

PIGTALE

Formale Modelle für die Entscheidungsunterstützung in der Lebensmittelproduktion

Kathrin Hausmann¹, Thomas Rose^{1,2}

¹ Fraunhofer FIT, Schloss Birlinghoven, D-53757 Sankt Augustin
{Kathrin.Hausmann, Thomas.Rose}@fit.fraunhofer.de

² RWTH Aachen University, Ahornstraße 55, D-52056 Aachen
{Rose}@i5.informatik.RWTH-Aachen.de

Abstract. Veränderungen in der Tierhaltung, wie beispielsweise die Stalltechnik, haben weitreichende Auswirkungen auf die Gesundheit der Tiere und die Effizienz der Haltungsprozesse. Gefordert sind formale Modelle, mit denen die Auswirkungen von Änderungen prognostiziert werden können. Heutige Prognosen sind vielfach qualitativer Natur. Dieses Papier stellt einen quantitativen Ansatz in den Mittelpunkt, der mit statistischer Analytik Kausalitäten zwischen Haltungsentscheidungen und Tiergesundheit ableitet. Hiermit lassen sich nicht nur Entscheidungen der Tierhaltung ökonomisch bewerten, sondern auch der Faktor Tierwohl rationalisieren. Konkret illustrieren wir unseren Ansatz mit Prognosemodellen für Lungenerkrankungen bei Mastschweinen. Unsere Modellentwicklung basiert auf operativen Datensätzen renommierter Stakeholder der Schweinemast, die ca. 80 Millionen Tiere pro Jahr in Deutschland produzieren. Ökonomische Effizienz und Tierwohl sind keine Kontrahenten, sondern genetische Zwillinge bei quantitativer Analyse und damit rationaler Betrachtung.

Keywords: Optimierung von Produktionsprozessen, evidenz-basierte Entscheidungsunterstützung, Rationalisierung von Tierwohl, erzeugerorientiertes Feedback von Haltungsänderungen

1 Einleitung

Die Wertschöpfungskette von Schweinefleisch ist komplex und die Prozessgestaltung sehr heterogen. Verteilt über verschiedene Parteien sind sehr unterschiedliche Prozessdaten zu integrieren. Neben der Fleischqualität legt der Verbraucher zunehmend mehr Wert darauf, dass die Tiere unter guten Bedingungen aufgezogen und transportiert werden. Um dies prüfen und dokumentieren zu können, wird der gesundheitliche Status der Schweine während des Schlachtprozesses durch tierärztliche Kontrollen erfasst. Die Rückmeldung dieser für den Landwirt essentiellen Informationen, um Verbesserungen hinsichtlich Tiergesundheit und Ökonomie treffen zu können, erfolgt

momentan jedoch sehr unzureichend. Neben der jeweils auf Lieferebene kumulierten Rückmeldung der ökonomischen Kennzahlen und Befundungen durch die Schlachthöfe, gibt es lediglich Tools, die es dem Landwirt erlauben, die Leistung seines Betriebs mit denen anderer Landwirte zu vergleichen. Ein Beispiel hierfür ist das Portal der IQ-Agrar¹, mit dem sich Daten einzelner oder akkumulierter Lieferpartien visualisieren und analysieren lassen (Abbildung 1).



Abbildung 1: Portal zur Visualisierung der Ergebnisse des Mastprozesses

Neuartig an der Idee von PIGTALE ist die Entwicklung eines formalen Modells über kausale Zusammenhänge zwischen Haltungentscheidungen und ihrer Auswirkungen auf das Tier. Hiermit wird Wissen über die Tierhaltung erzeugt, was über reine Vergleiche mit anderen Erzeugern hinaus geht, wie in Abbildung 1 illustriert. Es entsteht die Möglichkeit, Auswirkungen von Investitionen in Management und Infrastruktur auf die Gesundheit der Masttiere zu prognostizieren und damit gezielt gesündere Schweine zu produzieren. Landwirte können durch die Rückmeldung relevanter Informationen die Qualität der Prozesse steigern, Gefahrenpotentiale vermeiden sowie die Produktivität erhöhen.

Für die Entwicklung unserer quantitativen Modelle sind prozessbezogene Daten verschiedener Parteien entlang des Produktionsprozesses zu integrieren. Durch diese prozessübergreifende Sicht können nun Entscheidungen zu Änderungen mit ihren Auswirkungen prozessorientiert analysiert werden. Die Datenquellen umfassen Informationen über die Haltung der Tiere und wirtschaftlichen Kennzahlen der Mastbetriebe, Daten der betreuenden Bestandsveterinäre zur Optimierung der Tierhaltung sowie Befunddaten der Schlachthöfe als letztendliche Indikatoren sowohl für den ökonomischen als auch gesundheitlichen Haltungserfolg. Durch Analyse einzelner Datenquellen wäre eine Ableitung kausaler Zusammenhänge nicht möglich, da der prozessübergreifende Charakter der Daten und der Wirkzusammenhänge fehlen.

¹ <https://www.iq-agrar.de/>

Neben den Landwirten können weitere Akteure der Lebensmittelkette durch den Ansatz von PIGTALE profitieren. Beispielsweise bietet dieser die Möglichkeit, Prognosemodelle zu erstellen, die die Tiergesundheit in Abhängigkeit konkreter Haltungsbedingungen modellieren. Damit wäre die Bewertung des Tierwohls rational indiziert.

Es stellt sich ebenso die Frage nach der weiteren Verwertbarkeit dieses neuen Rohstoffes: Wissen über kausale Abhängigkeiten zwischen Haltung und Tiergesundheit wie letztendlich bei der Schlachtung diagnostiziert. Mit diesem Wissen lassen sich verschiedene Services für den Landwirt rund um die Tierhaltung entscheidend verbessern. Beispielsweise können Entscheidungen über Veränderungen der Haltungsbedingungen durch Investitionen in die Stalltechnik ökonomisch modelliert werden (*Auswirkungsanalysen*). Des Weiteren lassen sich Tieruntersuchungen auf dem Schlachthof risikoorientiert optimieren, da aus den bekannten Haltungsbedingungen mögliche Krankheitsbilder abgeleitet werden können (*Problemidentifikation*). Auf Grundlage der Haltungsbedingungen könnten Schlachttiere bereits vor dem Transport zum Schlachthof auf ihr Risiko hin bewertet werden, was nach unserem Wissensstand zum jetzigen Zeitpunkt noch nicht praktiziert wird. Zudem ermöglicht dieses Wissen dezidierte Prognosen der Maßnahmen zur Tiergesundheit. Ankerpunkt von PIGTALE ist die formale Modellierung der Abhängigkeiten in der *Produktionskette Schweinemast*, um die Entscheidungsunterstützung durch Auswirkungsanalysen zu optimieren.

In diesem Beitrag stellen wir PIGTALE als Proof of Concept (PoC) für eine evidenzbasierte Entscheidungsunterstützung auf Basis quantitativer Modelle vor. Die Modellbasierung von PIGTALE steht damit im Kontrast zu eminentbasierten Ansätzen, die auf rein qualitativen Betrachtungen basieren.

Dieser Beitrag ist folgendermaßen strukturiert. Kapitel 2 zeigt zunächst die Bedeutung der formalen Modelle für die Entscheidungsunterstützung. Kapitel 3 zeichnet den Umfang des verwendeten Datenpools, während Kapitel 4 die Methodik der Modellgenerierung herausstellt. PIGTALE stellt den zentralen Punkt eines Ökosystems für die Optimierung der Produktionsprozesse in der Lebensmittelindustrie dar. Die Elemente dieses Ökosystems für eine nachhaltige Umsetzung werden in Kapitel 5 vorgestellt, bevor Kapitel 6 noch einmal unseren Ansatz als evidenzbasierte Entscheidungsunterstützung für die Optimierung der Fleischproduktion zusammenfasst.

2 Modellbasiertes Wissen der Optimierung von Haltungsentscheidungen

Die Explizierung von Wissen ist seit jeher eine große Herausforderung vieler Ansätze zum Wissensmanagement. Mit ihrem SECI-Model (socialiation, externalizaion, combination, internalization) haben die Autoren (Nonaka & Takeuchi, 1995) grundlegende Arbeiten zur Transformation von implizitem zu explizitem Wissen und seiner Internalisierung gelegt, d.h. wie kann die Entscheidungsunterstützung in Änderungsprozessen auf eine rationale Basis gestellt werden.

In der Tierzucht liegt vieles Wissen ebenfalls auf individueller Basis vor und ist organisatorisch im Sinne von Fähigkeiten (*Capability Maturity Modelling*) zu generalisieren (Paulk & et. al., 1993). Ein Wissenskreislauf ließe sich hier also auch anwenden, um Wissen für die Tierhaltung zu externalisieren und es letztendlich für Optimierungsprozesse zu kombinieren und zu internalisieren. Im Kern sind die Beobachtungen in der Tierhaltung aber eher qualitativer Natur. Erkenntnisse quantitativer Natur sind eher dünn gesät und somit ökonomische Analysen bisher für eine Optimierung der Produktionsprozesse nur schwerlich möglich.

Der direkte Zusammenhang zwischen dem Gesundheitsstatus eines Nutztierbestandes einerseits und dem Tierwohl im Sinne von „animal welfare“ andererseits ist mittlerweile unbestritten. Neben anderen Faktoren schafft die Reduktion von Produktionskrankheiten die Voraussetzungen für ein hohes Maß an Wohlbefinden (Blaha & Sundrum, 2017) der gehaltenen Nutztiere. Als mögliches Maß zur Bewertung der Tiergesundheit in den Schweinebeständen und damit die Schaffung eines objektiven Tierwohlintegrators wird seit langem die Auswertung der Schlachtbefunderhebung diskutiert. Bei allen Problemen hinsichtlich der Vergleichbarkeit der erhobenen Daten, belegen einzeltierbezogene Auswertungen der biologischen Leistungen und der korrespondierenden Schlachtbefunddaten eindrucksvoll diesen Zusammenhang.

Von ca. 100.000 im Jahr 2017 von IQ-Agrar mittels RFID-Transpondern gekennzeichneten Schweinen konnten folgende Verluste von Gewichtszunahmen im Vergleich zu gesunden Tieren nachgewiesen werden (Hartmann, 2018):

- 74g geringere Tageszunahme für Tiere mit Befund,
- 100g geringere Tageszunahme für Tiere mit mehr als einem festgestellten Befund,
- 140g geringere Tageszunahme bei festgestellten hochgradigen Leberbefunden oder Brustfellentzündungen.

Die betroffenen Tiere waren also einem erheblichen Krankheitsstress und somit Schmerzen, Leiden und Schäden ausgesetzt. Verursacht wird diese Leistungseinbuße durch den erhöhten Nährstoffaufwand für die Aufrechterhaltung des Immunsystems im Krankheitsfall. Da die Schweine vor allem während der Jugendentwicklung Muskelmasse bilden und in der späteren Entwicklung der Fettansatz dominiert, belegen Studien direkte Zusammenhänge zwischen Merkmalen der Tiergesundheit und Parametern der Schlachtkörperzusammensetzung sowie der Fleischqualität (Klauke, et al., 2013). Diese ursprünglich rein ökonomischen Bewertungsansätze erlauben somit ebenfalls objektive Bewertungen der Tiergesundheit in den Beständen und somit Rückschlüsse auf die Tierwohlsituation. Tierwohl und Effizienz eines Haltungsprozesses müssen also keine konfligierenden Ziele sein, sondern sind eher komplementär zu betrachten.

Die Ursachen bei auffälligen Beständen sind häufig vielfältig und selten monokausal auf einzelne Haltungsbedingungen zurückführbar. Zwar existieren Publikationen, die einzelne Haltungsentscheidungen hinsichtlich ihres Effektes auf die Tiergesundheit analysieren (Hornauer, 2001) (Schumacher & et. al., 2011), jedoch beruhen diese Analysen zumeist auf qualitativen Beobachtungen.

PIGTALE integriert als PoC einen prozessübergreifenden Datentopf entlang der Produktionskette, der es erstmalig erlaubt in großem Stil statistisch belegte, quantitative Rückschlüsse aus realen Datenbeständen zu ziehen und zu kommunizieren.

3 Datenquellen

Während der Prozesskette vom Landwirt bis zum Schlachthof wird eine Vielzahl an Informationen erfasst, die im Rahmen dieses Vorhabens in eine Datenbank fließen, die wiederum die Basis für ein Softwaretool darstellt, das in der Lage ist, dem Landwirt automatisiert mögliche Handlungsstrategien zur Steigerung des Tierwohls und der Wirtschaftlichkeit vorzuschlagen.

Um das Konzept exemplarisch zu verdeutlichen wurden Projektpartner akquiriert und eine Datengrundlage geschaffen, welche große Teile der Produktionskette in Westfalen widerspiegelt (Abbildung 2).

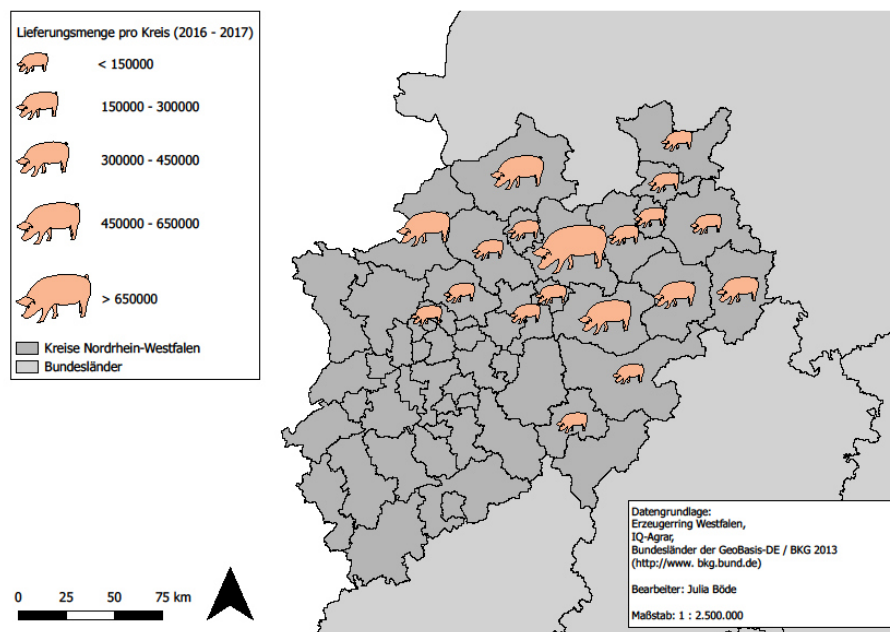


Abbildung 2 Anzahl der analysierten Schweine nach Kreisen

Informationen über die gesamten Lebensumstände der Mastschweine liefert der Datenbestand des Erzeugerrings Westfalen². Beginnend mit Herkunfts- und Abstammungsmerkmalen werden Informationen über die Haltungsumstände der Tiere erfasst. Die Kenntnisse über die Haltungsform, die technische Ausstattung der Betriebe,

² <https://www.erzeugerring.com/>

das Fütterungskonzept, einschließlich der eingesetzten Technik und Futtermittel, sowie weitere Managementaspekte machen es möglich, die Haltungsbedingungen der Masttiere zu modellieren.

Die bei der Schlachtung ante und post mortem erhobenen Merkmale wie Gewicht und Klassifizierungsergebnisse sowie insbesondere die amtlichen Befundungen (LmH), bieten eine Retrospektive auf den Verlauf der Mast und den Gesundheitszustand der Tiere. Diese Daten werden bei IQ-Agrar zentral verwaltet. Die Daten umfassen allgemeine Produktionsdaten (QS-Auditierung, Ergebnisse des Salmonellen- und Antibiotikamonitorings). Zudem enthält der Datentopf auch ergänzende Befundungen, welche am Schlachthof erhoben werden (Abbildung 3).

Im Ergebnis entsteht ein Datenpool, der die unterschiedlichen Lebensphasen (Life Cycle) der Schweine widerspiegelt und die Grundlage für eine Modellierung der gesundheitlichen Entwicklung eines Schweins von der Geburt bis zur Schlachtung bietet (Vorhersagemodell).

4 Methodischer Überblick der Modellentwicklung

4.1 Theoretische Grundlagen

Aus methodischer Sicht gilt es, die anspruchsvolle Datenstruktur (Ebene des Einzelschweins, Ebene der Lieferpartie, Ebene des Mastbetriebs, Ebene des Schlachthofes) möglichst umfänglich zu erfassen. Hierfür wurden die einzelnen Datensätze je Lieferpartie aggregiert. Die Befundungen wurden dabei prozentual an der Gesamtzahl der Schweine in einer Lieferpartie gespeichert. Nach mehreren Aufarbeitungsschritten erhielt man somit einen Datensatz mit ca. 26.000 vollständigen Lieferpartien.

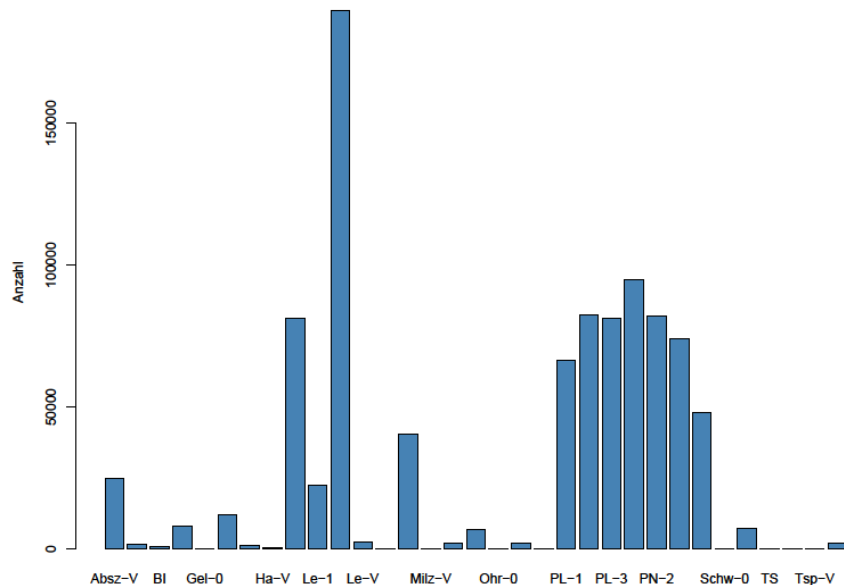


Abbildung 3 Befunde in integrierter Datenbasis³

Die Datenbasis weist diverse interne Abhängigkeiten (Korrelationen) auf. So hängt die Performance einer Partie vom jeweiligen Mastbetrieb, dem Schlachthof aber auch von saisonalen Einflüssen ab. Die zeitliche Komponente und komplizierte Struktur der Daten erfordert eine nicht-lineare Modellierung über ein Tobit Modell (Tobin, 1958), welches unter Berücksichtigung der Panel Struktur ein komplexes mathematisches Problem stellt.

Modelliert wird die Häufigkeit eines Lungenbefundes (genannt y) innerhalb einer Lieferpartie, d.h. die Variable y ist beschränkt zwischen 0 und 1. Es wird für jeden Landwirt i eine eigene Häufigkeit geschätzt, wobei die erklärenden Variablen \mathbf{x}_i mit Koeffizienten $\boldsymbol{\beta}$ multipliziert werden. Um die Häufigkeit zu schätzen, wird zunächst eine Hilfsvariable y_i^* geschätzt, welche keine Einschränkung auf 0 und 1 hat. Im Anschluss werden Werte, die außerhalb $[0,1]$ liegen auf 0 gesetzt.

$$y_i = \begin{cases} 0, & y_i^* < 0 \\ 1, & y_i^* > 1 \\ y_i^*, & y_i^* \in [0,1] \end{cases}$$

Da der Datensatz zusätzlich eine zeitliche Komponente aufweist (Lieferungen erfolgen nacheinander), wird das Modell um die Zeit t ergänzt. Um eine konsistente Schätzung zu erhalten, wird ein Ansatz mit „Random Effects“ gewählt (Baltagi, Song, & Koh, 2003). Ein Modell mit Fixed Effects (Borenstein & et. al., 2010) wurde zwar ebenfalls entwickelt, ist aber nicht Teil dieser Ausführung.

³ Die gängigen Abkürzungen können beispielsweise in (Meemken & et. al., 2016) nachgeschlagen werden.

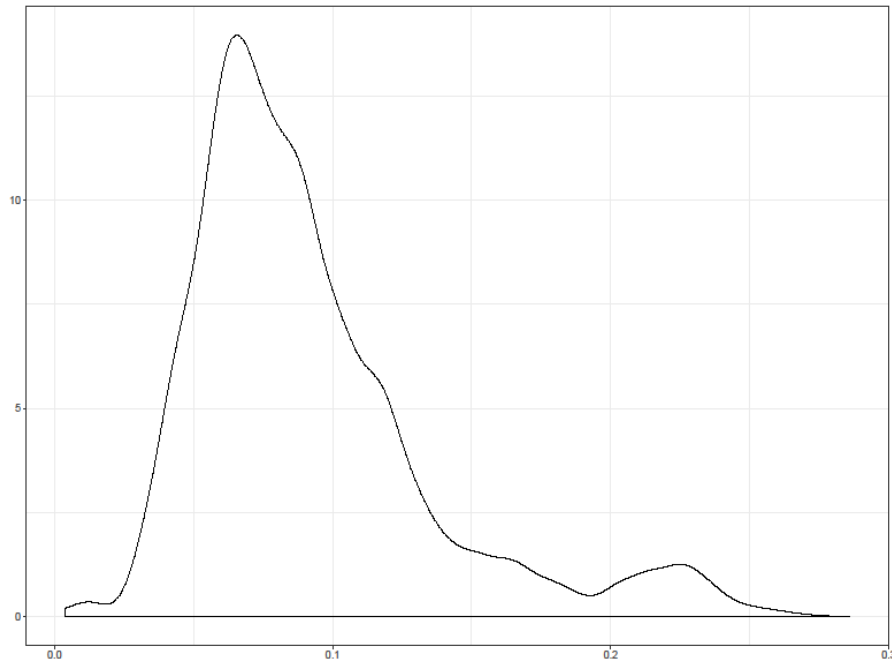


Abbildung 4 Verteilung der Häufigkeiten von Lungenbefunden

Jedes Individuum, in unserem Fall jeder Landwirt, erhält eine zusätzliche Variable, die seine persönlichen Präferenzen widerspiegelt μ_i . Diese Präferenzen sind normalverteilt. Das komplexe Modell folgt den Abhängigkeiten:

$$y_{it}^* = \mathbf{x}_{it}'\boldsymbol{\beta} + \mu_i + \varepsilon_{it}.$$

Henningsen diskutiert die Anwendung dieses Modells und dessen Einschränkungen (Henningsen, 2010). Mithilfe dieser funktionalen Form und einer Schätzung durch Maximum Likelihood Methoden, lassen sich die Häufigkeiten schätzen (Abbildung 4). Um den Umfang der Kausalitäten zu quantifizieren, wurden nach (Greene, 2004) die marginalen Effekte berechnet.

4.2 Ausgewählte Auswertungsergebnisse

Als *Proof of Concept (PoC)* wurde eine Analyse der Lungenbefunde durchgeführt, welche unterschieden nach drei Schweregraden in den Daten vorliegen:

- Lunge bis zu 10% verändert (PN 1),
- Lunge zwischen 10% und 30% verändert (PN 2),
- Lunge über 30% verändert (PN 3).

Pneumonische Veränderungen (PN 1-3) gehören zu den häufigsten Befunden am Schlachthof. Auch die Schwere der Veränderung ist über alle drei Kategorien gleichmäßig verteilt (Abbildung 5). Die Integration der diversen Datentöpfe über den Zeitraum von einem Jahr ergab einen exemplarischen Datensatz

über rund 1,9 Mio. Schweine, die entweder gesund oder von einer Lungenkrankheit betroffen sind.

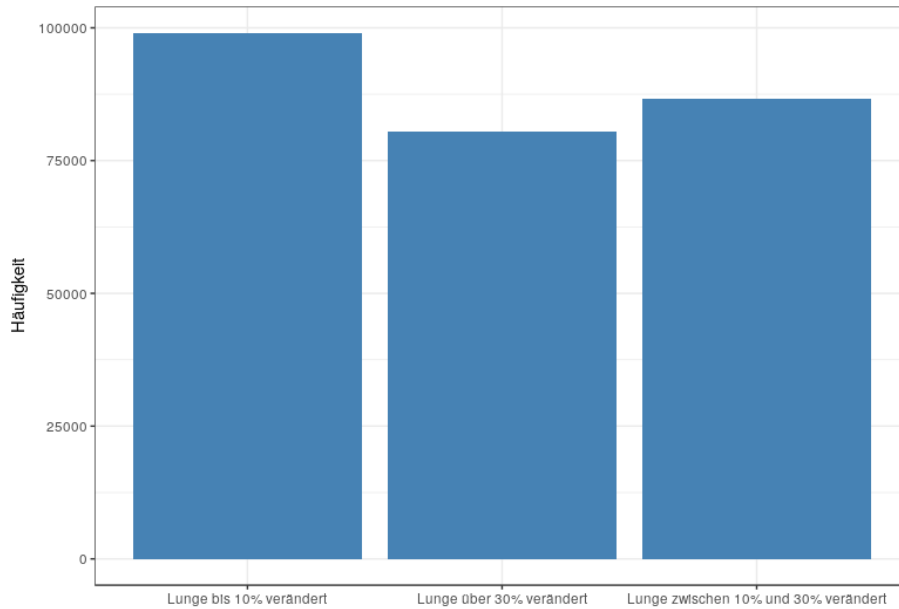


Abbildung 5 Anzahl der Lungenbefunde nach Schweregrad

Als erklärende Variablen wurden die Bereiche **Genetik** (Sauenlinie, Eberlinie, Herkunft der Ferkel), **Ernährung** (Hauptfutterkomponente, Fütterungstechnik, Phasenfütterung), **Stallausstattung** (Heizungstechnik, Aufstellungsart, Wasserangebot) sowie **Stallmanagement** (Gruppengröße, Geschlechtertrennung, Buchtfläche) zu Grunde gelegt. Nach weiteren Anpassungen des Modells wird es möglich sein, konkrete Rückschlüsse einzelner erklärender Variablen auf die Tiergesundheit und damit das Tierwohl zu ziehen.

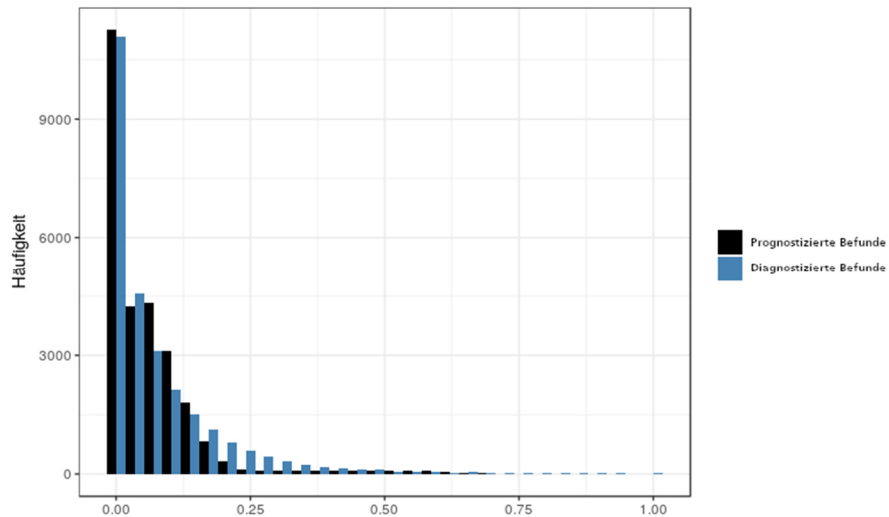


Abbildung 6 Test des Prädiktiven Modells

Im Rahmen dieser Initiative wurde ein prädiktives Modell aufgestellt und verifiziert, welches das Auftreten von Lungenveränderungen beschreibt. Für jede neue Lieferpartie von Schweinen kann hiermit der Anteil lungenkranker Schweine geschätzt und vorhergesagt werden. Der Datensatz wurde in Trainingsdaten und Testdaten geteilt. Abbildung 6 zeigt für den Testdatensatz, wie gut die Verteilung der Lungenkrankheiten bereits jetzt getroffen wird. Zum Vergleich werden auch die tatsächlichen Lungenbefundhäufigkeiten dargestellt. Die Anwendungsmöglichkeiten solcher prädiktiven Modelle ist vielfältig und bietet Raum zu Erweiterungen und Generalisierungen. Die Vorhersagekraft ist mit $R^2 \approx 40\%$ bereits jetzt auf einem Level, das vielversprechende Ergebnisse aufzeigt.

4.3 Potential des PIGTALE Ansatzes

Das mit PIGTALE abgeleitete Wissen bietet eine Vielzahl an Anwendungsmöglichkeiten. Vorwärts angewendet lassen sich mit dem integrierten Datenpool haltungsbedingte Schlüsse auf Fleischqualität und Ertrag pro Schwein ableiten. Rückwärts angewendet lassen sich aus den Untersuchungsergebnissen im Schlachthof Verbesserungspotentiale in Haltung und Stallausstattung ziehen. Weitere Arbeiten haben die folgenden Ziele vor Augen:

- **Entwicklung von objektiven Tierwohlindikatoren** – Zusammenfassung und Gewichtung diverser Tierwohlindikatoren. Analysen verschiedener Gesundheitsindices und Auditberichte hat hier eine Varianz gezeigt, die zu minimieren gilt, wenn Indices als Mehrwert vermarktet werden sollen.
- **Nutzung des generierten Wissens für die amtliche Schlachtier- und Fleischuntersuchung** – Dank des Wissens über die Zusammenhänge zwischen

Haltungsbedingungen und Gesundheit können Untersuchungsabläufe nach der Schlachtung evidenzbasiert optimiert werden, was aber der Abstimmung mit den Veterinäraufsichten bedarf.

- **Wahrung der Souveränität der Datenerzeuger in der Tierhaltung** – Die Nutzung der Daten bedarf natürlich der Zustimmung der Landwirte, die im Kern die Haltungsdaten erzeugen. Hier sind Konzepte für die Dateneinspeisung zu entwickeln, welche die betrieblichen Interessen der Landwirte wahren.
- **Incentivierung** – Die von unterschiedlichen Stakeholdern bereitgestellten Daten stellen wertvolle produktive Ressourcen dar, deren Nutzung zur Optimierung der Produktionsprozesse incentiviert werden muss. Gleiches gilt natürlich für die Analytik zur Ableitung formaler Modelle für die Wissensgenerierung.
- **Automation der Datenerhebung** – Essentieller Rohstoff für PIGTALE sind Daten aus dem Life Cycle Prozess der Schweinemast. Teils lassen sich diese Daten aus bestehenden Managementsystemen für die Tierhaltung extrahieren. In anderen Teilen bietet sich aber die Integration einer Sensorik (Internet of Things) oder eine automatische Bildklassifikation an, um beispielsweise die Schlachtkörperuntersuchung durch Bilderkennung zu rationalisieren.

PIGTALE hat gezeigt, dass auf der Basis integrierter Datentöpfe der Life Cycle übergreifenden Tierhaltung quantitative Evidenzen für die Auswirkungen der Haltung auf das Tierwohl und damit letztendlich auch die ökonomischen Erfolge nachhaltiger Tierhaltung gewonnen werden können.

Voraussetzung hierfür ist die Integration der Datentöpfe verschiedener Stakeholder in einem gemeinsamen Datenraum, wobei die individuellen Souveränitäten zu wahren sind. Die Schaffung dieses gemeinsamen Datenraums steht somit im Fokus zukünftiger Arbeiten zu Tierwohl und ökonomischer Effektivität in der Tierhaltung.

5 Konzeptionelle Gestaltung eines Ökosystems für PIGTALE

PIGTALE illustriert die Entwicklung formaler Modelle für die Optimierung von Produktionsprozessen in der Fleischproduktion in Form eines Proof of Concept. Dieser PoC basiert auf einer genossenschaftlichen Idee zur gemeinsamen Datenbereitstellung, Analyse und Wissensnutzung. Grundlage ist die Integration prozessübergreifender Daten im Lebenslauf eines Tierbestandes. Ziel ist die Demonstration der Machbarkeit eines datengetriebenen Wissensmanagements, aus Rohstoffen (Produktionsdaten) mit geeigneten Werkzeugen (statistische Analytik) marktfähige Produkte für die Landwirtschaft (Regeln zum Änderungsmanagement) zu entwickeln. Daten werden somit selbst Quelle einer Monetarisierung (Rose T. , Gruber, Hausmann, & Osterland, 2021).

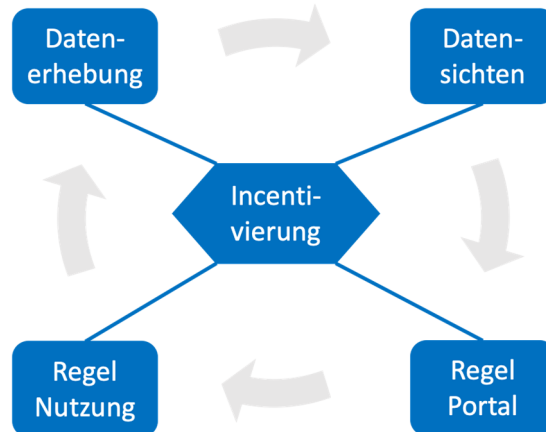


Abbildung 7: Das Ökosystem zur Nutzung der Modelle von PiGTALE

Unser PoC baut auf der Gemeinschaft eines Netzwerks von Kernpartnern entlang der Produktionskette auf. Für eine weitere Öffnung dieses Ansatzes ist ein Ökosystem mit einem gemeinsamen Geschäftsmodell zu gestalten, in dem Rohstoff- und Werkzeuglieferanten für ihre Beiträge von Produktkonsumenten entlohnt werden. Abbildung 7 zeigt die Kernkomponenten dieses Ökosystems (Rose, Gruber, & Hausmann, September 2020). Im Hintergrund stehen immer Dienste der Analytik, mit denen Regeln zur Optimierung aus Datenpools abgeleitet werden können. Zudem werden Kernkomponenten für fünf Aufgaben benötigt:

- *Incentivierung* – Im Zentrum dieses Ökosystems steht ein Geschäftsmodell, mit dem die Bereitstellung von Rohstoffen (Daten der Prozesskette) sowie die Generierung von Mehrwerten durch Analytik (Werkzeuge) incentiviert werden.
- *Datenerhebung* – Daten bilden den Ausgangsstoff unseres Wissenskreislaufs. Datenanbieter müssen aber immer Souverän ihrer Quellen bleiben. Nur unter Wahrung strikter Datensouveränität lässt sich ein Marktplatz für Optimierungswissen umsetzen, ohne die Interessen Beteiligter potentiell zu verletzen.
- *Datensichten* – Je mehr Parteien prozessbegleitende Daten bereitstellen, umso mehr stellt sich die Frage nach der Konfiguration verschiedener Sichten für die Analytik. In einer Sicht sollen beispielsweise nur die Futtermittel eines Herstellers auf Wirkung überprüft werden, wohingegen zur Analyse von Gegenmaßnahmen bestimmter Erkrankungen nur Teilsichten aller Datensätze berücksichtigt werden sollen. Zudem können auch Abonnentenmodelle realisiert werden.
- *Regelportal* – Über ein Portal können die Regeln von verschiedenen Maklern über spezifische Kanäle unterschiedlichen Marktsegmenten angeboten werden. Die Regeln werden somit zu einem handelbaren Gut, dessen Bereitstellung durch Makler zu incentivieren ist.
- *Regelnutzung* – Die Nutzung der Regeln ist durch eine geeignete Governance zu steuern. Da die Nutzung einer erworbenen Regel für jeden einzelnen Tierbestand

einen Mehrwert darstellt, sind individuelle Anwendungen zu dokumentieren und zu entlohnen. Regeln können also nicht im Sinne von Zugriffs- oder Lizenzrechten genutzt werden, sondern individuelle Zugriffe sind zu protokollieren.

Die statistische Analytik ist bereits als Dienstleistung orthogonal zu diesen Elementen zu Beginn genannt worden. Ergänzend dazu basiert das Ökosystem gleichermaßen auf der Distributed Ledger Technology (DLT) für die Prozessunterstützung über alle Komponenten hinweg (Prinz, Rose, Osterland, & Putschli, 2019). Da sich im Rahmen der Digitalisierung von Prozessen zunehmend Fragen nach Identitäten beteiligter Parteien (wer war der Auditor mit welcher Qualifizierung) und des Vertrauens in ein Datum (mit welcher Güte wurden Handlungsbedingungen erhoben) stellen, ist ein Identitätsmanagement von Daten, Akteuren und Prozessen zunehmend wichtiger (Osterland & Rose, September 2020). Eine Blockchain protokolliert nicht nur die Nutzung von Regeln für die Optimierung der Produktionsprozesse, sondern auch die Verwertung von Daten für die Analytik oder die Bereitstellung von Regeln in Portalen. DLT wirkt somit als umfassendes Clearing House für Rohstoffe, Werkzeuge und Verwertungen. Zudem wird die Verwaltung der Daten über eine Variante des Industrial Data Space (Otto, et al., 2016) realisiert, um die Souveränität der Anbieter zu wahren und die Nutzung der Daten über Zugriffspolitiken steuern zu können. Alternativ können sich auch Datenplattformen wie Gaia-X⁴ entwickeln.

6 Zusammenfassung und Ausblick

PIGTALE illustriert die Ableitung formaler Modelle für die Entscheidungsunterstützung in der Fleischproduktion am Beispiel von Mastschweinen. Es basiert auf statistischer Analytik zur Generierung von Wissen über kausale Abhängigkeiten zwischen Handlungsentscheidungen in produzierenden Betrieben und den Befundungen am Schlachthof. Schlachtbefunde beschreiben einerseits den Gesundheitszustand der Tiere (*Tierwohl*), haben aber gleichzeitig bedeutende ökonomische Auswirkungen für die Erzeuger (*Haltungsökonomie*), da Schlachthöfe über *Masken* ihre Erwartungen an angelieferte Tiere (Anteile von Fett gegenüber Muskelfleisch oder Anteil Schinken am Körper) aus ihrer Vermarktungssicht formulieren. Wissen über die Auswirkungen von Handlungsentscheidungen kann somit ökonomisch für die Erhöhung der Haltungseffizienz als auch des Tierwohls als Vermarktungsmöglichkeit genutzt werden.

Zudem gibt es im Ökosystem der Schweineproduktion noch verschiedene Möglichkeiten der Verwertung dieses Wissens, wie beispielsweise die Beratung von der Futtermittelmischung durch Mühlen oder von Investitionen in der Stallgestaltung bis hin zu einer risiko-orientierten Befundung in der Schlachtkörperuntersuchung.

In diesem Beitrag haben wir uns auf die Modellentwicklung zur Verbesserung der Entscheidungsunterstützung in der Lebensmittelproduktion konzentriert. Kernaussage ist die Verbesserung der Entscheidungsunterstützung von Handlungsänderungen, wenn

⁴ <https://www.data-infrastructure.eu/GAIA/Navigation/EN/Home/home.html>

quantitative Methoden angewendet werden, anstatt die Haltung eminentbasiert zu bewerten. PIGTALE repräsentiert somit ein evidenzbasiertes Verfahren der Entscheidungsunterstützung, das das Wissen über Auswirkungen von Haltungsentscheidungen in der Produktion transparent wiedergibt und für verschiedene Parteien einen greifbaren Nutzen dieses Wissens darstellt.

Wir haben damit *Data Science & Business Analytics* in eine industrielle Domäne implantiert, die in maschinell-orientierten Bereichen wie dem Precision Farming oder dem Monitoring von Produktionsketten zwar bereits Digitalisierung und Big Data werkzeuggetrieben vorantreibt, in tiernahen Bereichen aber immer noch emotional beeinflusst wird. Neben einer Rationalisierung der Entscheidungen im Hinblick auf Tierwohl als Marketingoption lassen sich Haltungsprozesse auch ertragsorientiert steuern.

Wenn dieses Wissen als *Bonanza* in den Mittelpunkt eines Ökosystems für die Datenbewirtschaftung zur Optimierung von Produktionsprozessen gestellt wird, stellen sich Fragen für eine nachhaltige Umsetzung:

- Datenraumkonzept für die Wahrung von Daten- und Analysesouveränität,
- Incentivierung der Bereitstellung von Daten und Analyseergebnissen,
- Distributed Ledger Technology für die Abrechnung und Entlohnung von Wissensbereitstellung sowie
- Geschäftsmodell für die Governance des Betriebs einer vertrauensvollen Plattform.

PIGTALE hat hierfür als Proof of Concept die Basis gelegt, dass Information und Wissen die Wahrung der Digitalisierung⁵ sind und *nicht* Rohdaten wie von einigen Marktplatzplattformen entgegen tradierter Erfahrungen⁶ propagiert.

7 Acknowledgements

Diese Arbeit wäre nicht möglich gewesen ohne die Bereitstellung der umfangreichen Datenbestände (teils konkurrierender) Parteien aus operativen Prozessen und die fachliche Unterstützung zur Identifikation signifikanter Attribute und kritischer Reflektionen von Zwischenergebnissen aus verschiedenen Forschungssichten: Veterinärmedizin, Tierzucht, Farmmanagement, etc.. Unser besonderer Dank gilt hierfür Peter Schwaer, Stephanie Bering, Tobias Horn (IG Agrar), Thorsten Klauke (Erzeugerring Westfalen), Martin Frettlöh und Tanja Haag (QuH Lab). Zudem danken wir der b-it Foundation für ihre Unterstützung⁷ und unseren Kolleginnen Marlene Bubb und Julia Gruber.

⁵ Carly Fiorina, former CEO of Hewlett-Packard, Oracle Open World, San Francisco, 2004 <http://www.hp.com/hpinfo/execteam/speeches/fiorina/04openworld.html>

⁶ Analog zum Rockefellerprinzip (<https://de.wikipedia.org/wiki/Rockefellerprinzip>) ist nicht der Rohstoff das Ziel, sondern seine Nutzung: *Verkauft Lampen und nicht Öl.*

⁷ <http://www.b-it-center.de/>

Referenzen

1. Baltagi, B., Song, S., & Koh, W. (2003). Testing panel data regression models with spatial error correlation. *Journal of econometrics*, 1, S. 123-150.
2. Blaha, T., & Sundrum, A. (2017). Tierärztliche Kompetenz und Zielorientierung erforderlich! *Deutsches Tierärzteblatt*, 65(11), S. 1518-1521.
3. Borenstein, M., & et. al. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research synthesis methods*, 2(1), 97-111.
4. Greene, W. (2004). Fixed effects and bias due to the incidental parameters problem in the Tobit model. *Econometric reviews*, 23.2, S. 125-147.
5. Hartmann, F. (2018). Jahresbericht IQ-Agrar 2017.
6. Henningsen, A. (2010). Estimating censored regression models in R using the censReg Package. (R. p. vignettes, Redakteur)
7. Hornauer, N., Haidn, B., & Schön, H. (2001). Tiergesundheit im Aussenklima-Kistenstall für Mastschweine. *Agrartechnische Forschung*, 7(1).
8. Klauke, T., Piñeiro, M., Schulze-Geisthövel, S., Plattes, S., Selhorst, T., & Petersen, B. (2013). Coherence of animal health, welfare and carcass quality in pork production chains. *Meat Science*, 95(3), S. 704-711.
9. LmH. (kein Datum). Allgemeine Verwaltungsvorschrift über die Durchführung der amtlichen Überwachung der Einhaltung von Hygienevorschriften für Lebensmittel und zum Verfahren zur Prüfung von Leitlinien für eine gute Verfahrenspraxis.
10. Meemken, D., & et. al. (2016). *Harmonisierung der Schlachtbefunderfassung und -dokumentation bei der Schweineschlachtung in Niedersachsen*.
11. Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford university press.
12. Osterland, T., & Rose, T. (September 2020). Decentralised identity management for DLT-based cooperation support – Enabling new business relationships for inland waterway transport. *Proc. EnviroInfo 2020 Digital Twins for Sustainability*. Nikosia, Zypern.
13. Otto, B., Auer, S., Cirullies, J., Jürjens, J., Menz, N., Schon, J., & Wenzel, S. (2016). *Industrial Data Space-Digitale Souveränität über Daten*. (F.-G. z. eV, Hrsg.) München.
14. Paulk, M., & et. al. (4. Juli 1993). Capability maturity model, version 1.1. *IEEE Software*, 10(4).
15. Prinz, W., Rose, T., Osterland, T., & Putschli, C. (2019). Blockchain – Reliable Transactions. In R. Neugebauer, *Digital Transformation* (S. 301-309). Springer Vieweg.

16. Rose, T., Gruber, J., & Hausmann, K. (September 2020). PigFarm – Developing decision support for the pork production industry. *Proc. EnviroInfo 2020, Digital Twins for Sustainability*. Nikosia, Zypern.
17. Rose, T., Gruber, J., Hausmann, K., & Osterland, T. (2021). PigConomy – Evidenzbasierte Analyse von empirischen Daten der Nutztierhaltung und deren Verwertung. In D. T. al., *Monetarisierung von Technischen Daten*. Springer.
18. Schumacher, U., & et. al. (2011). *Wissensstandsanalyse zur Tiergesundheit aller Nutztierarten im Ökologischen Landbau und 100% Biofütterung*.
19. Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica*, 26, p. 24-36.