotating sample cup"; Raumtemperatur; Detektor: Si (600-750 nm bzw. -1100 nm) bzw. PbS (1000-2000 nm); Datenabstand 1,0 nm; Referenz: Luft; Messungen an **Zeitungspapier**: Technik: ATI Unicam UV/VIS-Spektrometer UV2-300; Lichtquelle: 190-1100 nm; Spaltbreite: 4mm; Abstand 2 mm; Probe zwischen Quarzküvette; Detektor: Photodiode; Referenz: Luft; [AG K. Schlothauer, Hochschule Merseburg]; Messungen an **Zeitungs- und Druckerpapier**: Spektrometer: NIRSystems 6500 von Foss/Perstorp; Lichtquelle: via Sonde: IRP; Technik: Diffuse Reflexion; Sonde aufgesetzt auf Glasscheiben; Raumtemperatur

15 unpublizierte Ergebnisse, H.W. Zwanziger, 2008; Messungen mit FT-IR-Spektrometer Avatar 380 von Thermo Nicolet unter Verwendung des ATR-Messmoduls Smart Golden Gate

16 IUPAC-Festlegung von 1985

17 Oberschwingungen sind energetische Übergänge zwischen dem Schwingungsgrundniveau ($\nu = 0$) und angeregten Schwingungsniveaus ($\nu = 2$) eines Moleküls.

nung denkbar einfach und schnell beschrieben. An einem zigarrenkistengroßen Spektrometermodul befindet sich einerseits der Anschluss für eine Reflexionssonde und andererseits der Datenausgang zu einem Rechner, auf dem die erforderliche Software zur Steuerung des Spektrometers und zur Auswertung installiert ist. Die Reflexionssonde ähnelt einem dicken Bleistift, der direkt auf die Probe aufgesetzt wird. Dessen „Mine“ ist ein Lichtleitkabel für das eingestrahlte Infrarotlicht, das von Probe diffus reflektiert und über ein die zentrale Faser umhüllendes Lichtleiterbündel ins Spektrometer gelangt und dort spektral zerlegt wird.

5 Wie sehen die Spektren aus?

Die Spektren der Pergamente können mittels der ANIR-Software des Leihstellers des Spektrometers in drei verschiedenen Darstellungen gezeigt werden.

Die Spektrendarstellung in folgender Abbildung zeigt die „Rohspektren“ mit ziemlich hohen „absoluten Intensitäten“,¹⁸ in denen bereits Probenunterschiede deutlich werden.

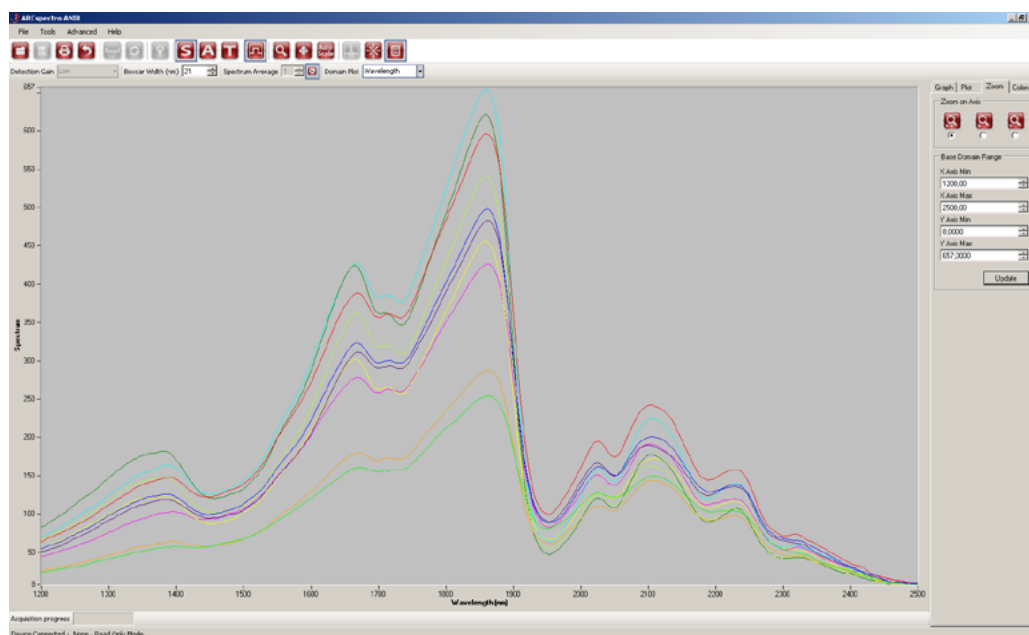


Abb. 1: Pergament-Spektren [ARCSpectro ANIR, geglättet]

Die beiden anderen Darstellungen beziehen sich auf ein Referenzspektrum, das mit einem weißen Reflexionsstandard gemessen wurde.¹⁹

18 Der Rechenweg ist im Handbuch nicht genau angegeben. Es wird aber vermutlich in Analogie zur Definition der Transmission zunächst ein Reflexionsgrad R' als Quotient aus den Lichtintensitäten R berechnet, die von der Probe bzw. dem Standard reflektiert werden: $R' = R_{\text{PROBE}} / R_{\text{REFL_STAND}}$. Das Absorptionsspektrum dürfte dann die berechneten scheinbaren Absorptionen AA wiedergeben: $AA = \lg 1/R' = \lg (R_{\text{REFL_STAND}}/R_{\text{PROBE}}) = \lg R_{\text{REFL_STAND}} - \lg R_{\text{PROBE}}$.

19 labosphere SRS-99-10

In der folgenden **Absorptionsdarstellung** (der fast nicht geglätteten Spektren) erkennt man deutliche Absorptionsmuster im Bereich der 1. Oberschwingungen unterhalb von ca. 1950 nm und im Bereich der Kombinationsschwingungen oberhalb von 1950 nm.

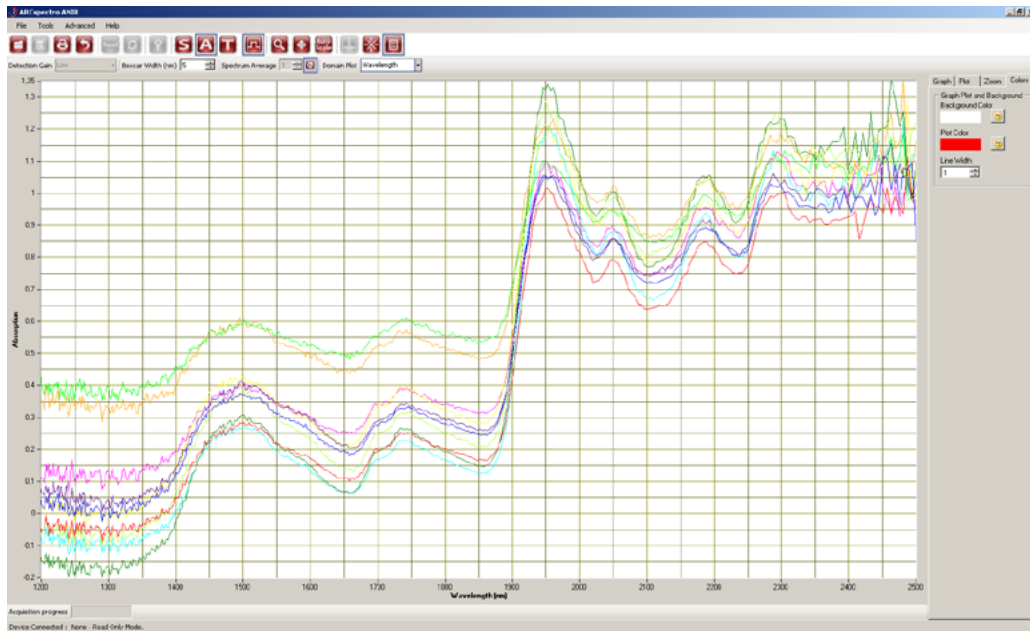


Abb. 2: Pergament-Spektren. [ARCspectro ANIR, Absorption]

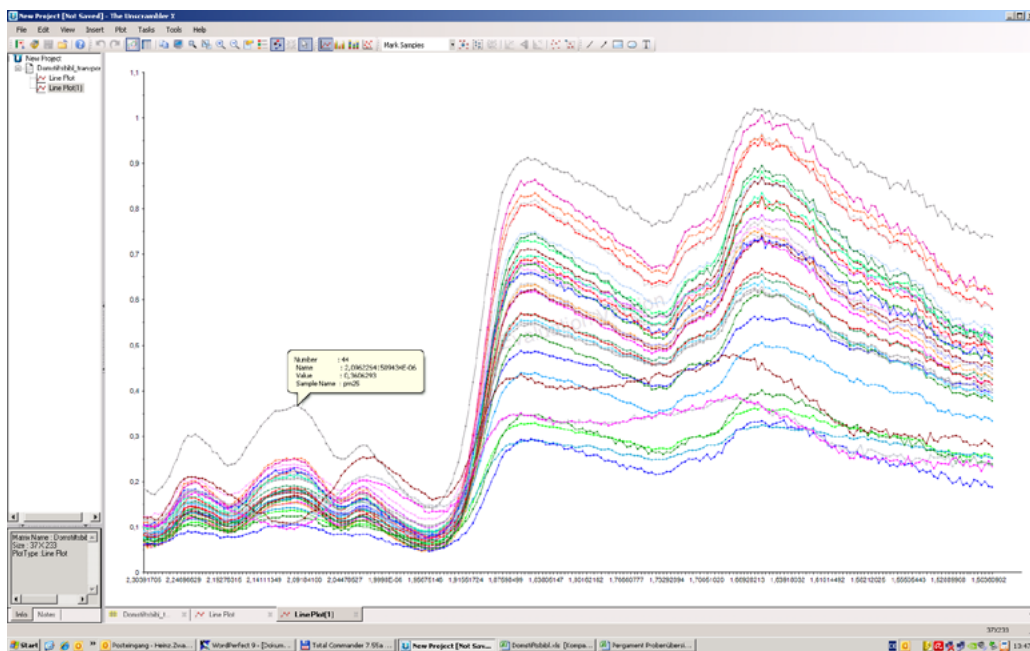


Abb. 3: Pergament- und Papier-Spektren. [Unscrambler, line plot]

In der Abbildung der nicht geglätteten Spektren erkennt man wieder deutlich unterscheidbar den Bereich der **Kombinationsschwingungen** von 2304 bis 1957 nm und den Bereich der **1. Obertöne** von 1953 bis 1501 nm.

Sofort auffällig, und die gesamte Darstellung verzerrend, war Probe pm21 (Jahr 1649), die mit veränderter Verstärkereinstellung²⁰ gemessen wurde. Sie wurde deshalb aus dem Datensatz gestrichen.²¹

Außer dem markierten Pergament pm25 (von 1637) sind als nächste die Proben pm33 (1708), pm33_2 (1708) und pm34_2 (1745) auffällig, also alles Papiere aus dem 18. Jh. Nachdem diese Spektren aus dem Line-Plot entfernt sind, erweist sich zuletzt die Papierprobe pm34 (1745) als auffällig. Wir entfernen auch sie aus dem Bild und dem Datensatz.

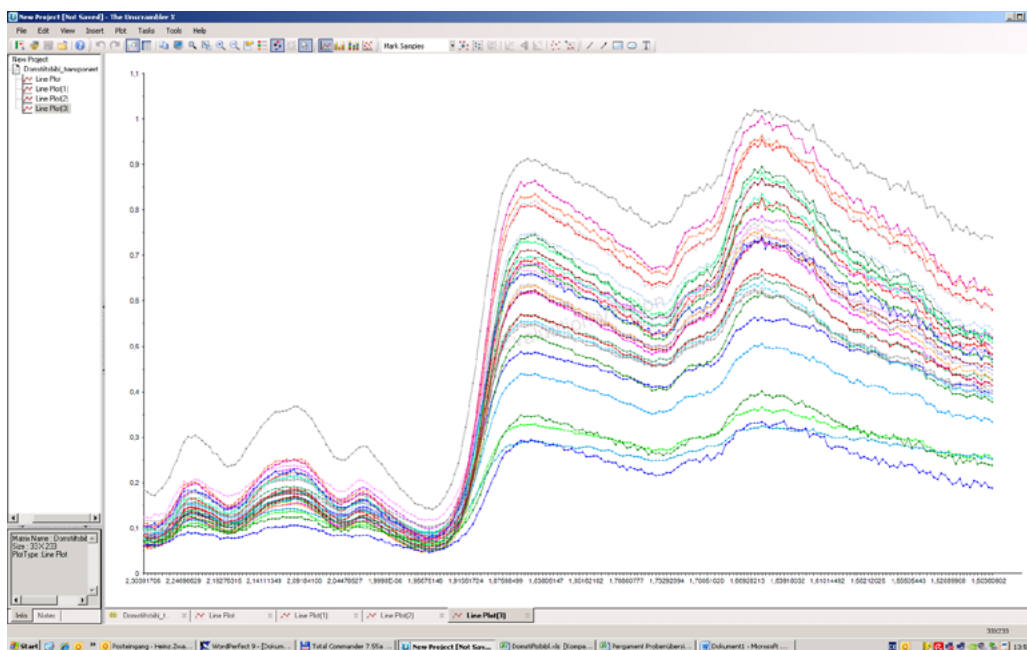


Abb. 4: Pergament-Spektren. [Unscrambler, „fallbereinigt“]

In dieser Abbildung ist nur noch pm25 (1637) auffällig. Dies war ein extrem dünnes Pergament, das deshalb mit untergelegtem Reflexions-Standard beleuchtet wurde.²²

Mitunter lassen sich nach adäquater, d. h. zur Datengewinnung passender, Spektren-Behandlung gerade bei der diffusen Reflexion Unterschiede in den Probeneigenschaften noch deutlicher erkennen. So bietet die Unscrambler-Software u. a. Spektren-Ableitungen an. Eine aktuelle Publikation nutzt z. B. die 4. Ableitung zur Verbesserung von Nah-Infrarot-Spektren.²³ Wir raten von dieser Variante ab, solange nicht alle anderen „Fehlerquellen“ (Detektorrauschen u. ä.) ausgeschlossen oder abgeklärt werden konnten.

20 gain = extreme

21 Um derartige „Ausreißer“ ggf. doch im Datensatz zu belassen, könnte man sie zu skalieren versuchen.

22 Der Einfluss der evtl. „durchscheinenden“ Unterlage müsste experimentell abgeklärt werden. Bei Probe pm7 (1336) hat man den Eindruck, dass ein darunter liegendes Siegel stört.

23 K. N. Norris, A. M. C. Davies: Examining diffuse and transmission spectra more thoroughly: Part 2. Sample temperature and sample composition; Spectroscopy Europe, Vol. 24, No. 2 (2012) 15-20

6 Sind die Spektren plausibel?

6.1 Was sagen Literaturquellen?

Zur detaillierten Interpretation der Banden von Pergament-Spektren im Nah-Infrarot-Bereich fand man in der Literatur bis 2012 nur wenige Hinweise, vor allem aber aus rumänischen Forschergruppen.

Die meisten Publikationen beziehen sich dabei auf den Infrarotbereich und diskutieren Wellenzahlen ν , deren Einheit cm^{-1} ist.²⁴

Wo erforderlich, wurden diese Werte in Wellenlängen λ (Einheit Nanometer, nm) umgerechnet: $\lambda = 10^7 / \nu$ und wir verwenden ausschließlich diese Einheit.

Zur Erinnerung: der von uns ausgewertete Bereich entspricht 1501 bis 2304 nm.

2008 untersuchten **Badea** et al. historische Pergamente 14. bis 16. Jahrhunderts aus dem Turiner Stadtarchiv in **diffuser Reflexion**.²⁵ Sie ordnen die Banden im NIR-Bereich wie folgt zu:

um 1450, um 1950	dominieren Wasserbanden
1730	CH in Lipiden und Proteinen
1820 w (=weak)	CH ₃ und CH ₂ (1. OT)
1500w, 2000w	NH-Streck, NH ₂ -Streck (1. OT) der CONH ₂ -Gruppen (Kollagen)
1490-1500	OH-Streck (1. OT, H ₂ O); neuzeitliches Ziegen-Pergament; Verwendung als Referenz; „appears to be very specific for absorption of water“
1458, 1468-1470	dito in den historischen Pergamenten.

Die breite Bande 1468-1470 wird aufgrund von Messungen anderer Autoren zwei Wasserspezies („second- and first-order absorbed water“) mit schwächeren bzw. stärkeren Wasserstoff-Bindungen zugeordnet.

Die Banden im Bereich 1490-1510 ordnen sie überlappenden Banden von stark „hydrogen bonded water and collagen“ zu, weil Proteine zwei Banden nahe 1400 und 1500 zeigen, „due to the combinations of CH vibrations and the first overtone of NH stretching“.

24 Für die schwingungsspektroskopische Zuordnung von Wellenzahlen im Infrarotbereich hat sich die Colthup-Tabelle bewährt. Vgl. R. Kellner, J.-M. Mermet, M. Otto, H. M. Widmer (Hrsg.): Analytical Chemistry. The Approved Text to the FECS Curriculum Analytical Chemistry. Wiley-VCH. Weinheim, New York u. a., 1998. S. 896, oder K. Doerffel, R. Geyer, H. Müller (Hrsg.): Analytikum. Methoden der Analytischen Chemie und ihre theoretischen Grundlagen. Deutscher Verlag für Grundstoffindustrie. Leipzig, Stuttgart, 1994. S. 222. (Bereichsteilung in cm^{-1} und nm) Leider beginnen diese Tabellen meist bei 4000 cm^{-1} , 2500 nm, wo unser untersuchter Bereich endet.

25 E. Badea, L. Miu, P. Budrugaec, M. Giurginca, A. Masic, N. Badea, G. Della Gatta: Study of Deterioration of Historical Parchments by Various Thermal Analysis Techniques Complemented by SEM, FTIR, UV-VIS-NIR and unilaterial NMR Investigations; J. of Thermal Analysis and Calorimetry, 91(2008)1, 17-27

Tite et al. publizierten 2006 Untersuchungen zum natürlichen Abbau von Pergament und Leder.²⁶ Sie geben folgende Bandenzuordnung bei **diffuser Reflexion** an:

900	OH-, CH-Gruppen
1010	dto.
1177	C-H-Bindungen (HC=CH, CH ₃)
1197	dto.
1480	NH (CONH ₂ + CONHR), OH
1570	NH (CONH)
1726	CH (CH ₂), SH (cystein, ...)
1940	C=O (CONH), OH
2037	C=O (CONH ₂), NH „symmetric stretching + amide II“
2170	2x amide I + amide III (CONHR)

Sie finden bei thermo-oxidativer Alterung eine abnehmende Empfindlichkeit in der Reihenfolge Schwein > „kid“ > Lamm > Kalb.

Miu und **Giurginca** bezogen sich 2006 bei ihren Auswertungen auf die Bandenzuordnung bei historischen Pergamenten.²⁷ Ihre Werte sind:

1180-1200	v(CH ₂)
1485-1500	v(OH) „associated inter/intra-molecular“
1930-1950	v(OH) + δ(OH) ²⁸

Sie stellen in allen Fällen eine Veränderung von Position und Intensität im Vergleich mit neuzeitlichem Pergament (vom Lamm) fest.

Meghea et al. untersuchten 2004 die Pergamente vom Lamm, Kalb, Schwein, Schaf und Rind.²⁹ Sie ordnen typische Banden den Absorptionen der –CONH-Struktur zu: 1182, 1494, 1728, 1936, 2040 und 2157 nm.

6.2 Welche Spektralbereiche wurden erfasst?

Viele der oben genannten Bandenlagen finden wir in den von uns aufgezeichneten Spektren wieder, gehen aber auf Feinheiten der Zuordnung hier nicht ein.

26 T. Tite, L. Miu, M. Giurginca, A. Meghea: Evaluation of degradation of parchment and leather by FT-IR and UV-vis-NIR spectroscopy; Revista de Chimie (Bucharest, Romania) (2006), 57(5), 466-471

27 L. Miu, M. Giurginca: Spectral and kinetic characteristics of natural and accelerated aging of parchment. II; Revista de Chimie (Bucharest, Romania) (2006), 57(1), 42-44

28 Valenzschwingungen werden mit dem griechischen Buchstaben ν bezeichnet, der Buchstabe δ wird für Deformationsschwingungen benutzt.

29 Aurelia Meghea, Maria Giurginca, Nicoleta Iftimie, Lucretia Miu, Bocu Viorica, Petru Budrugaec: Behaviour to accelerate aging of some natural biopolymer constituents of parchment. Mol. Cryst. Liq. Cryst. 118 (2004)285-290.

Zur Übersicht diene die folgende Tabelle.

Wellenlänge	Zuordnung	Literatur	Erfasste Bereiche						
900	OH, CH	Tite et al.	Tite						
1010	OH, CH	Tite et al.							
1177	C-H (HC=CH, CH ₃)	Tite et al.							
1180- 1200	CH ₂	Miu und Giurginca		Miu					
1182	CONH	Meghea et al.			Meghea				
1197	C-H (HC=CH, CH ₃)	Tite et al.						ANIR	
um 1450	Wasser	Badea et al.				Badea			
1458	Wasser	Badea et al.							
1468-1470	Wasser	Badea et al.							
1480	NH, OH	Tite et al.							
1485-1500	OH	Miu+Giurginca							
1490-1500	OH (Wasser)	Badea et al.							
1494	CONH	Meghea et al.							
1500w	NH, NH ₂	Badea et al.						DOM	
1570	NH	Tite et al.					HWZ	2012	
1726	CH, SH	Tite et al.					2012		
1728	CONH	Meghea et al.							
1730	CH	Badea et al.							
1820w	CH ₃ und CH ₂	Badea et al.							
1930-1950	OH	Miu+Giurginca							
1936	CONH	Meghea et al.							
1940	C=O in CONH, OH	Tite et al.							
um 1950	Wasser	Badea et al.							
2000w	NH, NH ₂ in CONH ₂	Badea et al.							
2037	C=O in CONH ₂	Tite et al.							
2040	CONH	Meghea et al.							
2157	CONH	Meghea et al.							
2170	CONHR	Tite et al.							
2300									
2400									

Tabelle 1: Bereiche und Zuordnungen

6.3 Fazit Spektrenbewertung

Die aufgezeichneten Spektren der Pergamente weisen im Spektralbereich des nahen Infrarots typische Absorptionen im Bereich der ersten Obertöne von Grundschwingungen und der Kombinationsschwingungen auf. Sie sind mit Angaben aus der Literatur über Pergament-Untersuchungen vergleichbar.

7 Welche Kofaktoren sind möglich?

Bei der Interpretation von Spektren, besonders aber der Deutung datenanalytischer Ergebnisse darf man mögliche **Kofaktoren** nie außer Acht lassen. So könnte z. B. die tierische Herkunft oder die Herstellung eines Pergaments bestimmte spektroskopische Eigenschaften stärker prägen als Alterungsvorgänge.

Wir gehen deshalb nachfolgend auf einige Kofaktoren ein.

7.1 Kofaktor Tierart

Die mit dem ANIR-Spektrometer aufgenommenen Spektren von kommerziellen Pergamentproben von verschiedenen Tieren zeigen ebenfalls das oben erkennbare typische Absorptionsmuster. Die evtl. beste **Tierunterscheidung** deutet sich dort zwischen 1600 und 1750 nm an, der typischen Region der 1. Obertöne von S-H- und C-H-Bindungen.³⁰

Wir verwenden für die folgenden Auswertungen wieder die Werte des (transmissionsanalogen) Reflexionsgrads; die Umkehrung der Richtung der Abszisse (x-Achse) ist der Ausgabereihenfolge der Spektren-Daten durch die ANIR-Software geschuldet.

Wegen des evtl. hohen Rauschanteils, vor allem aber, um die Zahl der verarbeitenden Variablen etwas zu verringern, hatten wir nur die **Region von 1500 nm bis 2300 nm** ausgewählt. Damit war prinzipiell auch der aus eigenen Vorversuchen bekannte tierrelevante Bereich von etwa 1600 bis 1750 nm enthalten. Allerdings unterschieden sich die Peaks von Kalb- und Ziegenpergamenten mehr in der Intensität als in der Bandenlage.³¹

7.2 Kofaktor Fett

Der Fettgehalt in Pergament ist bislang kaum als Verursacher von Problemen angesehen worden. Fett kann in Pergamenten als Rest bei der Verarbeitung zurückbleiben. Man hat vermutet, dass Lipide zusätzlich durch mikrobielle Zersetzung der Proteine des Kollagens entstehen.³² „The presence of lipids acting as a free radical generator upon interaction with atmospheric sulfur dioxide is suggested as a possible means of the degradation.“

Auch Lipid-Signale können sich in unseren NIR-Spektren zeigen.

30 Für **alle** Auswertungen gilt: Evtl. könnte man ausschließlich die „tierrelevante“ Region 1600- 1750 nm näher untersuchen. Andererseits spricht vielleicht die interessante Struktur der **kommerziellen** Pergamente, gerade dagegen, diese Region bei den **historischen** Pergamenten zu verwenden.

31 Diese Beobachtung berichten auch Aurelia Meghea et al. 2004

32 vgl. C. Ghioni, J. C. Hiller, C. J. Kennedy, A. E. Aliev, M. Odlyha, M. Boulton, T. J. Wess: Evidence of a distinct lipid fraction in historical parchments: a potential role in degradation? in: Journal of Lipid Research 46(2005)2726-2734

7.3 Kofaktor Feuchte

Einen auffälligen aktuellen Feuchtegehalt würden wir in unseren Pergamenten nur als Reflexion um 1450 und 1950 nm erkennen, wenn sie nicht im Signalbereich anderer funktioneller Gruppen, vor allem der Peptidbindungen, untergehen. In unserem Fall könnte nur die Region um 1950 nm ausgewertet werden, deren charakteristisches Muster wir allerdings nicht näher ausgewertet haben.

Durch Feuchte verursachte Folgeschäden ließen sich ggf. durch Signale von Abbauprodukten oder an Feinheiten der Kollagen-Regionen erkennen.

7.4 Kofaktor Tinte

Die Zuordnung der untersuchten Pergamente zu „Jahrhundert-Klassen“ ist willkürlich. Bei Annahme einer weitgehend gleichbleibenden Bearbeitungstechnologie der Tierhäute wäre es nicht verwunderlich, wenn Klassifikationsverfahren bezüglich der **Fragestellung (2)** versagten. Andererseits könnte deren Gelingen auch auf die analytischen Signale von Farbstoffen der Illustrationen oder der Schrift (Tinten) zurückzuführen sein.

Wir schließen ein unmittelbares Wirken eines solchen Kofaktors vorerst aus, da die meisten Farb-Pigmente bzw. Tinten-Bestandteile anorganischen Ursprungs sind. Bekannte mineralische Pigmente wie Azurit, Ultramarinblau und Malachit, aber auch Bleiweiß und Karbonationen (Kalk) sind mit spektroskopischen und mikroskopischen Techniken nachzuweisen.³³

7.5 Kofaktor Palimpsest

Gelegentlich werden beschriebene Pergamente mehrmals wieder verwendet. Dazu werden diese rasiert und erneut beschriftet. Ist das geschehen, so sind auffällige Abweichungen zwischen der Datierung des Pergaments sowie der Datierung des Textes erklärbar. Als Kofaktor konnte die Verwendung als Palimpsest nach Augenschein ausgeschlossen werden.

Wir werden dennoch an mehreren Stellen auf diesen möglichen Kofaktor hinweisen.

33 vgl. S. Bruni, F. Cariati, F. Casadio, L. Toniolo: Identification of pigments on a XV century illuminated parchment by Raman and FTIR microspectroscopies. *Spectrochimica Acta Part A* 55(1999)1371-1377 oder Robin J. H. Clark, Peter J. Gibbs: Raman Microscopy of a 13th-Century Illuminated Text. *Analytical Chemistry News & Features* 99(1998)99A-104A

8 Welche Datenanalysen passen?

Zur Auswertung von wenigen Messdaten einer überschaubaren Menge von Objekten gibt es eine Reihe einfacher grafischer Möglichkeiten, die schnell Erkenntnisse bringen können. Bei hochdimensionalen Datenräumen bzw. großen Datenmengen greift man gewöhnlich zu mathematischen Methoden der Datenanalyse. (Zwanziger, Sorkau 2020)

Zur Klärung der oben formulierten **Frage (1)** wenden wir zwei Methoden der explorativen Datenanalyse (Mustererkennung, pattern cognition) an.

Mit Methoden der **explorativen Datenanalyse** lassen sich Hinweise auf das tatsächliche Vorhandensein von Datenstrukturen finden. So kann die **Clusteranalyse** Gruppen strukturell oder historisch ähnlicher Pergament-**Objekte**.

Gruppen von **Messgrößen**, mit denen diese Pergamentgruppen besonders deutlich hervortreten, ließen sich mittels **Faktoranalyse** finden. Mit besonders aussagekräftigen Messgrößen lassen sich die gefundenen, erwarteten oder a priori bekannten Gruppen sogar „modellieren“. Im Idealfall findet man vielleicht sogar Messgrößen, mit deren Hilfe das Alter direkt berechnet werden kann.

Die beiden **Fragen (2)** und **(3)** werden wir mittels **modellierender Methoden** klären (pattern recognition).

Mit diesen Methoden, die unser Vorwissen über die Pergamente einbeziehen, werden wir zunächst versuchen, mittels der **Diskriminanzanalyse** die vorgegebenen „Jahrhundert-Klassen“ zu modellieren. Schließlich werden wir eine noch feinere Modellierung mittels **multipler Regression** versuchen, um die Zielgröße „Jahr der Ausstellung“ anhand spektroskopischer Information zu berechnen.

Die jeweiligen Modellfunktionen erlauben es also prinzipiell, bislang unbekannte Pergamente, d. h. Pergamente, deren Jahrhundert- oder Jahres-Zuordnung zweifelhaft ist, fehlerarm zuzuordnen oder gar zu datieren.

Man Beachte: Sollte unser Vorwissen fehlerhaft sein, auf dem die Modellannahmen beruhen, werden auch die mathematischen Methoden keine zuverlässigen Ergebnisse liefern. Aber Modelle vermitteln ohnehin keine Wahrheiten, sondern sind immer nur Hilfsmittel bei Erklärungsversuchen.

8.1 Kurzbeschreibung Clusteranalyse

Routinemäßig kann man immer zuerst versuchen, Proben-Ähnlichkeiten, die sich in den Spektren verbergen könnten, mit einer Clusteranalyse zu entdecken.

Bei hierarchischen Clustermethoden kann die Anzahl erwarteter Cluster vorgegeben werden. Wir wollen aber ohne weitere Annahmen die Daten explorieren und wählen deshalb agglomerative Clustermethoden aus, bei denen die Zahl der Cluster nicht vorgegeben wird.

Die agglomerativen Algorithmen führen zuerst die zwei Proben mit der geringsten Distanz zu einem neuen „Pseudoobjekt“ zusammen, ermitteln dann dessen Distanz zu den verbleibenden Objekten, fusionieren dann wieder die beiden Objekte mit der geringsten Distanz usw. usf. Die Algorithmen gehen von einem räumlichen **Distanzmaß** aus. Die entsprechende Distanzmatrix enthält die Euklidischen Abstände jeder Probe zu allen anderen Proben. In unserem Fall sind das Abstände im 233-dimensionalen Raum der Wellenlängen, für die die Intensitäten der Absorption des Lichts bzw. der „Reflexionsgrad“ gemessen wurden.

Die Datenstrukturen werden in einer baumartigen Grafik dargestellt, dem **Dendrogramm**. Links befinden sich die ähnlichsten Objekte, quasi die Blätter, rechts kommt man zu größeren Strukturen, also zu Zweigen, Ästen, Stämmen.

Zur Interpretation kann das Baumdiagramm wie ein Baum „geschnitten“ werden. Beginnt man damit rechts, so gehören zu den Ästen zunächst nur wenige Zweige mit jeweils vielen ähnlichen Objekten. Je weiter man nach links fortfährt, umso deutlicher werden die Unterschiede zwischen Gruppen ähnlicher Objekte, die Inhomogenität der nun zerfallenden vorigen Gruppen wird deutlicher.

Erfahrungsgemäß bringt das Verfahren nach Ward für „gruppenartige“ Strukturen bereits gut interpretierbare Ergebnisse.

Alternativ kann man z. B. mit dem Single-Linkage-Verfahren versuchen, „kettenartige“ Strukturen zu entdecken. Da unsere Pergamentstücke zumindest inhaltlich in einer zeitlichen Reihenfolge stehen, werden wir auch den Single-Linkage-Algorithmus probieren. Etwaige Palimpseste könnten allerdings die erwartete inhaltlich bedingte Reihung „stören“.

Für die Auswertung betrachten wir zunächst das Dendrogramm von der größten Distanz her (im Druckbild von rechts), also zunächst „aus dem größten Abstand, aus dem ein Cluster ähnlicher Objekte sichtbar“ wird. Schrittweise können wir dann die Schnitte bei immer geringeren Distanzen legen und die Ähnlichkeit der Objekte diskutieren, die jeweils ein Cluster bilden.

8.2 Kurzbeschreibung Faktoranalyse

Die Faktoranalyse geht im Gegensatz zur Clusteranalyse nicht von einem Distanzmaß, sondern einem **Ähnlichkeitsmaß** aus, dem **Korrelationskoeffizienten**. Die Faktoranalyse könnte also versuchen, die maximal 233x233-dimensionale Matrix der Korrelationskoeffizienten aller Wellenlängenpaare in ihre Hauptfaktoren zu zerlegen.

Mit der Faktoranalyse (Extraktionsmethode: Hauptkomponenten) finden wir maximal 233 differenzierungsstarke Linearkombinationen von Wellenlängen, sofern bestimmte mathematische Voraussetzungen erfüllt sind.

Von diesen Linearkombinationen interessieren uns nur **zwei**, also **zwei** Faktoren, die eine grafische Darstellung der Proben in der **zweidimensionalen** Papierebene ermöglichen.

8.3 Kurzbeschreibung Diskriminanzanalyse

Ähnlich der Faktoranalyse sucht die Diskriminanzanalyse nach statistisch gesicherten optimalen Kombinationen der Originalmessgrößen. Zusätzlich erlaubt sie aber, nach Aussondern „redunder Information“, mit nur wenigen Originalmessgrößen trennscharfe Merkmalsmengen zu bilden. Diese **Diskriminanzfunktionen** können als **Klassifikationsfunktionen** dienen und eine (hoffentlich) fehlerfreie Probenzuordnung zu definierten Klassen (in unserem Fall: die Jahrhunderte) ermöglichen.

Als „Vorwissen“ verwendet die Methode lediglich die Klassenzugehörigkeit der Objekte in einem Lerndatensatz.

Es gibt theoretisch N solche Klassifikationsfunktionen df , wobei

$$N = \text{Klassenzahl} - 1$$

oder

$$N = \text{Anzahl Messgrößen}$$

ist, je nachdem, welches die kleinere Zahl ist.

Die Diskriminanzfunktionen sind vom Typ

$$df_n = f_n(Y_1, Y_2, \dots) = b_0 + b_1 Y_1 + b_2 Y_2 + \dots,$$

wobei $n = 1, 2, \dots, N$.

Da wir 6 Jahrhundertklassen „modellieren“ möchten, erwarten wir also 5 Diskriminanzfunktionen. Im vorliegenden Anwendungsfall wurden diese Diskriminanzfunktionen mit einer Lernstichprobe ermittelt, die alle unsere Pergamente enthielt.

8.4 Kurzbeschreibung Regressionsanalyse

Mit der multiplen Regression versucht man, die kausale Abhängigkeit einer Zielgröße Z von mehreren unabhängigen Einflussgrößen Y durch ein Regressionsmodell zu berechnen. Aufgrund bestimmter statistischer Kennwerte wird die „Vorhersagekraft“ jeder unabhängigen Variablen bewertet.

In unserem Fall ist Z das Jahr, dem das Pergament zugeordnet wurde, und Y ist eine Wellenlänge.

In das Modell werden dann nach und nach die aussichtsreichsten Variablen einbezogen oder die unbrauchbaren Variablen weggelassen. Wegen der sehr vielen Einzelvariablen bietet sich in unserem Anwendungsbeispiel der schrittweise Aufbau des Modells an.

Im einfachsten Fall verwendet man polynomiale Modelle des Typs

$$Z = f(Y_1, Y_2, \dots) = a_0 + a_1 Y_1 + a_2 Y_2 + \dots$$

Ein mögliches Maß für die Güte des Modells ist das Bestimmtheitsmaß B .³⁴

Auch bei der multiplen Regression kann die Modellgüte mit der „Lernstichprobe“ selbst, d. h. den zur Modellberechnung ausgewählten Pergamente, oder einer geeigneten „Validierstichprobe“ ermittelt werden. Auf Details zu Modellfehlern und Gütemaßen wird hier nicht eingegangen. Sie können an anderer Stelle nachgelesen werden. (Zwanziger, Sorkau 2020)

³⁴ Bei linearen Modellen mit nur einem Y entspricht das Bestimmtheitsmaß dem Quadrat des Korrelationskoeffizienten R_{ZY}^2 .

9 Was entdeckten die explorativen Methoden?

9.1 Clusteranalyse

9.1.1 ALGORITHMUS NACH WARD

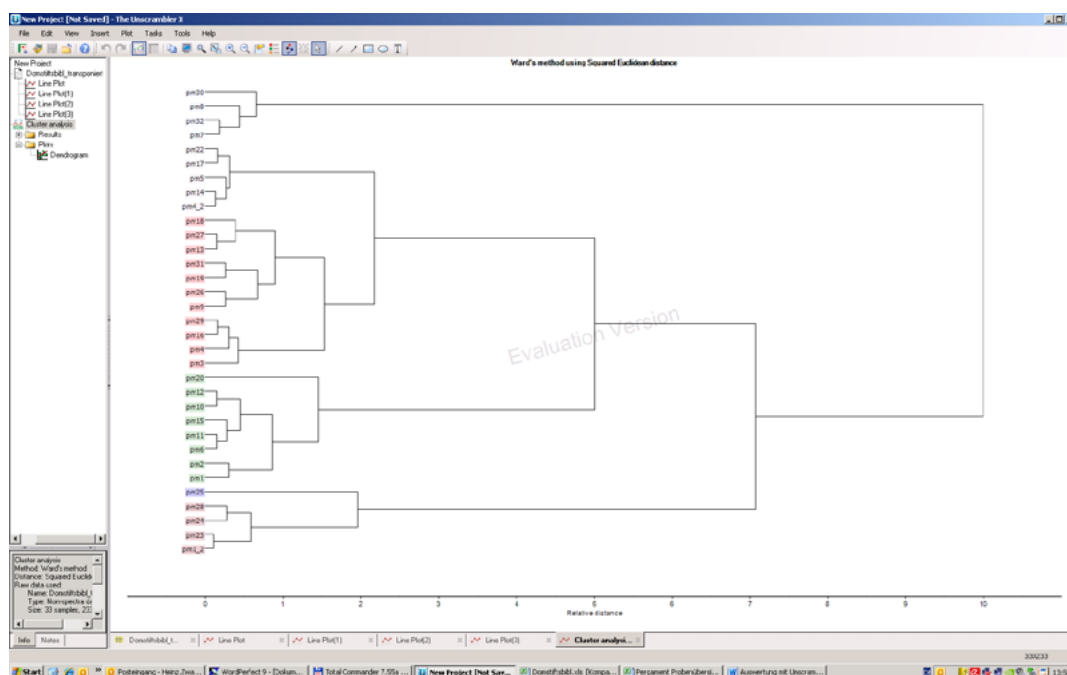


Abb. 5: Clusteranalyse. Gesamter Spektralbereich. Dendrogramm nach Ward [Unscrambler]

In dieser Abbildung wird bei einem „Schnitt“ bei einer relativen Distanz = 8 ist zuerst das obere Cluster mit den Pergamenten pm7 (1336, Siegelreflex?), pm8 (1334), pm30 (1745, oberflächenbehandelt?) und pm32 (1755, sehr glatte Oberfläche) auffällig.

Das verbleibende Cluster teilt sich bei Distanz = 7 in ein kleineres Cluster, in dem Pergamente des 17. Jahrhunderts dominieren, und ein größeres, deutlich strukturiertes Cluster.³⁵

Im kleineren Cluster finden wir die Proben pm1_2 (1240, wie pm1, nur unter stärkerem Andruck beleuchtet), pm23 (1646), pm24 (1635), pm25 (1637) und pm28 (1706).

Das größere Cluster teilt sich bei Distanz = 5 in ein kleineres Cluster (8 Objekte), das 3 der 5 Pergamente des 15. Jahrhunderts enthält, und ein größeres mit 11+5 Objekten. Die Pergamente im

³⁵ Bei Objekt pm 1_2 von 1240 war ein stärkerer Anpressdruck der Sonde erforderlich; die anderen Pergamente stammen aus den Jahren 1646, 1635 und 1637 (das extrem dünn erschien) und von 1706.

kleinen Cluster sind pm1 (1240), pm2 (1258), pm6 (1336), pm10 (1335), pm11 (1456), pm12 (1457), pm15 (1453) und pm20 (1538).

Das Cluster mit den 11+5 Objekten löst sich erst bei einer Distanz ≈ 2 in zwei Cluster auf, die hinsichtlich der Datierung sehr inhomogen sind.

Abgesehen davon, dass die Besonderheiten bestimmter Proben erkannt werden (Ausreißerkennung), bringen uns diese Ergebnisse nicht auf neue Erkenntnisse oder Vermutungen. Wir versuchen daher eine gesonderte Betrachtung der beiden Schwingungsbereiche (1. Obertöne, Kombinationsschwingungen).

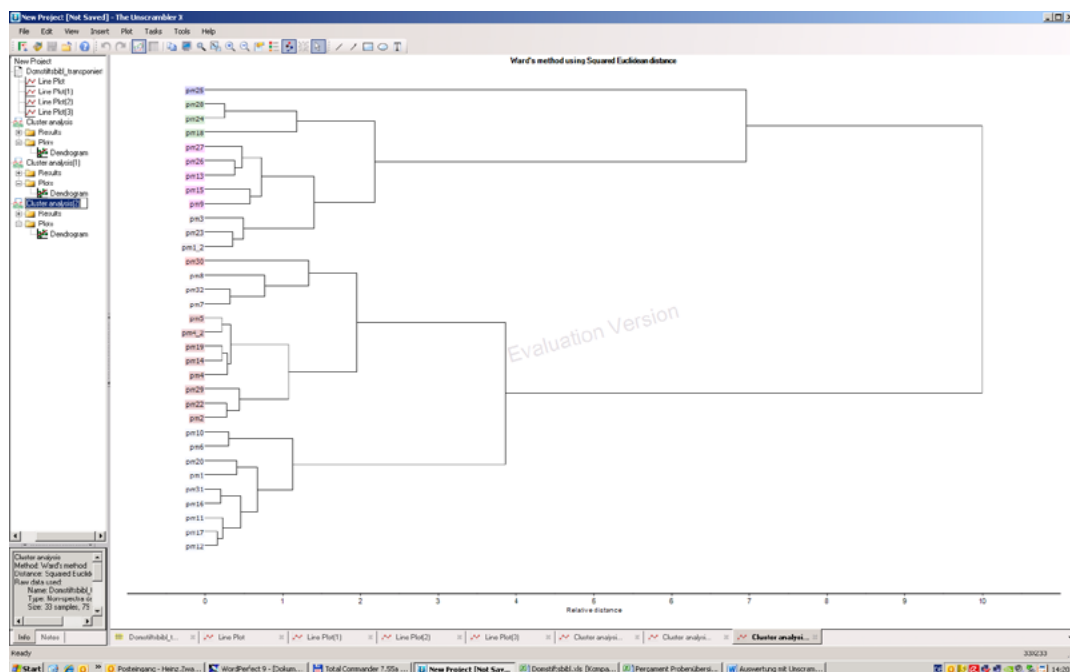


Abb. 6: Clusteranalyse. Kombinationsschwingungen. Dendrogramm nach Ward [Unscrambler]

Für den Bereich der **Kombinationsschwingungen** erhalten wir ein Baumdiagramm, das sich bei Distanz = 10 in zwei Cluster teilt, in dem schon bald, ab Distanz ≈ 7 , die Sonderrolle von pm25 (1637) deutlich wird.

Bei Distanz ≈ 4 schließlich verbleiben drei große Cluster, in denen Pergamente aus allen Jahrhunderten vermischt sind. In einem dieser Cluster dominieren mit 3 von 5 Objekten die des 16. Jh., in einem anderen Cluster finden wir 3 von 5 Schriftstücke des 13. Jahrhunderts.

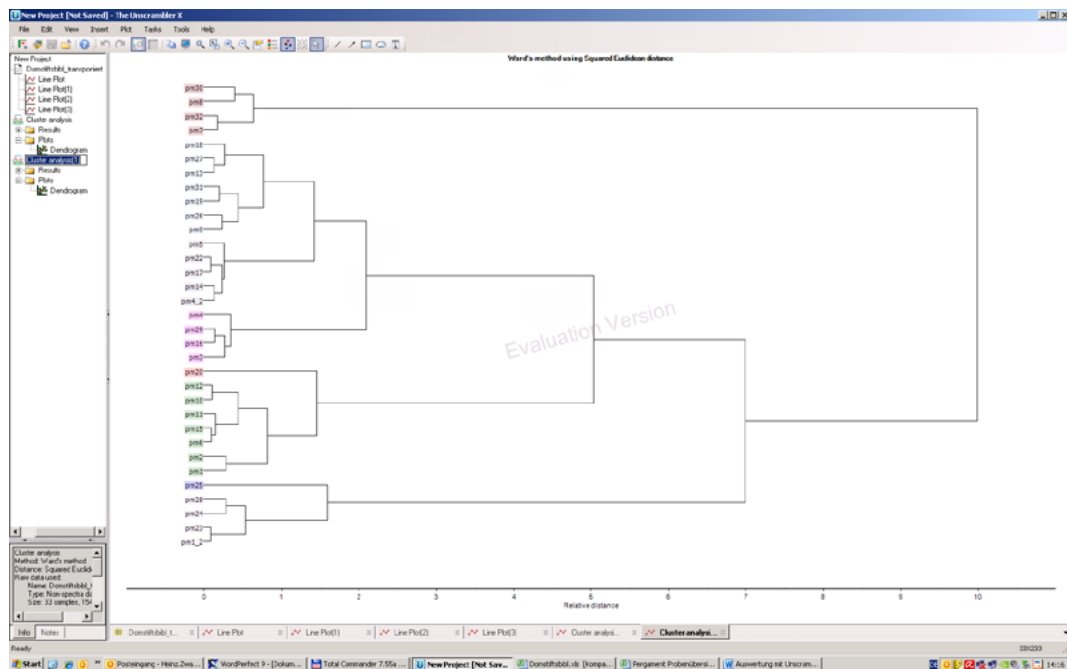


Abb. 7: Clusteranalyse. Obertonschwingungen. Dendrogramm nach Ward [Unscrambler]

Für den **Oberton-Bereich** gibt es ebenfalls eine deutliche Struktur, in der man Gruppierungen mit zwei, drei oder vier Clustern interpretieren kann. Bei Distanz = 5 kann man unterscheiden:

- Cluster 1 mit pm7, pm8, pm30 und pm32}, identisch zum ersten Cluster mit allen Wellenlängen und deutlicher Teil eines Clusters im Kombinationsbereich,
- Cluster 2 mit pm1_2, pm23, pm24, pm25 und pm28 ist identisch zu Cluster 2 des Gesamtbereichs („17. Jh.“) und zu
- Cluster 3 mit pm1, pm2, pm6, pm10, pm11, pm12, pm15 und pm20,
- Cluster 4 mit pm3, pm4, pm16, pm29; pm4_2, pm5, pm9, pm13, pm14, pm17, pm18, pm19, pm22, pm27 und pm31, das in der Summe identisch ist mit Cluster 4 des Gesamtbereichs.

So gesehen, könnte die Empfehlung gegeben werden, die Auswertung statt auf den gesamten Spektralbereich auf den **Oberton-Bereich**, den der höheren Intensitäten, zu beschränken.

9.1.2 SINGLE-LINKAGE-ALGORITHMUS

Nach der nicht sehr erkenntnisfördernden ersten Clusterung nach Ward³⁶ mit allen Intensitäten versuchen wir nun noch den alternativen Clusteralgorithmus **Single Linkage** mit **allen Messwerten**. Das sowieso auffällige Pergament pm25 (1637) wurde nicht mehr einbezogen.

Statt mit den Doppelbezeichnungen der Proben zu arbeiten, verwenden wir fortan nur die Jahreszahlen als Bezeichner.

In der folgenden Abbildung erörtern wir bei Distanz = 0,6 die Cluster „von unten“ her:

- Cluster 1 enthält Pergamente der Jahre 1336, 1334, 1745 und 1755. Neben den zwei Proben aus dem 14. Jh. sind das zwei haptisch auffällige aus dem 18. Jh. Für die letztgenannten Pergamente hat sich der „Palimpsest-Verdacht“ nicht bestätigt. (vgl. Abschnitt Kofaktoren)
- Cluster 2 umfasst ein Pergament von 1240, zwei aus der Administratoren-Zeit (1646 und 1635) und eins aus Herzog Moritz Wilhelms Zeit (1637).
- Cluster 3 besteht nur aus einem nach Augenschein „unauffälligen“ Pergament (1538).
- Cluster 4 enthält die restlichen Pergamente.

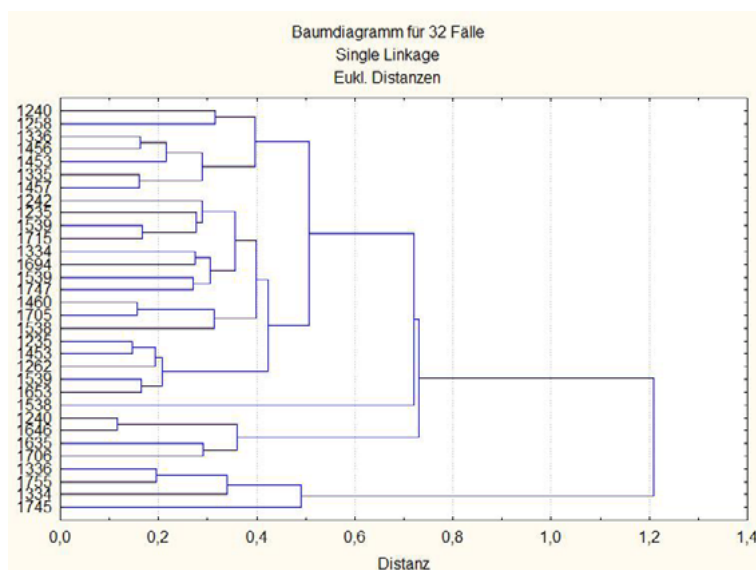


Abb. 8: Clusteranalyse. Gesamter Spektralbereich. Dendrogramm nach Single Linkage

36 ...deren Ergebnisse man vielleicht mit einer Zusatzinformation über die Tierspezifität der Proben besser verstehen kann...

Auch mit diesem Algorithmus lassen sich also wieder keine Vermutungen oder Bestätigungen zu dateninhärenten Strukturen ableiten. Wir verzichten aber auf eine Wiederholung der Single-Linkage-Clustering mit getrennten Wellenlängen-Bereichen.

9.2 Faktoranalyse

9.2.1 FAKTORANALYSE ALLER WELLENLÄNGENDATEN

Die den Berechnungen zugrunde liegende Korrelationsmatrix ist erwartungsgemäß schlecht konditioniert, weil wir es mit sehr ähnlichen Spektren zu tun haben. Die Mathematik findet in unserem Fall nur drei nicht-negative Eigenwerte, wobei der erste für einen Gesamtvarianz-Anteil der Daten von 77% steht und der zweite für 21% extrahierte Varianz. Mit verschiedenen methodischen Varianten gelingt es aber, den Varianzanteil, der durch den zweiten Faktor extrahiert wird, zu senken. Dies gelingt z. B. durch die Beschränkung auf einen ausgewählten Spektralbereich.

Andererseits kann es aus grafischen Gründen, zumal für zweidimensionale Darstellungen, nützlich sein, wenn fast die gesamte Varianz nicht nur auf eine Koordinatenachse entfällt.

Mit allen Wellenlängen ergibt sich zunächst folgendes Bild der Faktorladungen, in dem man zwei Gruppen erkennen kann.

Die numerischen Werte belegen, dass die erste Gruppe (der erste Faktor) durch die Wellenlängen der Kombinationsschwingungen gebildet wird (1501 bis 1908 nm). In der zweiten Gruppe (im zweiten Faktor) sind nur die Beiträge der Obertöne signifikant (1912 bis 2304 nm).

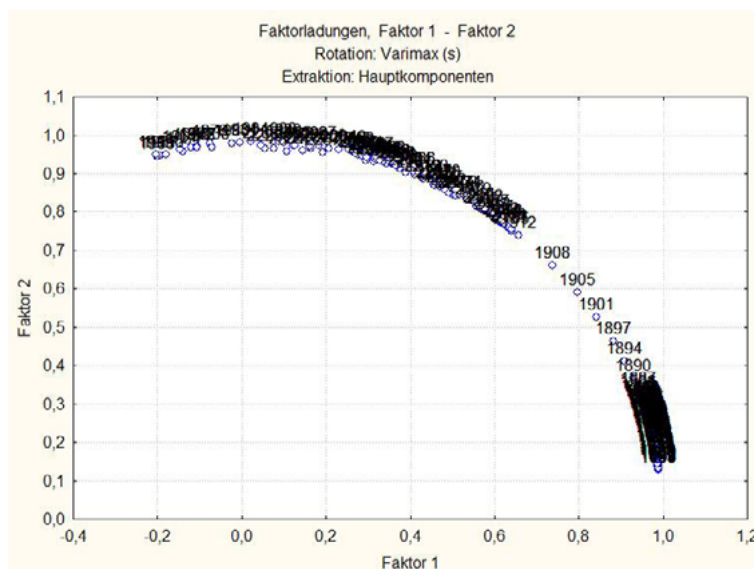


Abb. 9: Faktoranalyse. Alle Wellenlängen. Gruppierung der Messstellen

Eine entsprechende grafische Darstellung der Pergamente im Raum der beiden Faktoren zeigt die folgende Abbildung.

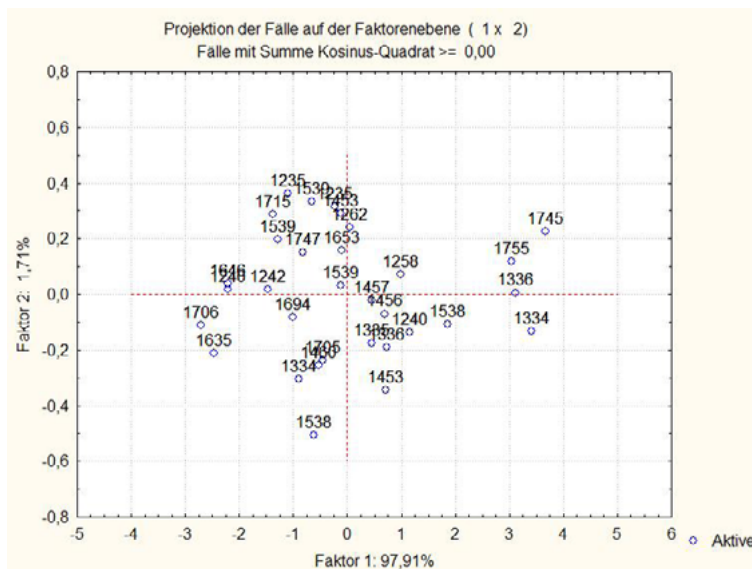


Abb. 10: Faktoranalyse. Alle Wellenlängen. Pergamente im Raum zweier Faktoren

Eine Struktur im Sinne von Gruppen „jahrhundertgleicher“ Pergamente lässt sich noch nicht erkennen.

9.2.2 FAKTORANALYSE DER OBERTONSCHWINGUNGEN

Wir erinnern uns, dass die meisten Interpretationen von Pergamenten (in der Literatur) sich auf den Oberton-Bereich der Schwingungsspektren bezogen. Auch aus der Clusteranalyse ergab sich die Vermutung, dass mit den Messwerten des **Oberton-Bereichs** eine Datenstruktur erkennbar werden könnte. Im Folgenden wiederholen wir eine Faktoranalyse mit diesem Bereich.

Die beiden Hauptfaktoren extrahieren nun 93% und 6% der Varianz, erfassen also 99% der Proben-Diversität.

Besonders deutlicher wird die Struktur der Daten dadurch aber noch nicht, wie man in folgender Abbildung erkennt, die der vorigen sehr ähnelt.

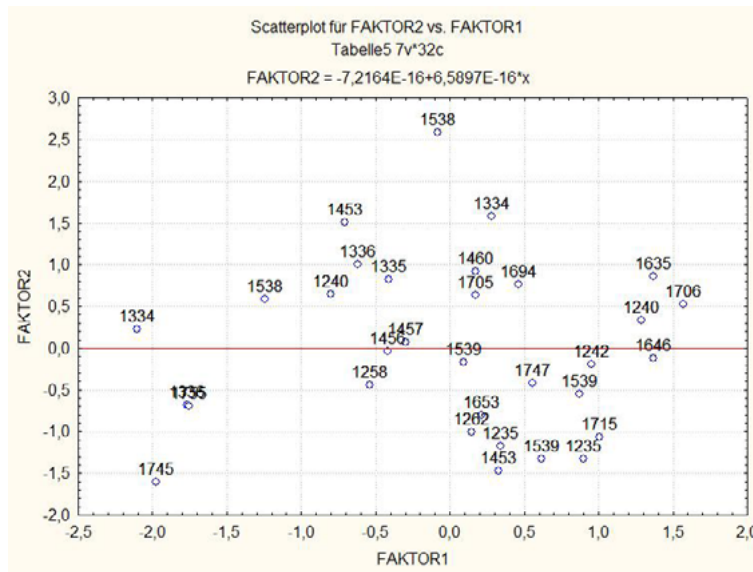


Abb. 11: Faktoranalyse. Oberton-Bereich. Pergamente im Raum zweier Faktoren

9.2.3 FAKTORANALYSE DER KOMBINATIONSSCHWINGUNGEN

Einfach zur endgültigen Gewissheit zeigen wir nun noch das visuelle Ergebnis mit den Messwerten des Bereichs der Kombinationsschwingungen. Auch hiermit gelangen wir zu keiner verbesserten Erkenntnis.

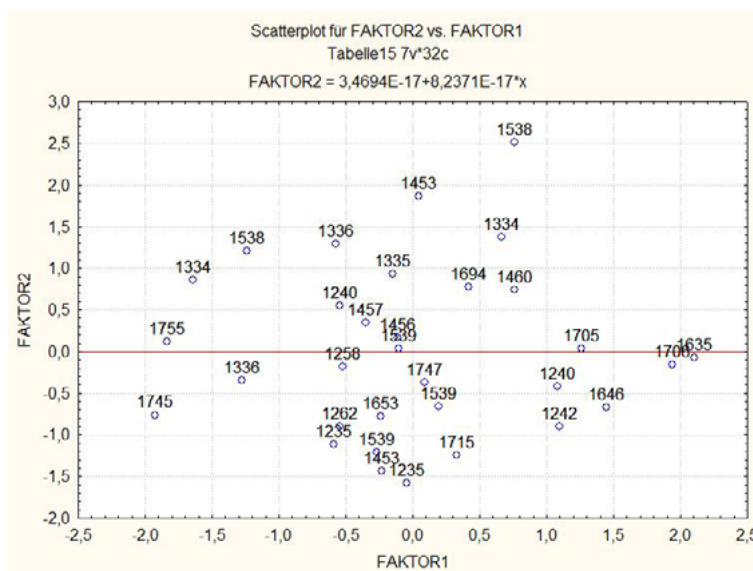


Abb. 12: Faktoranalyse. Kombinationsschwingungen. Pergamente im Raum zweier Faktoren

10 Fazit der explorativen Methoden

Auf den ersten Blick widerspiegeln die **Clusterergebnisse** nicht die anzunehmenden Pergament-Gruppen (Jahrhundert-Gruppen). Dieses Ergebnis ist nicht ermutigend, wenn man von einer gewissen „Ausreißer-Erkennung“ absieht.

Eine weiterführende, gar endgültige Diskussion der mit der Clusteranalyse gefundenen Ähnlichkeiten der in die Analyse einbezogenen Pergamente ist derzeit nicht möglich.

Es scheint ratsam, vor der Clusteranalyse den Merkmalsraum einzuschränken. Eine Möglichkeit besteht darin, nur enge Bereiche um die in der Literatur diskutierten Bandenlagen einzubeziehen. Wir haben dies entsprechend Tabelle 1 vorgenommen und beide Cluster-Strategien erneut angewandt. Es ergaben sich dadurch aber keine neuen Erkenntnisse.

Man sollte jedoch die Möglichkeit nicht ganz ausschließen, dass die Clusterergebnisse einen Hinweis auf das tatsächliche Alter der Pergamente geben. Dann sollte evtl. „Palimpsest“ als eine (binäre) Kovariable dem Datentableau hinzugefügt werden.

Mit Vorliegen der Ergebnisse der **Faktoranalyse** ist nun insgesamt festzuhalten, dass die (rein) explorativen Methoden keinen klaren und überzeugenden Hinweis auf eine „Jahrhundertstruktur“ der Pergamente im Sinne deutlich abgegrenzter Klassen ergeben.

Da die handwerkliche Bearbeitung etc. der Pergamente sich aber nicht sprunghaft mit den Jahrhunderten ändert, ist auch die Klassifizierungsvariable „Jahrhundert“ keine eigentlich ordinale Variable, sondern eher eine intervallskalierte Variable („immer bessere Qualität“).

11 Wie erfolgreich waren die modellierenden Methoden?

11.1 Diskriminanzanalyse

Es wurde zunächst vom gesamten Spektralbereich ausgegangen.

Wegen der hohen Korrelation der Spektren war ein Arbeiten mit der „herkömmlichen“ Diskriminanzanalyse nicht möglich.³⁷ Die gewählte alternative „Allgemeine Diskriminanzanalyse“ ist insbesondere für große Datensätze besser geeignet.

Bemerkenswert ist, dass diese Variante zu Diskriminanzfunktionen führt, die nur Merkmale aus dem **Kombinationsschwingungsbereich** enthalten. Im Beispiel sind es die Wellenlängen 2304 bis 2178 nm (26 Merkmale).

Zur Probe („Kontrolle“) wurde die genutzte Diskriminanzanalyse mit den Daten des Obertonbereichs gestartet. Dabei tritt jedoch ein Fehler auf, der die weitere Analyse unmöglich macht. Mit den Daten des Bereichs der Kombinationsschwingungen ergibt sich die bereits bekannte Wellenlängenkombination, mithin die gleiche visuelle Darstellung.

11.1.1 VISUELLE DARSTELLUNG

In der folgenden Grafik der Objekt-Koordinaten, den Diskriminanzmerkmalen, die mit nur zwei Diskriminanzfunktionen errechnet wurden, zeigt sich die gute Trennung der Jahrhundert-Klassen. Die optisch deutliche Trennung der Proben des 14. und des 17. Jahrhunderts bedarf einer weiteren Dimension der Darstellung, was bei Vorliegen von insgesamt 5 Diskriminanzfunktionen aber kein Problem darstellen sollte. (→ 3D-Darstellung)

37 Eine Gegenüberstellung mit den Daten aus dem BKA-Forschungsprojekt, „Möglichkeiten zur kriminaltechnischen Differenzierung und Identifizierung von **Schreibpapieren** mittels NIR-Spektroskopie“, Bericht 23.11.1999, Abschlussbericht 06.01.2000, zur Modellierung von 40 **Papier**-Klassen verwendeten Merkmalsmengen ergibt folgendes Bild:

$P_{\text{AUSWAHL}} = 50$	$P_1 = 23$ (1. OT)	$P_2 = 18$ (Kombin.)	$P_3 = 9$ (Kombin.)
Wellenlängen (in nm)	1372 - 1416	1970 - 2004	2198 - 2214
Entspräche bei ANIR-Spektren		1968 - 2004	2198 - 2217
			$P_4 = 26$ (Kombin.)
Diskriminanzanalyse (ADM)* fand			2178 - 2304

* Allgemeine Diskriminanzanalytische Modelle. Das Teilmodul ADA findet eine Kombination aus 26 Wellenlängen.

11.1.3 KLASSIFIKATIONSERGEBNISSE

Zur Re-Klassifikation wurden alle Diskriminanzfunktionen und alle Pergamentdaten benutzt. Die Klassifikationsgüte wurde somit ebenfalls anhand der Lernstichprobe ermittelt, da eine extra Validierstichprobe nicht zur Verfügung stand.³⁸

Mit allen fünf Diskriminanzfunktionen erhält man die folgende Zuordnungstabelle, die keine Fehlklassifikationen aufweist.

Interessant wäre eine inhaltliche Diskussion der **nächstwahrscheinlichen Zuordnung**. Danach werden die Pergamente klassenweise (bis auf eine Ausnahme) anderen als den vorgegebenen Jahrhunderten zugeordnet. Als Erklärung käme wohl weniger eine (massenhafte!) **Fehlzuweisung** der Pergamente in Frage. Eher muss man über die Möglichkeit vermehrten Vorhandenseins von **Palimpsesten** nachdenken. Für das 18. Jahrhundert wurde diese Möglichkeit bereits ausgeschlossen. (siehe Abschnitt Kofaktoren)

Schließlich darf man nicht vergessen, dass wir mit sehr kleinen Klassenumfängen arbeiten. Dadurch könnten schon „statistische Feinheiten“ zu anderen Sichtweisen führen.

Datum des Pergaments	Jahrhundert	Zuordnung	Zweithöchste Wahrsch.	Dritthöchste Wahrsch.	Vierthöchste Wahrsch.	Fünfhöchste Wahrsch.
1240	13	13	16	17	14	18
1240	13	13	16	17	14	18
1258	13	13	16	17	14	18
1242	13	13	16	17	14	18
1235	13	13	16	17	14	18
1235	13	13	17	14	16	18
1262	13	13	16	17	14	18
1336	14	14	17	13	16	18
1336	14	14	17	13	16	18
1334	14	14	17	13	16	18
1334	14	14	17	13	16	18
1335	14	14	17	13	16	18
1456	15	15	16	18	13	17
1457	15	15	16	18	13	17
1460	15	15	16	18	13	17
1453	15	15	16	18	13	17
1453	15	15	16	18	13	17
1539	16	16	13	15	18	17
1539	16	16	13	15	18	17
1538	16	16	13	15	18	17
1539	16	16	13	15	18	17
1538	16	16	13	15	18	17
1653	17	17	14	13	16	18
1646	17	17	14	13	16	18
1635	17	17	14	13	16	18
1694	17	17	14	13	16	18
1705	18	18	16	15	13	17
1706	18	18	16	15	13	17
1715	18	18	16	15	13	17
1745	18	18	16	15	13	17
1747	18	18	16	15	13	17
1755	18	18	16	15	13	17

Tabelle 2: Diskriminanzanalyse. Klassifikationsergebnisse

³⁸ Für ein sinnvolles Verfahren der Kreuzvalidierung anhand der Lernstichprobe ist diese zu klein.

11.2 Multiple Regression (MR)

11.2.1 SCHRITTWEISER MODELLEAUFBAU

Startet man mit den Daten des gesamten Wellenlängen-Bereichs und baut schrittweise ein Regressionsmodell auf, so bezieht der Algorithmus die in der folgenden Tabelle aufgeführten Wellenlängen ein. Alle diese Wellenlängen sind **statistisch signifikant**.

Es handelt sich dabei um 19 Wellenlängen, von denen 13 dem **Obertonbereich** (ca. 1475 bis 2050 nm) und 6 dem **Kombinationsschwingungsbereich** (ca. 1950 bis 2500 nm) zugeordnet werden können. Nur zwei Wellenlängen liegen im „Überlappungsbereich“ (ca. 1950 bis 2050 nm). Die Nützlichkeit der Oberton-Schwingungsdaten wird damit erneut bestätigt.

Dies wird durch die numerischen Werte der standardisierten Regressionskoeffizienten ($a_{i,St}$ und a_i) unterstrichen. Ihre Standardfehler (stdf.), die Werte der statistischen t-Verteilung sowie der Wahrscheinlichkeiten p müssen wir an dieser Stelle nicht diskutieren.

Anhand der standardisierten Regressionskoeffizienten $a_{i,St}$ erkennt man die Wellenlängen, die das größte Gewicht im Regressionsmodell haben. Wir wählen willkürlich alle $a_{i,St} > 2$ aus. Aus dem Obertonbereich sind dies vor allem 1767 und 1751 nm, gefolgt von 1686, 1658, 1639, 1930, 1984, 2160, 1779 und 1897 nm. Dazu könnte man noch 2008 und 2020 nm zählen, die im Übergangsbereich zu den Kombinationsschwingungen. Aus diesem Bereich ist eigentlich nur die Schwingung bei 2202 nm beitragsstark.

Eine detaillierte Interpretation der Bedeutung der Wellenlängen im Sinne von Tabelle 1 nehmen wir in dieser Studie nicht vor. Es bestätigt aber sich einmal mehr die Bedeutung des Bereichs der Oberschwingungen.

<Index> i		$a_{i,St}$	stdf.	a_i	stdf.	t(12)	p-Niveau
0				867		27	0,00
1	2278	-1,8	0,2	-17516	2308	-7,6	0,00
2	1930	-4,2	0,2	-37589	2047	-18,4	0,00
3	1984	4,2	0,4	38127	3358	11,4	0,00
4	1992	-1,3	0,2	-10950	1946	-5,6	0,00
5	1658	5,1	1,2	5069	1212	4,2	0,00
6	2008	-5,5	0,5	-40824	3623	-11,3	0,00
7	2202	7,0	0,3	51547	1898	27,2	0,00
8	2169	-0,9	0,3	-6256	2500	-2,5	0,03
9	2020	3,5	0,3	22911	1918	11,9	0,00
10	1767	-11,4	1,2	-16651	1805	-9,2	0,00
11	2288	1,9	0,1	19143	1339	14,3	0,00
12	2257	-1,3	0,3	-9574	2120	-4,5	0,00
13	1751	9,0	0,8	13957	1287	10,8	0,00
14	1686	-5,7	0,8	-6889	933	-7,4	0,00
15	2160	-2,8	0,6	-18294	4104	-4,5	0,00
16	1639	4,4	1,0	4462	1037	4,3	0,00
17	1957	1,7	0,3	17550	3023	5,8	0,00
18	1897	2,1	0,5	4884	1121	4,4	0,00
19	1779	-2,8	1,2	-3990	1641	-2,4	0,03

Tabelle 3: Multiple Regression. 20 Aufbauschritte

Die folgende Abbildung zeigt dazu die recht gute **Übereinstimmung** der **Datierung** (der gegebenen Werte) und der **berechneten Jahreszahlen** (Prognosewerte).

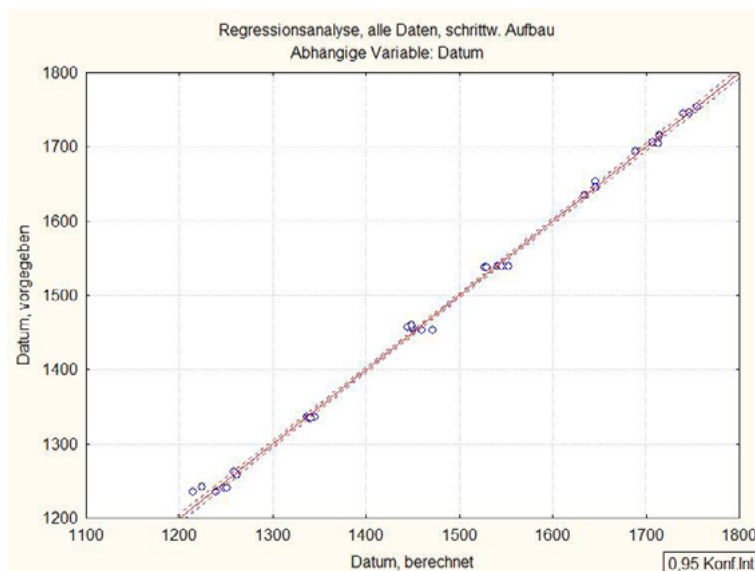


Abb. 15: Multiple Regression. Diplom-Jahr vs. berechnete Jahreszahl

Mit dem linearen Regressionsmodell, d. h. der gefundenen Variablenmenge, können nun auch prinzipiell die Jahreszahlen für bislang undatierte Dokumente vorausberechnet werden.³⁹

Allerdings erkennt man in der vorigen Abbildung schon, dass viele Voraussagewerte außerhalb des Vertrauensbandes liegen, in dem 95% Sicherheit der Zuordnung gegeben ist ($p < 0,05$).

Genauer erkennt man die Abweichungen, mit denen man rechnen muss, anhand eines Residualplots, wie in der nächsten Abbildung gezeigt.

³⁹ Der dabei zu erwartende Fehler könnte jedoch in diesem Stadium nicht sicher quantifiziert werden, da weitere sicher datierte Pergamente für einen Validierdatensatz noch nicht spektroskopiert wurden.

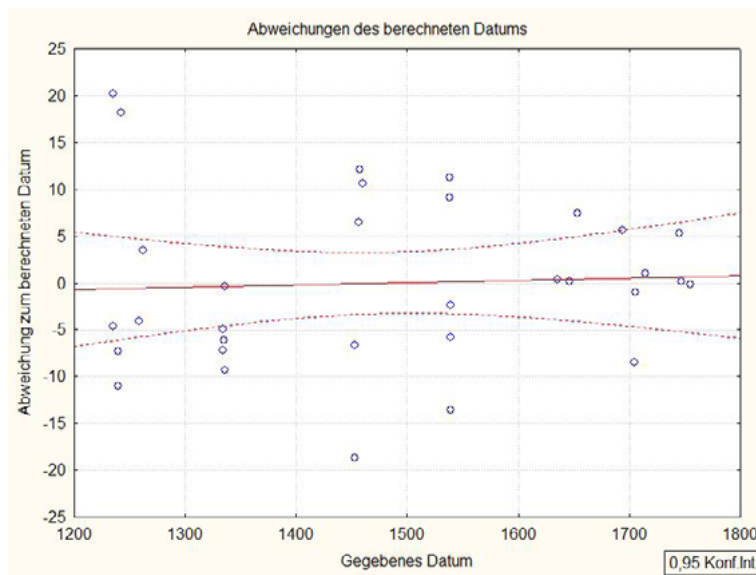


Abb. 16: Multiple Regression. Abweichungen der berechneten Jahreszahlen

Sieht man von den Diplomen des 14. Jahrhunderts ab, könnte man vermuten, dass die Vorausberechnung der Jahresangaben umso genauer wird, je jünger die Pergamente sind. Man muss von dieser Genauigkeit oder Präzision die Richtigkeit unterscheiden. Sie ist bei den Pergamentpunkten am höchsten, die auf oder nahe der Null-Linie liegen.

12 Fazit der modellierenden Methoden

Nach den fast „enttäuschenden“ Ergebnissen der explorativen Methoden musste die Erfolgsaussicht für eine **Diskriminanzanalyse**, zumindest im Raum der zahlreichen originalen Messgrößen, als gering eingeschätzt werden.

Überraschend war, dass dennoch Linearkombinationen, d. h. Diskriminanzfunktionen, der Originalmessgrößen gefunden werden konnten, die eine 100% richtige Zuordnung aller Pergamente zu den Jahrhundertklassen ermöglichten.

Dies lässt den Versuch aussichtsreich erscheinen, eine noch feinere Modellierung, nämlich der Datumsabhängigkeit von Spektraldaten, mittels regressionsanalytischer Methoden zu versuchen.

Wie auch schon bei der Diskriminanzanalyse findet man beim Einbeziehen von „Vorwissen“ zu den Messdaten, hier bei der **Multiplen Regression** sind das die Werte der Zielgröße „Dokumentenjahr“, hinreichend statistisch gesicherte Modelle.

Mit diesen Regressionsfunktionen kann das (wahrscheinliche) Entstehungsjahr von Pergamenten berechnet werden. Dabei muss allerdings mit einem weiten Konfidenzbereich der Prognosewerte gerechnet werden.

Überhaupt muss man immer vor einer zu weit gehenden „Zahlengläubigkeit“ warnen, solange nicht alle Kofaktoren bekannt sind und berücksichtigt wurden.

13 Zusammenfassung

Dieser Studie an historischen Pergamenten aus mehreren Jahrhunderten liegen Messwerte der zerstörungsfrei anwendbaren **Nah-Infrarot-Reflexionsspektroskopie** zugrunde.

Die sonst im ersten Schritt meist erfolgreichen **explorativen Datenanalysen** brachten keine befriedigenden Ergebnisse – wahrscheinlich aufgrund der hohen Anzahl der hochkorrelierenden Messgrößen.

Wird diese Anzahl mittels Hauptfaktorzerlegung u. ä. reduziert oder wird a-priori-Wissen um die Pergamente einbezogen, werden dateninhärente Strukturen deutlicher.

Die **konfirmativen Datenanalysen** waren wesentlich erfolgreicher. Mit den gefundenen **Diskriminanzfunktionen** gelang eine zu 100% richtige Jahrhundertzuordnung der Pergamente.

Mit einem **Regressionsmodell** unter Verwendung von 19 statistisch signifikanten Wellenlängen gelang eine Jahresberechnung, die maximal 20 Jahre vom gegebenen Datum abwich.

Damit eröffnet sich ein Ansatz wahrscheinlicher Zuordnung von zweifelhaft datierten Diplomen.

Quellen

Peter Rück (Hrsg., 1991): Pergament – Geschichte. Struktur. Restaurierung. Herstellung. Jan Thorbecke Verlag Sigmaringen

Karl Theodor Weiss (1962): Handbuch der Wasserzeichenkunde. Bearbeitet und herausgegeben von Wisso Weiß. Fachbuchverlag, Leipzig 1962. Reprint 1983

Heinz W. Zwanziger, Eduard Sorkau (2020): Kalibration analytischer Methoden. Theorie und Techniken, Hochschulverlag Merseburg 2020, ISBN 978-3-948058-33-3

Anhang 1: Spektrometer, Software, Rohdaten

Spektrometer

Als Spektrometer wurde eine befristete Leihstellung der Fa. AcalBfi Germany genutzt (23. bis 30.04.2012). Es handelte sich um das Miniatur-Spektrometer ANIR 09 – 2.6 der Fa. Arcspectro ANIR mit Lichtquelle⁴⁰ HL-2000-HP-FHSA und Reflexionssonde R600-7-VIS der Fa. OceanOptics.

Eine Optimierung der Reflexionsunterlage war aus Zeitgründen nicht möglich.

Software

Die Spektren wurden aus der Gerätesoftware in Form von Textdateien exportiert und unter MS Excel⁴¹ in einer Datenmatrix zusammengeführt.

Erste Auswertungen am 02.05.2012 wurden mit Unscrambler 10.2 (Demo-Version für 30 Tage) vorgenommen. Da die Demo-Version keine Ergebnisausgabe in Dateiform ermöglichte, wurde die Datenanalyse mit dem Programmpaket STATISTICA⁴² wiederholt und fortgeführt. Die schlechte Qualität der mit Unscrambler gewonnenen Abbildungen ist dem Umweg über *screen shots* geschuldet.

Die Graphiken enthalten Titel, Achsenbezeichnungen etc. im Original der jeweiligen Software und wurden nicht bearbeitet.

⁴⁰ Vor künftigen Untersuchungen ist die Strahlungscharakteristik zu beachten bzw. zu dokumentieren.

⁴¹ Microsoft Office, Excel 2007, Einzelplatzlizenz

⁴² STATISTICA 8.0, StatSoft Inc., Tulsa, USA, 2008, Einzelplatzlizenz

Rohdaten

Die Pergamente wurden von den Vereinigten Domstiftern zu Merseburg und Naumburg und des Kollegiatstifts Zeitz am 27.04.2012 bereitgestellt. Die Messungen, bei denen Domstiftsarchivar Markus Cottin assistierte, erfolgten direkt im **klimatisierten** Domstiftsarchiv Merseburg.

Ausgangspunkt aller Auswertungen ist ein Datentableau, dessen typische Struktur in der folgenden Tabelle gezeigt wird: Zeilenweise sind die Pergamente aufgeführt, die gelegentlich auch Objekte, Fälle, Proben genannt werden. In der ersten Spalte sind ihre Bezeichnungen aufgeführt. Die nächsten beiden Spalten enthalten für jede Probe die Werte zweier (zunächst lediglich) beschreibender Variablen, dann eine Klassifikationsvariable (Jahrhundert). Die restlichen Variablen sind die eigentlichen Messgrößen. Sie sind durch die jeweilige Wellenlänge λ (Einheit **Nanometer**: nm) bezeichnet und enthalten die Messwerte aus dem nah-infraroten Spektralbereich. In unserem Fall sind es Werte des transmissionsanalogen Reflexionsgrads.

Probe	Signatur	Datum	Klasse	2304 nm	2299 nm	2293 nm	2288 nm
pm1	41	1240	13	0,077110518	0,07583897	0,07409699	0,07896396
pm1_2	41	1240	13	0,08431121	0,08857184	0,08420174	0,08195776
pm2	52	1258	13	0,08424563	0,07622869	0,07911892	0,07757806
pm3	43	1242	13	0,08030286	0,08033422	0,07889634	0,07339362
pm4	40	1235	13	0,05889904	0,05840012	0,06041808	0,06072261
pm4_2	40	1235	13	0,06386287	0,06248142	0,05443345	0,05517133
pm5	55	1262	13	0,06099703	0,06135105	0,05541892	0,05872691
pm6	241	1336	14	0,09070972	0,09437509	0,09081084	0,0869662
pm7	244	1336	14	0,06920986	0,06465221	0,06968123	0,06869465
pm8	229	1334	14	0,07391376	0,07915057	0,07671726	0,07589577
pm9	233	1334	14	0,09920764	0,09989835	0,10076426	0,10331441
pm10	234	1335	14	0,09924966	0,09432905	0,09589723	0,09480652
pm11	543	1456	15	0,0915055	0,08751731	0,08699497	0,08988165
pm12	546	1457	15	0,08786552	0,08446907	0,08287818	0,08496967
pm13	550	1460	15	0,09905877	0,10409225	0,09900973	0,0988898
pm14	536	1453	15	0,06615454	0,05412146	0,05639554	0,05815502
pm15	538	1453	15	0,11544325	0,11654951	0,11470852	0,11307121

Tabelle 4: Prinzipieller Aufbau des Datentableaus

Insgesamt sind 34 Objekte spektroskopisch untersucht worden, davon 2 Papierproben. Bei einem Pergament (Objekt pm4_2) und den beiden Papieren wurde die Reflexionssonde an zwei verschiedenen Stellen platziert.⁴³

Die Zuordnung der Proben zu den „Jahrhundert-Klassen“ geht aus folgender Tabelle hervor.

⁴³ Wenn mehr Messzeit zur Verfügung steht, sollte man unbedingt Mehrfachmessungen vornehmen oder mit einer Sonde arbeiten, die größere Flächen erfasst. Vgl. frühere Messungen mit der „Bügeleisensonde“ (NIRSystems SY-1711 (MC 6500), Steuersoftware VISION 1998).

Jahrhundert	13.	14.	15.	16.	17.	18.
Objekt	pm1 - pm5	pm6 - pm10	pm11 -pm15	pm16 - pm20	pm21 - pm26	pm27 - pm32
Anzahl	5 (7)	5	5	5	6	6
Personen	div. Bischöfe*	Bf. Gerhard von Schraplau	Bf. Johannes von Bose	Bf. Sigismund von Lindenau	Kurfürst Johann Georg I.	Herzog Moritz Wilhelm
Amtszeiten	1204-1265	1320-1340	1431-1463	1535-1544	1603-1656	1694-1731

* Dietrich von Meißen (1204-1215), Ekkehard (1216-1240), Rudolf von Webau (1240-1244) sowie Heinrich von Wahren (1244-1265)

Tabelle 5: Klassenzuordnung (Jahrhunderte), Amtszeiten der zugeordneten Personen

Zur besseren Identifizierung der einzelnen Objekte in grafischen Darstellungen und der leichteren fachlichen Verständigung verwenden wir direkt die Jahreszahl als Bezeichner für die Pergamente, weil diesen damit das Jahr unmittelbar grafisch, meist sogar eindeutig, zugeordnet werden kann.

Anhang 2: Ergebnis-Reproduktion

Um die berichteten Ergebnisse reproduzieren zu können, bedarf es des Einhaltens notwendiger technischer Datenvorbereitungsschritte, insbesondere bei der Auswertung mit dem hochkomplexen Programmsystem STATISTICA.

Datenvorbereitung MS Excel

Zur Umwandlung der Wellenlängenwerte (die Wellenzahlen werden nicht verwendet) sind diese zunächst mit 1.000.000.000 zu multiplizieren, dann zu runden ($f(x)=\text{runden}(Zelle*1000000000;0)$) und die entsprechende Spalte danach als Text zu formatieren (Arbeitsblatt ‚Spektren‘) und nochmals die Spalte mit Einfügeoption (Werte übernehmen). Der ausgewählte Bereich kann dann in ein neues Datenblatt (z.B. ‚transponiert‘) transponiert eingefügt (Einfügeoption oder Ctrl-Alt-V). Ggf. kann der beschriebene Zwischenschritt (Werte übernehmen) umgangen werden, indem gleich mit Ctrl-Alt-V kopiert wird und dann ‚Werte übernehmen‘ gewählt wird.

Datenvorbereitung STATISTICA

Über Datei/Öffnen wird die oben erzeugte Excel-Datei geöffnet. Wenn die Datei mehrere Blätter enthält, sollte sie im xls-Format abgespeichert sein, nicht xlsx verwenden! Ausgewählt wird das Datenblatt ‚transponiert‘; anzukreuzen ist: Fallnamen aus Spalte 1, Variablennamen aus Zeile 1, Zellenformatierung importieren.

Will man später andere Fallnamen verwenden, kann man schon in der Excel-Tabelle die erste Spalte entsprechend einrichten. STATISTICA bietet aber auch einen Fallnamen-Manager an.

Vor den Analysen sind ggf. Proben zu löschen bzw. bei den jeweiligen Analysen mit Auswahlbedingungen (z.B. $v_0 = \langle \text{Wert} \rangle$, $v_0 > 33$) auszuschließen: pm21 (extreme Ordinatenwerte wegen Spektren-Aufnahme mit ‚gain=extreme‘), pm33, pm33_2, pm34, pm34_2 (Papierproben) und ggf. pm25 (extrem dünnes Pergament, gegen untergelegten Reflexionsstandard aufgenommen).

Schließlich arbeiteten wir mit dem Datensatz „Proben Domstiftsbibliothek **fallbereinigt**“.

Ggf. ist es nützlich, die Daten des ersten Oberton- und des Kombinationsschwingungsbereichs getrennt auszuwerten (z.B. als Variablengruppen Nr. 3-81 und Nr. 82-235).

Datenauswertung STATISTICA/ Faktoranalyse

Die Korrelationsmatrix (mit allen Wellenlängen) ist schlecht konditioniert; offenbar gibt es nur 3 positive Eigenwerte, folglich nur drei extrahierbare Hauptkomponenten. Für die Graphik wurden rotierte Faktorwerte verwendet (Varimax, standardisiert).

Datenauswertung STATISTICA/ Diskriminanzanalyse

Man beachte bei der traditionellen Diskriminanzanalyse:

- Es sind nur max. ca. 85 Merkmale auswählbar.
- Bis ca. 65 Merkmale wird auf Modul **ADM** verwiesen.
- Bis ca. 24 Merkmale wird das Toleranz-Kriterium verletzt.
- Analyse ist nur im Aufbauverfahren möglich.
- Verfahren funktioniert nur bis zur Menge {2304, 2278, 2293, 2267, 2188}, dann wieder Verletzung des Toleranzkriteriums.
- Mit {2304, 2278, 2293, 2267} ist die Klassifikationsrate noch nicht perfekt.
- Die Klassifikationsmatrix ist nur während der Einzelschritte bis zum Abbruch erreichbar.
- Man folge deshalb gleich der Empfehlung, unter Statistik/Multivariate explorative Technik/ **Allgemeine diskriminanzanalytische Modelle (ADM)** auszuwählen.

Brauchbare Ergebnisse erhält man im Modul **Allgemeine diskriminanzanalytische Modelle** mit folgender Verfahrensweise:

Als Analysetyp: Traditionelle Diskriminanzanalyse im ‚Standarddialog‘ auswählen und mit [OK] bestätigen. Dann:

- im Register ‚Standard‘ Variablen auswählen: Abhängige Variable: century , Prädiktorvariablen: Alle (Wellenlängen) [OK]
- Im Register ‚Details‘ alle Standardeinstellungen übernehmen (A-Priori-Klassifikationswahrscheinlichkeiten: Geschätzt, Modellaufbau: Alle Effekte, Kreuzvalidierung: aus [OK]

Im erscheinenden Ergebnisfenster verfähre man wie folgt:

- Im Register ‚Funktionen‘ wird mit [Klassifikationsfunktions-Koeffizienten] ausgewiesen, welche Wellenlängen für die Diskriminanzfunktionen verwendet wurden.
- Im Register ‚Funktionen‘ erhält man auch die [Klassifikationsmatrix der Analysestichprobe], sie weist das Ergebnis für alle Klassen aus.
- Im Register ‚Fälle‘ kann die [Stichproben-Klassifikationsmatrix]ebenfalls eingesehen werden, interessanter ist es, nach Häkchen-Setzen bei ‚Klassifikationen‘ sich mit [Statistiken zeigen] die wahrscheinlichen Zuordnungen für alle Proben anzusehen
- Setzt man das Häkchen bei ‚Kanonische Werte‘, so erhält man mit [Statistik zeigen] die Probenkoordinaten für jede Diskriminanzfunktion (*root*), die zusammen mit weiteren Variablen der Datenmatrix über [Statistiken speichern] als Tabelle für spätere Plots gespeichert werden kann.
- Im Register ‚Effekte‘ können auch die ‚Koeffizienten‘ für den **Regressionsansatz** gefunden werden.

Anhang 3: Liste der Abbildungen und Tabellen

Abbildungen

Abb. 1: Pergament-Spektren [ARCSpectro ANIR, geglättet]	10
Abb. 2: Pergament-Spektren. [ARCSpectro ANIR, Absorption]	11
Abb. 3: Pergament- und Papier-Spektren. [Unscrambler, line plot]	11
Abb. 4: Pergament-Spektren. [Unscrambler, „fallbereinigt“]	12
Abb. 5: Clusteranalyse. Gesamter Spektralbereich. Dendrogramm nach Ward [Unscrambler]	22
Abb. 6: Clusteranalyse. Kombinationsschwingungen. Dendrogramm nach Ward [Unscrambler]	23
Abb. 7: Clusteranalyse. Obertonschwingungen. Dendrogramm nach Ward [Unscrambler]	24
Abb. 8: Clusteranalyse. Gesamter Spektralbereich. Dendrogramm nach Single Linkage	25
Abb. 9: Faktoranalyse. Alle Wellenlängen. Gruppierung der Messstellen	26
Abb. 10: Faktoranalyse. Alle Wellenlängen. Pergamente im Raum zweier Faktoren	27
Abb. 11: Faktoranalyse. Oberton-Bereich. Pergamente im Raum zweier Faktoren	28
Abb. 12: Faktoranalyse. Kombinationsschwingungen. Pergamente im Raum zweier Faktoren	28
Abb. 13: Diskriminanzanalyse. Pergamente im Raum zweier Diskriminanzmerkmale	31
Abb. 14: Diskriminanzanalyse. Pergamente im Raum dreier Diskriminanzmerkmale	31
Abb. 15: Multiple Regression. Diplom-Jahr vs. berechnete Jahreszahl	34
Abb. 16: Multiple Regression. Abweichungen der berechneten Jahreszahlen	35

Tabellen

Tabelle 1: Bereiche und Zuordnungen	15
Tabelle 2: Diskriminanzanalyse. Klassifikationsergebnisse	32
Tabelle 3: Multiple Regression. 20 Aufbauschritte	33
Tabelle 4: Prinzipieller Aufbau des Datentableaus	39
Tabelle 5: Klassenzuordnung (Jahrhunderte), Amtszeiten der zugeordneten Personen	40