

## Recenzja pracy doktorskiej mgr inż. Dawida Słomiana

pt. „Ocena wartości hodowlanej bydła przy zastosowaniu modelu jednostopniowego”

Promotor: prof. dr hab. Joanna Szyda  
oraz promotor pomocniczy dr inż. Kacper Żukowski

Instytut Zootechniki – Państwowy Instytut Badawczy w Balicach

### Ocena formalna

Przedstawiona do recenzji dysertacja ma formę rozprawy opartej na spójnym cyklu czterech prac naukowych poświęconych metodologicznym i aplikacyjnym aspektom wykorzystania modeli jednostopniowych w ocenie wartości hodowlanej bydła mlecznego. Rozprawa została przygotowana w dyscyplinie zootechnika i rybactwo, zawiera poprawnie wyodrębnione części obejmujące wstęp, cel badań, hipotezy, materiał i metody, wyniki i dyskusję, konkluzje oraz bibliografię. Na cykl składają się dwie opublikowane prace oryginalne oraz dwa preprinty:

- P1. Słomian, D.,** Żukowski, K., Szyda, J. (2023). Heterogeneity in convergence behaviour of the single-step SNP-BLUP model across different effects and animal groups. *Genetics Selection Evolution*, 55(1). <https://doi.org/10.1186/s12711-023-00856-5>; **P2. Słomian, D.,** Żukowski, K., Szyda, J. (2025). A comparison of genomically enhanced breeding values predicted by different single-step approaches. *Annals of Animal Science*. <https://doi.org/10.2478/aoas-2025-0088>; **P3. Słomian, D.,** Jakimowicz, M., Suchocki, T., Szyda, J. (2025). Comparison of BLUPF90IOD3 and MiXBLUP implementations of the single-step model applied to the Polish national dairy cattle evaluation. PREPRINT (Version 1) available at Research Square [<https://doi.org/10.21203/rs.3.rs-8398690/v1>]; **P4. Słomian, D.,** Vandenplas, J., Ten Napel, J., Żukowski, K., Skarwecka, M., Szyda, J. (2025). Modeling missing parents in single-step test-day SNP-BLUP evaluation of dairy cattle. [Preprint]. bioRxiv. <https://doi.org/10.64898/2025.12.02.691779>

Wkład doktoranta został jednoznacznie określony jako wiodący w zakresie przygotowania danych, wykonania analiz oraz opracowania w/w manuskryptów. Układ pracy jest logiczny, prezentuje spójny, konsekwentny opis przebiegu badań założonych w tezie rozprawy

doktorskiej, dokumentujący właściwe przyporządkowanie zakresu tematycznego poszczególnych publikacji głównemu problemowi badawczemu, czyli ocenie użyteczności różnych wariantów modeli jednostopniowych w rutynowej krajowej ocenie bydła.

## **Ocena merytoryczna**

### ***1) Tematyka badań i ich znaczenie***

Tematyka rozprawy dotyczy jednego z kluczowych i wciąż aktualnych zagadnień nowoczesnej hodowli zwierząt, jakim jest ocena wartości hodowlanej. Integracja informacji fenotypowej, rodowodowej i genomowej w ramach jednolitego modelu predykcji stwarza unikalną możliwość zwiększenia precyzji szacowania wartości hodowlanej bydła. Podjęty w badaniach temat ma zarówno walor poznawczy, jak i praktyczny, gdyż doskonalenie metod rutynowej oceny wartości hodowlanej bezpośrednio wpływa na trafność selekcji zwierząt, a tym samym na tempo postępu hodowlanego oraz poprawę efektywności produkcji.

Rozprawa koncentruje się na analizie populacji bydła holsztyńsko-fryzyjskiego objętej polskim systemem rutynowej oceny wartości hodowlanej, z wykorzystaniem rzeczywistego materiału badawczego obejmującego dane o wartości użytkowej zwierząt oraz dane genomowe o dużej skali, co nadaje prowadzonym analizom szczególnie złożony i obliczeniowo wymagający charakter.

Autor podejmuje ważny problem metodycznych i obliczeniowych uwarunkowań rutynowej oceny wartości hodowlanej bydła holsztyńsko-fryzyjskiego z zastosowaniem modeli jednostopniowych. Analizuje wpływ zakresu informacji rodowodowych na zbieżność modelu, porównuje warianty modeli jednostopniowych i wyniki uzyskiwane w programach MiXBLUP i BLUPF90 oraz ocenia wpływ różnych sposobów kodowania brakujących danych rodowodowych na oszacowania wartości hodowlanych i wyniki ich walidacji. Tak określony problem badawczy jest bardzo dobrze osadzony w aktualnych potrzebach genomowej oceny wartości hodowlanej oraz praktycznego wdrażania tych rozwiązań w krajowej hodowli bydła.

### ***2) Oryginalność tematyki i wartość poznawcza***

Oryginalność przedstawionej do oceny dysertacji wynika przede wszystkim z kompleksowego, wielowątkowego opracowania zagadnienia modeli jednostopniowych w odniesieniu do krajowej rutynowej oceny bydła mlecznego, opartego na rzeczywistych danych fenotypowych, genotypowych i rodowodowych. Autor nie ogranicza się do porównania

pojedynczych algorytmów, lecz analizuje cały ciąg zagadnień decydujących o praktycznej przydatności tych modeli, obejmujący stabilność numeryczną, tempo zbieżności, zapotrzebowanie na zasoby obliczeniowe, zgodność oszacowań wartości hodowlanych oraz wrażliwość modeli na niekompletność rodowodów.

Na podkreślenie zasługuje także fakt, że badania zostały wykonane na bardzo dużych zbiorach rzeczywistych danych, obejmujących między innymi ponad 1,09 mln krów z fenotypami, ponad 141 tys. buhajów z pseudo-fenotypami w postaci wartości hodowlanych pochodzących z metody wielocechowej oceny międzynarodowej wartości hodowlanej (MACE), poddanych deregresji (DRP), blisko 135 tys. osobników zgenotypowanych oraz rodowody liczące od 1,56 do 8,45 mln zwierząt, a w analizach próbnych udojów ponad 63,6 mln rekordów. Takie ujęcie nadaje uzyskanym wynikom wysoki walor aplikacyjny, zwiększa ich wiarygodność i tym samym istotnie podnosi ich znaczenie dla praktyki rutynowej oceny wartości hodowlanej.

### **3) Zdefiniowanie celów i hipotez badawczych**

Cele rozprawy zostały sformułowane jasno i pozostają w pełni zgodne z przedstawionym zakresem badań. Autor wyodrębnił cztery hipotezy odnoszące się do wpływu liczby pokoleń uwzględnionych w rodowodzie na zbieżność modelu, porównywalności wyników uzyskiwanych przy zastosowaniu różnych wariantów modeli jednostopniowych, zgodności oszacowań otrzymywanych z wykorzystaniem oprogramowania MiXBLUP i BLUPF90 oraz wpływu sposobu kodowania brakujących informacji rodowodowych na oszacowania wartości hodowlanych i wyniki walidacji. Hipotezy przedstawione przez autora są logiczne, precyzyjnie sformułowane i dobrze powiązane z przyjętym planem badawczym.

Należy podkreślić, że struktura pracy, przy zachowaniu jednolitego podejścia metodologicznego, umożliwia systematyczną weryfikację każdej z postawionych hipotez w poszczególnych publikacjach składających się na przedstawioną rozprawę doktorską. Świadczy to o przemyślanym zaplanowaniu badań oraz wysokiej dojrzałości warsztatowej doktoranta.

### **4) Zakres i metody badań**

Zastosowany warsztat badawczy należy ocenić bardzo wysoko. Autor wykorzystał i porównał zarówno modele jednostopniowe SNP-BLUP, jak i G-BLUP, analizując ich właściwości w zróżnicowanych wariantach metodycznych i obliczeniowych. Uwzględnił dwa

typy cech pokrojowych o odmiennej odziedziczalności, różne warianty doboru osobników rdzeniowych w podejściu APY (Algorithm of Proven and Young), dwa niezależne systemy oprogramowania oraz trzy sposoby kodowania brakujących rodziców: surowy rodowód, grupy genetyczne i kody metafounders.

Na uznanie zasługuje również zastosowanie procedur walidacyjnych opartych na porównaniu oszacowań uzyskanych z pełnych i zredukowanych zbiorów danych, z uwzględnieniem nie tylko dokładności predykcji, ale także zgodności oszacowań wartości hodowlanych oraz wymagań obliczeniowych związanych z zastosowaniem poszczególnych modeli. Szczególnie cenne jest rozwinięcie analiz od modeli jednocechowych do bardziej zaawansowanego modelu jednostopniowego SNP-BLUP z regresją losową dla próbnych udojów, co lepiej oddaje realia rutynowej oceny wartości hodowlanej bydła mlecznego.

#### **Synteza wyników prezentowanych w dysertacji**

W pracy P1 autor wykazał, że głębokość rodowodu w istotny sposób wpływa na tempo zbieżności procesu iteracyjnego szacowania wartości hodowlanej. Warto podkreślić, że rodowód zredukowany do pięciu pokoleń zapewniał około dwukrotnie szybsze osiągnięcie zbieżności, przy zachowaniu wysokiej zgodności oszacowań wartości hodowlanych z wynikami uzyskanymi dla rodowodu pełnego, choć to pełny rodowód dawał oszacowania bardziej zbliżone do rozwiązań ostatecznych. Jednocześnie wskazano, że w przypadku osobników nieposiadających ani informacji fenotypowej, ani genotypowej rozbieżności oszacowań w trakcie procesu iteracyjnego były największe, co ma znaczenie dla praktyki obliczeń wielkoskalowych.

W pracy P2 doktorant porównał różne warianty modeli jednostopniowych. Wyniki walidacji dla cech wysokość w krzyżu i kąć racycy wskazały na dużą zgodność podejść SNP-BLUP i GT, a jednocześnie na niższą jakość oszacowań w wariantach APY wykorzystujących jedynie 3000 osobników rdzeniowych. Autor trafnie wskazał, że odpowiedni dobór liczby osobników rdzeniowych jest istotnym warunkiem jakości oszacowań oraz wyników walidacji, a także że model SNP-BLUP cechuje się wysoką efektywnością obliczeniową, wyrażającą się krótszym czasem osiągania zbieżności i mniejszym zapotrzebowaniem na pamięć operacyjną.

W pracy P3 zestawiono wyniki uzyskane w dwóch programach: MiXBLUP i BLUPF90. Uzyskane korelacje oraz wyniki walidacji, obejmujące współczynniki regresji i współczynniki

determinacji, wskazały na wysoką zgodność wyników uzyskiwanych w obu programach, co potwierdza, że oba systemy mogą być wykorzystywane w wiarygodnej rutynowej ocenie wartości hodowlanej, a obserwowane różnice mają głównie charakter techniczno-implementationalny.

W pracy P4 przeanalizowano wpływ brakujących informacji rodowodowych na oszacowania wartości hodowlanych. Autor wykazał, że metody kodowania brakujących rodziców z użyciem grup genetycznych (GG, Genetic Groups) oraz kodów metafounders (MF, Metafounders) są na ogół korzystniejsze niż wykorzystanie surowego rodowodu, a zarazem trafnie wskazał ograniczenia podejścia metafounders przy wzrastającej niekompletności rodowodów. Wniosek ten ma istotne znaczenie praktyczne, ponieważ pokazuje, że dobór sposobu kodowania brakujących rodziców powinien być ściśle dostosowany do kompletności i struktury danych.

#### **Zestawienie badań i wyników z aktualnym stanem wiedzy**

Prezentowane w pracy badania i uzyskane wyniki zostały w sposób adekwatny odniesione do literatury światowej, co potwierdza ich znaczenie zarówno poznawcze, jak i aplikacyjne. Autor wykazał zbieżność własnych obserwacji z wcześniejszymi pracami dotyczącymi szacowania wartości hodowlanej z zastosowaniem modeli G-BLUP i SNP-BLUP, wskazał na istotne znaczenie liczby osobników rdzeniowych w podejściu APY (Algorithm of Proven and Young), a także wpływ różnych sposobów kodowania brakujących danych rodowodowych na uzyskiwane oszacowania. Należy podkreślić, że rozprawa wnosi do tego obszaru wiedzy istotny nowy element w postaci usystematyzowanego porównania metod szacowania wartości hodowlanej, przeprowadzonego w warunkach krajowej rutynowej oceny wartości hodowlanej, opartego na danych rzeczywistych.

W mojej ocenie szczególnie wartościowe jest połączenie wyników stricte metodologicznych z perspektywą wdrożeniową. Autor nie poprzestaje na porównaniu różnic między badanymi podejściami, lecz interpretuje je w kontekście czasochłonności i wymagań obliczeniowych, zgodności oszacowań wartości hodowlanych oraz przydatności praktycznej analizowanych rozwiązań. Nadaje to rozprawie wyraźny wymiar praktyczny, wykraczający poza typowe porównanie modeli, algorytmów i sposobów ich zastosowania.

### **Mocne strony i ograniczenia przedstawionej rozprawy**

Do najważniejszych mocnych stron pracy zaliczam wykorzystanie bardzo dużych i rzeczywistych zbiorów danych, logicznie zaplanowany cykl publikacyjny, wysoki poziom metodologiczny analiz, jednocześnie uwzględnienie trafności predykcji i wydajności obliczeniowej oraz wyraźny walor aplikacyjny uzyskanych wyników. Rozprawa dobrze pokazuje, że efektywność modelu w ocenie wartości hodowlanej nie może być rozpatrywana wyłącznie przez pryzmat jakości predykcji, lecz powinna być oceniana łącznie z wymaganiami obliczeniowymi oraz wrażliwością na strukturę i kompletność danych.

Za szczególnie cenne uważam również to, że autor nie ograniczył się do eksponowania wyników potwierdzających przydatność analizowanych rozwiązań, lecz wskazał także obszary, w których określone podejścia wykazują ograniczenia lub mogą prowadzić do niedoszacowania wartości hodowlanych. Taka rzetelność interpretacyjna zwiększa wiarygodność całej rozprawy.

Pewne ograniczenia pracy mają charakter formalny, ale w istocie pozostają neutralne dla ogólnej oceny jej wartości naukowej. Dotyczy to przede wszystkim faktu, że dwie z czterech prac wchodzących w skład cyklu mają formę preprintów, a więc nie przeszły jeszcze pełnej, niezależnej procedury recenzyjnej. Nie podważa to wartości naukowej wyników przedstawionych w preprintach, choć pełniejsze potwierdzenie ich znaczenia przyniesie przyjęcie tych prac do publikacji oraz ich opublikowanie w renomowanych recenzowanych czasopiśmie.

Drugim elementem, który mógłby zostać szerzej rozwinięty, byłaby pogłębiona dyskusja praktycznych kryteriów doboru modelu w zależności od struktury populacji, wymaganej infrastruktury obliczeniowej oraz wskazanej częstotliwości rutynowych wycen. Rozprawa dostarcza w tym zakresie bardzo mocnych podstaw, pozostawiając zarazem przestrzeń na dalsze opracowania wdrożeniowe.

**W związku z powyższym proszę Autora o ustosunkowanie się do następujących kwestii:**

1. Które z analizowanych podejść Autor uznałby za najbardziej perspektywiczne do długoterminowego wdrożenia w krajowej rutynowej ocenie wartości hodowlanej, skoro wyniki pracy wskazują na zbliżoną trafność predykcji różnych modeli, a zarazem na wyraźne korzyści obliczeniowe modelu SNP-BLUP??

2. Na ile, zdaniem Autora, obserwowane w pracy różnice między podejściami APY, GT i SNP-BLUP powinny być uznawane za istotne z punktu widzenia praktyki selekcyjnej, zwłaszcza gdy decyzje hodowlane dotyczą bardzo wąskiej grupy najlepszych osobników?
3. Czy Autor rozważał lub planuje rozszerzenie analiz o scenariusze, w których wartości wariancji genetycznej i resztowej byłyby estymowane dla poszczególnych wariantów analiz, zamiast być przyjmowane jako stałe na podstawie rutynowych wycen?
4. Jak Autor ocenia możliwość opracowania praktycznych rekomendacji dotyczących wyboru metody kodowania brakujących rodziców w zależności od poziomu kompletności rodowodów oraz udziału osobników zgenotypowanych w populacji?
5. Czy zdaniem Autora wyniki uzyskane dla populacji polskiej mogą być bezpośrednio przenoszone do innych populacji krajowych, czy też wymagają uprzedniego dostosowania do specyfiki struktury danych, głębokości rodowodów oraz zakresu genotypowania?

Powyższe pytania i zawarte w nich uwagi nie obniżają mojej wysokiej oceny rozprawy. Przeciwnie, traktuję je jako przyczynek do szerszej dyskusji nad pracą, która wyróżnia się dojrzałością metodologiczną, bardzo dobrym osadzeniem w realiach krajowej oceny wartości hodowlanej z wykorzystaniem informacji genomowej, a także dużą użytecznością praktyczną uzyskanych wyników dla organizacji i doskonalenia systemu oceny wartości hodowlanej bydła mlecznego.

#### **Na szczególne podkreślenie zasługują następujące aspekty:**

- 1. Kompleksowe ujęcie problemu** – autor połączył w jednej rozprawie analizę zbieżności, trafności predykcji, zgodności wyników uzyskiwanych w różnych implementacjach oraz wrażliwości modeli na braki rodowodowe, tworząc spójny i merytorycznie dobrze zintegrowany obraz badanego problemu.
- 2. Wysoka wartość aplikacyjna** – badania zostały oparte na rzeczywistych danych krajowej oceny rutynowej, dzięki czemu sformułowane wnioski mają bezpośrednie znaczenie dla praktyki hodowlanej oraz organizacji systemu oceny wartości hodowlanej.
- 3. Silny komponent metodologiczny** – rozprawa pokazuje bardzo dobre opanowanie zaawansowanych metod statystycznych, numerycznych i obliczeniowych wykorzystywanych we współczesnej genomice zwierząt gospodarskich.

**4. Rzetelność interpretacji** – autor nie tylko wskazuje zalety poszczególnych rozwiązań, lecz również precyzyjnie identyfikuje ich ograniczenia, co świadczy o dojrzałości naukowej i krytycznym podejściu do własnych wyników.

**5. Znaczenie dla dalszego rozwoju dyscypliny** – uzyskane wyniki mogą stanowić istotny punkt odniesienia dla przyszłych badań i wdrożeń dotyczących oceny wartości hodowlanej w dużych populacjach zwierząt.

### **Konkluzja**

Rozprawa mgr. inż. Dawida Słomiana prezentuje wysoki poziom naukowy, wyróżnia się oryginalnym ujęciem problemu oraz wnosi istotny wkład do rozwoju metod genomowej oceny wartości hodowlanej bydła. Autor wykazał się bardzo dobrym przygotowaniem teoretycznym, biegłością w stosowaniu nowoczesnych narzędzi analitycznych oraz umiejętnością krytycznej interpretacji wyników uzyskanych na wielkoskalowych danych rzeczywistych.

### **Podsumowanie i wniosek końcowy**

Stwierdzam, że rozprawa doktorska mgr. inż. Dawida Słomiana pt. „Ocena wartości hodowlanej bydła przy zastosowaniu modelu jednostopniowego” spełnia wymagania określone w art. 187 ust. 1–4 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce. Praca jest poprawna metodologicznie, ma wyraźny walor poznawczy i aplikacyjny oraz wnosi istotny wkład do rozwoju zootechniki i metod oceny wartości hodowlanej zwierząt. Wnoszę o dopuszczenie mgr. inż. Dawida Słomiana do dalszych etapów postępowania doktorskiego.

Jednocześnie, biorąc pod uwagę wysoki poziom merytoryczny rozprawy, nowatorskie i kompleksowe ujęcie problemu oraz znaczący potencjał aplikacyjny uzyskanych wyników, wnoszę o wyróżnienie niniejszej rozprawy doktorskiej.

Z wyrazami szacunku,  


Prof. dr hab. Mariusz Pierzchała