

Modemo - Construction et mise à disposition d'un modèle de prévision démographique pour l'élevage bovin français, et test d'application à l'espèce porcine

Groshens E.¹, Lecuyer B.², Maigne E.³, Pauly O.³

¹ Idele - Institut de l'Elevage – 149 rue de Bercy, F-75012 Paris

² IFIP - Institut du porc, la Motte au Vicomte, BP 35104, F-35651 Le Rheu Cedex

³ INRAE – 24, Chemin de Borde Rouge - Auzeville - CS 52627, F-31326 Castanet Tolosan Cedex

Correspondance : eva.groshens@idele.fr

Résumé

Afin de répondre aux demandes croissantes des acteurs des filières d'élevage en direct ou via les interprofessions, l'Institut de l'Elevage (Idele) et l'Institut du porc (IFIP) ont souhaité mener un travail méthodologique commun sur la prévision des cheptels et de leur production, à partir des bases de données animales bovines et porcines. Si la base porcine, plus récente ne demandait au préalable que des travaux de qualification de la donnée, la profondeur d'utilisation des bases bovines a permis la construction d'un outil d'exploration du modèle dans le cadre du projet. Ce dernier, pour lequel l'INRAE-ODR a apporté son expertise en matière de construction d'applications, a abouti fin 2018, et notamment sur l'outil de prévision Modemo. Cet outil s'est révélé particulièrement utile pour éclairer l'impact de la crise Covid-19 sur le marché des jeunes bovins perturbé par le premier confinement. De cette expérience résulte un indicateur pérenne de suivi de la dynamique d'écoulement des jeunes bovins, publiés chaque quinzaine pour l'interprofession bovine Interbev. L'IFIP est également sollicité pour déployer ses prévisions de production.

Mots-clés : modélisation démographique, prévision, SPIE-BDNI, BDPORC, Normabev, bœuf, porc

Abstract: Modemo: modelling and forecasting the demographics of French cattle and pig herds from animal identification data

To be able to meet the growing demand from the meat sector for production forecasts, the Livestock Institute (Idele) has committed alongside the Pork Institute (IFIP) and the Rural Development Observatory (INRAE-ODR) in a demographic forecasting project, relying on national identification data analysis. The pig database was quite recent, and needed further analysis before using it for forecasting. As we already had good knowledge of the cattle database and its exhaustivity, we were able to develop a monthly demographic forecast model for the bovine species, broken down by racial groups within different perimeters (France, administrative regions). The Modemo'Bovins tool is constantly evolving. Nevertheless, it has already made it possible to respond reactively to the demands of the sectors, and in particular in the context of the current Covid-19 crisis. From this experience emerged a new indicator of the fluidity of male cattle, which is still in use, and produced fortnightly.

Keywords: demographic model, forecast, SPIE-BDNI, BDPORC, Normabev, cattle, pork

Introduction

Dans un contexte de libéralisation accrue des marchés des produits agricoles, les filières d'élevage expriment des besoins croissants de prévisions de productions. Les instituts techniques Idele (Institut de l'Élevage) et IFIP (Institut du porc) ont développé à la demande des interprofessions des modèles de prévision de productions, qui s'appuient sur des processus de type ARMA (AutoRegressive Moving

Average), qui combinent auto-régression et moyennes mobiles pour prévoir une série en fonction de ses valeurs passées (Charpentier, 2011). Ces modèles ne prenaient en compte que très peu des évolutions des cheptels, qui pourtant peuvent être décrites finement grâce aux bases d'identification animales. Idele et l'IFIP ont donc souhaité mener un travail méthodologique commun de prévision démographique en s'appuyant sur ces bases, avec pour objectifs (i) d'améliorer les prévisions de productions, (ii) de mieux connaître les dynamiques des cheptels, et (iii) de susciter de nouvelles valorisations des bases de données d'identification animale. L'INRA – ODR s'est associé au projet pour développer l'outil de mise à disposition des données du modèle bovin. Le projet Modemo a bénéficié du soutien du Casdar (compte d'affectation spéciale « Développement agricole et rural »), dans le cadre de l'appel à projets « Recherches & technologies ».

Au-delà des équipes techniques impliquées et des experts sollicités auprès du Cirad, de l'INRA et du SSP (Service de la Statistique et de la Prospective), les utilisateurs finaux ont été associés aux travaux menés au sein du comité de pilotage : interprofessions porcine (Inaporc) et bovine (Interbev), fédération nationale bovine (FNB) et porcine (FNP), Coop de France. Ainsi, ces utilisateurs ont pu suivre le projet dès la phase exploratoire, et orienter les développements pour assurer l'adéquation de l'outil à leurs besoins. Les gestionnaires des données SPIE, Normabev et BDPORC ont également participé aux comités. En plus des trois comités de pilotage, une réunion de travail avec les premiers utilisateurs a permis de valider le premier prototype présenté en cours de projet, et de décider des évolutions à y apporter.

Malgré une réflexion méthodologique commune sur les espèces bovine et porcine, deux voies de modélisation différentes ont ensuite été suivies pour aborder au mieux la prévision de production pour chacune des espèces. En effet, si les bases SPIE-BDNI et BDPORC recensent toutes deux des données exhaustives sur les mouvements d'animaux et les sites d'élevages, elles ont néanmoins des structurations bien différentes, et ne fournissent pas le même type de données.

Les bovins sont identifiés de façon individuelle dès leur naissance jusqu'à leur mort, dans les données SPIE-BDNI et Normabev. Ces informations permettent de comptabiliser les animaux présents sur un périmètre donné et à une date donnée, et leurs mouvements entre deux dates (naissances, importations, exportations, abattage, mort naturelle). Ces informations sont valorisées depuis de nombreuses années par le département Economie de l'Institut de l'Élevage, pour le suivi des cheptels et des productions dans le cadre du suivi de conjoncture mensuel : publication Tendances Lait et Viandes (Idele, 2022), Bulletin économique pour Interbev, fourniture d'indicateurs pour l'interprofession laitière CNIEL, études ponctuelles. La qualité avérée de ces données et leur connaissance préalable nous ont ainsi permis de nous engager dans la construction d'un modèle démographique abouti dans les temps du projet.

BDPORC fournit des informations sur les flux de porcs entre élevages, et des élevages vers les abattoirs et à l'export. En revanche, elle ne permet pas de connaître directement la situation du cheptel. Les animaux sont identifiés par lot lors de mouvements et non pas individuellement, et les naissances ne sont pas déclarées. La construction d'un modèle démographique similaire n'était donc pas envisageable, et la voie d'une prévision par modélisation de séries temporelles a été privilégiée. La base BDPORC est par ailleurs plus récente que la BDNI Bovine, et son usage par l'IFIP en est également plus récent. Le projet Modemo a notamment permis une forte montée en expertise des équipes de l'IFIP sur cette base de données.

Dans cet article, une présentation du modèle démographique bovin sera faite (partie 1) avant de s'intéresser aux méthodes de prévisions (partie 2) et en particulier au modèle porcin puis à l'approche de prévision simplifiée adoptée pour le modèle bovin. Dans la 3^{ème} et dernière partie, l'outil Modemo'Bovins sera présenté avec les premières utilisations concrètes et fructueuses de cet outil pour répondre aux interrogations des filières.

1. Un Modèle démographique en temps discret pour l'espèce bovine

1.1 Un modèle démographique, décliné par race et par périmètre

Pour modéliser les dynamiques démographiques de l'espèce bovine, un modèle en temps discret (pas de temps de 1 mois), qui s'appuie sur une stratification de la population par sexe, âge et stade reproductif, a été choisi. Ce choix a pour avantage de correspondre aux catégories de population et à la temporalité couramment utilisées par la filière. Ce modèle est décliné pour 11 groupes raciaux (Charolaise, Limousine, Blonde d'Aquitaine, autres races allaitantes, Prim'Holstein, Montbéliarde, Normande, autres races laitières, croisés viande x viande, lait x lait, lait x viande) (Figure 1).

Chacun des modèles en race pure est alimenté par les seules naissances issues de mères de la même race. Les cas de naissance de race pure portée par une mère d'une autre race pure (mère porteuse suite à transfert embryonnaire), ou les naissances de race pure issue d'une mère de type croisé (croisement d'absorption) ne sont pas directement modélisés, mais gérés à travers l'application d'un taux de correction global. Les trois modèles croisés prennent quant à eux en compte les naissances croisées issues des modèles en race pure.

Le modèle a été développé uniquement au niveau national dans les temps du projet, mais a d'emblée été pensé pour pouvoir être décliné sur tout périmètre constitué d'un ensemble de numéros de cheptel (régions administratives, adhérents d'une coopérative, cheptels biologiques...). Il a de fait été décliné à l'échelle régionale pour les besoins du suivi de l'impact de la crise Covid-19 sur le marché des jeunes bovins.

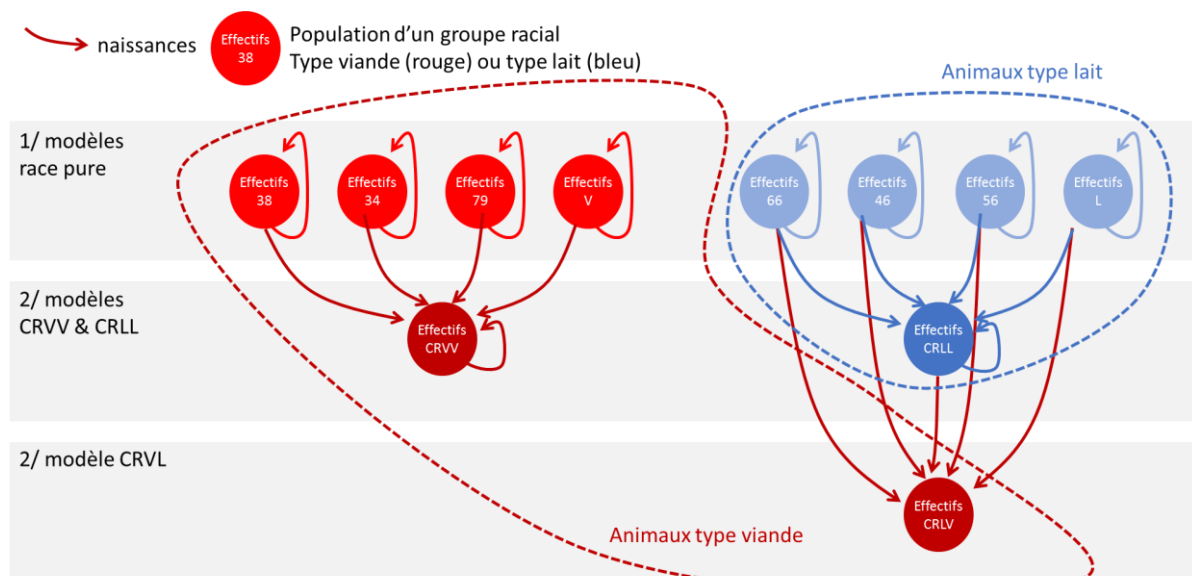


Figure 1 : Liens entre les modèles par race (38 : Charolaise, 34 : Limousine, 79 : Blonde d'Aquitaine, V : autres races pures de type viande, 66 : Prim'Holstein, 46 : Montbéliarde, 56 : Normande, L : Autres races pures laitières, CRVV : Croisés viande x viande, CRLV : Croisés lait x viande, CRLL : Croisés lait x lait).

1.2 Segmentation de la population bovine grâce aux données d'identification

Pour chaque race et chaque périmètre, la population bovine complète est segmentée à chaque début de période (1^{er} jour du mois) depuis 2010 en mâles, génisses et vaches selon le sexe et l'existence ou non d'un premier vêlage connu à la date (Figure 2). Les mâles et les génisses sont segmentées selon leur âge, par tranche d'âge de 1 mois environ jusqu'à 36 mois, la dernière tranche regroupant l'ensemble des animaux de 37 mois et plus. Pour les vaches, une forme de stade reproductif a été approché grâce au délai écoulé depuis le dernier vêlage, essentiel pour la prévision des naissances.

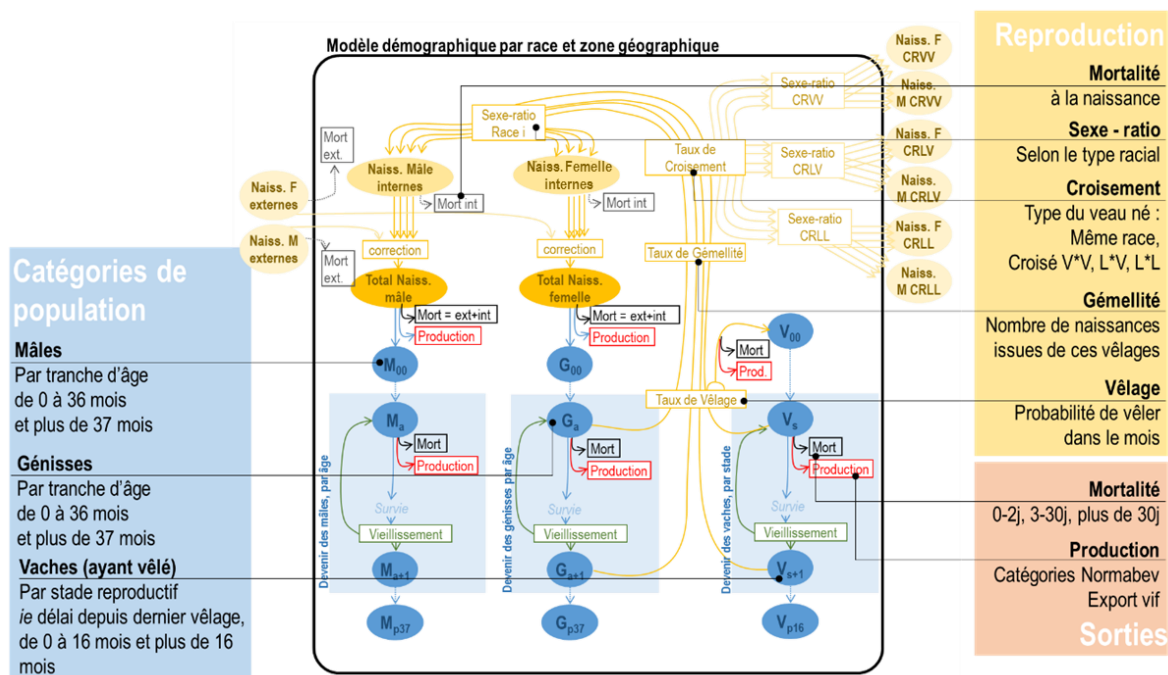


Figure 2 : Schéma général du modèle démographique par race et par péricimètre. La population est segmentée en catégories, et chacun des flux est représenté par une flèche. Les importations ne sont pas représentées sur ce schéma général.

1.3 La reproduction, mécanisme central du modèle

1.3.1 Femelles et vêlage

Les fonctions de reproduction (Figure 2) ne sont appliquées qu'aux catégories de femelles susceptibles de vêler : génisses d'au moins 11 mois et vaches ayant vêlé depuis 5 mois ou plus. Un taux de vêlage est calculé pour chacune de ces catégories : il correspond à la part des femelles ayant vêlé dans le mois rapporté aux effectifs totaux de la catégorie en début de période.

1.3.2 Gémellité, croisement et sexe-ratio, entreconnectés

Un taux de gémellité est calculé pour prendre en compte les naissances multiples : il correspond au nombre de naissances issues de chacune des catégories de femelles rapporté au nombre de vêlages de la même catégorie. Sont alors identifiés le sexe du veau et le type de croisement (même race que la mère, croisé lait x lait, lait x viande ou viande x viande) afin de pouvoir calculer respectivement le sexe-ratio et les taux de croisement de chacun des types décrits précédemment. Il a été décidé de lier ces caractéristiques du veau né sous la mère, suite à de précédents travaux ayant montré que les éleveurs avaient des pratiques différentes de croisement et de sexage en fonction des caractéristiques des mères et du rang de la tentative d'insémination (Idele, 2018). La mortalité du veau à la naissance a également été étudiée en lien avec l'âge de la primipare, ou avec le délai depuis le dernier vêlage de la multipare. Les veaux survivants et de même race que la mère sont regroupés comme « naissances internes », et réintégrés en entrée dans le modèle : ils viennent alimenter les catégories de mâles et de génisses les plus jeunes (M00 et G00, (Figure 2)).

1.3.3 Gestion des naissances externes et correction

Dans le cas des trois modèles croisés CRVV, CRLV et CRL, les naissances croisées issues des mères de race pure, nécessairement calculées au préalable, et appelées « naissances externes », ont été réintégrées. Une fois ces naissances intégrées dans « naissances totales », certaines naissances échappent toujours au modèle, notamment pour le cas des naissances de race pure issues de mères

croisées évoqué précédemment. Un taux de correction global permet alors de redresser les naissances mâles et les naissances femelles de la race.

1.4 Gestion des importations

Une importation correspond à l'arrivée d'un animal au cours d'un mois donné, alors qu'il était hors périmètre en début de mois. Pour le périmètre « France métropolitaine », il s'agit des importations en provenance de l'étranger. Pour une région administrative, les importations comprennent par exemple l'ensemble des broutards importés d'autres régions pour engraissement. Un animal importé et sorti du périmètre dans le même mois n'est pas pris en compte.

1.5 La mortalité

La mortalité est dénombrée sur l'ensemble des animaux présents en début de période, ou nés en cours de période (Figure 2, encadrés noirs). Les morts sont caractérisés selon l'âge de l'animal présent en début de mois, et selon qu'il ait vêlé ou non en cours de mois, pour tenir compte de la surmortalité périnatale de la mère. Enfin, pour les veaux nés dans le mois et comme déjà évoqué, la mortalité est étudiée selon la catégorie de la mère, ce qui permet par exemple de tenir compte de la surmortalité liée aux naissances prématurées ou issues de jeunes génisses. Les morts sont ensuite regroupés selon leur âge : 0 à 2 jours, 3 à 30 jours et à plus de 30 jours.

1.6 Les productions : enjeu principal en matière de prévision

Pour chacune des catégories de bovins présents, les productions qui en sont issues sont quantifiées (Figure 2). Les animaux nés dans le mois peuvent également donner lieu à une production de veaux dits « de 15 jours ». Cette sortie du modèle intéresse plus particulièrement les opérateurs des filières associés au comité de pilotage. Les productions sont regroupées selon les grandes catégories habituellement utilisées au sein des filières :

- Veau de 15 jours : exporté hors périmètre à moins de 1 mois
- Veau de boucherie : abattu et déclaré « V » dans Normabev
- Bovin jeune, déclaré « Z » dans Normabev
- Broutard exporté hors périmètre de 4 à 15 mois inclus
- Jeune bovin mâle, abattu et déclaré « A » dans Normabev ou exporté entre 16 et 24 mois
- Bœuf castré, abattu et déclaré « C » dans Normabev
- Taureau non castré, abattu et déclaré « B » dans Normabev
- Génisse, abattue à de plus de 12 mois et n'ayant pas vêlé
- Vache de réforme, ayant vêlé avant abattage
- Production résiduelle, regroupant les animaux restants.

Les productions sont suivies selon la catégorie de population d'origine : animaux présents en début de période ou nés et produits dans la même période.

1.7 Description de l'historique à l'aide des données d'identification animale

L'analyse des données d'identification SPIE-BDNI et Normabev a permis de décrire l'historique de l'état de la population ainsi segmentée, et des flux entrants et sortants associés à chacune de ces catégories. Pour cela, un appui a été nécessaire sur l'expertise préexistante en la matière pour la production d'indicateurs ; la plupart des flux présentés précédemment ayant déjà fait au préalable l'objet d'extractions récurrentes ou d'investigations ponctuelles au fil des demandes des filières. Chaque historique des flux

entrants ou sortants et de la population constitue ainsi une série chronologique, du 1^{er} janvier 2010 à la dernière date connue, comme par exemple abattages de génisses de race charolaise âgées de 28 à 29 mois, en région Bourgogne-Franche-Comté.

2. La prévision des séries, enjeu commun aux deux filières

2.1 Travaux de prévision menés pour le modèle porcin

2.1.1 Le modèle théorique

L'IFIP a privilégié de s'appuyer sur le cycle du porc pour construire son modèle théorique et prédire la production porcine.

A l'échelle d'un élevage, la production de porcs charcutiers est composée de 4 périodes. A l'issue de l'insémination des truies reproductrices, la gestation dure 3 mois, 3 semaines et 3 jours jusqu'à la mise-bas. Les porcelets sont sevrés lorsqu'ils atteignent un poids de 8 kg environ après une maternité de 1 mois. Après la période de post-sevrage de 6 à 8 semaines, les porcelets pèsent environ 25 kg. A l'issue des 3 à 4 mois d'engraissement, les porcs charcutiers sont envoyés à l'abattoir à un poids vif équivalant à 110 kg environ (moyenne France). La durée totale du cycle de production des porcs charcutiers, de l'insémination jusqu'à l'abattage, est donc comprise entre 9 et 11 mois.

Les truies reproductrices proviennent d'un atelier de multiplication présent dans l'élevage ou d'élevages multiplicateurs spécialisés. Elles sont réformées et envoyées à l'abattoir après environ deux ans d'activité.

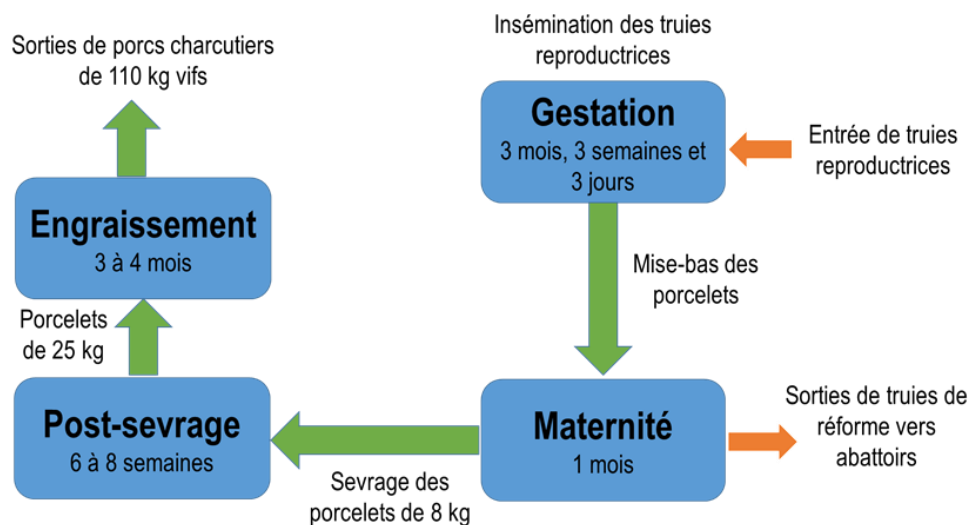


Figure 3 : Le cycle de production du porc. Source IFIP

En 2018, la France comptait 14 372 sites d'élevage ayant notifié le chargement de 22,78 millions de porcs charcutiers et de 10,86 millions de porcelets (Roguet, 2019). L'analyse des flux entrants et sortants de chaque site permet de connaître l'orientation des élevages.

Les naisseurs (avec ou sans post sevrage et engraissement) achètent des cochettes aux vendeurs de reproducteurs et/ou pratiquent l'auto-renouvellement. Ils produisent des porcelets de 8 et/ou 25 kg pour les sites de post-sevrage et d'engraissement. Les naisseurs engraisseurs engraissent la totalité ou une partie de leur production de porcelets sur place et vendent des porcs charcutiers aux abattoirs.

Les engraisseurs (avec ou sans post sevrage) se fournissent en porcelets auprès des élevages naisseurs et produisent des porcs charcutiers destinés aux abattoirs.

A l'échelle micro-économique d'un élevage, la production de porcs charcutiers et de porcelets dépend des flux passés entrants et sortants des différents types de porcins. Le nombre de porcs charcutiers produits par un élevage engraisseur à un temps t dépend par exemple du nombre de porcelets entrés dans l'élevage quelques mois auparavant. La production de porcs charcutiers française pourrait donc être estimée par les relations entre flux entrants et sortants à l'échelle de chaque élevage. Un modèle d'économétrie des panels pourrait répondre à ce besoin. Le faible nombre d'observations par site (60 mois d'observations) comparé au nombre important d'individus (plus de 14 000 sites) et les anomalies de notification de certains élevages (Roguet et al., 2018) rendent cependant l'exercice difficile.

Une approche macro-économique a donc été privilégiée pour expliquer et prédire la production de porcs charcutiers de la France en fonction de la somme des flux passés d'animaux renseignés dans la base de données BDPORC (échanges de porcelets et de reproducteurs entre sites et sorties de reproducteurs de réforme). Le modèle proposé met en relation la production de porcs charcutiers à un temps t avec les flux entrants et sortants antérieurs de porcs charcutiers et des autres catégories animales, en adéquation avec les durées des stades physiologiques d'élevage. Un décalage (Lag) de 5 à 11 mois a par exemple été choisi pour les flux de porcelets (8 kg et 25 kg) : la période de post-sevrage dure entre 6 et 10 semaines et la période d'engraissement de 3 à 4 mois. Le modèle est complété avec des indicateurs conjoncturels (prix de l'aliment et marge naisseur engraisseur) et de saisonnalité.

$$\begin{aligned} \text{Prod_CH}_t = & \\ & \text{Prod_CH}_{t-5 \text{ à } t-7} + \text{C_porcelet}_{t-5 \text{ à } t-11} + \text{C_reproducteurs}_{t-10 \text{ à } t-15} + \text{C_réformes}_{t-10 \text{ à } t-15} + \\ & \text{prix_porc}_{t-5 \text{ à } t-7} + \text{taux_var_prix_alim}_{t-5 \text{ à } t-7} + \text{marge_NE}_{t-5 \text{ à } t-7} + \\ & t_1 + t_2 + t_3 \end{aligned}$$

Figure 4 : Modèle théorique porcin. Source IFIP.

2.1.2 Le modèle empirique

Le pas de temps choisi pour le modèle est mensuel. La production de porcs charcutiers et les flux de porcelets, de reproducteurs, et de réforme ont été calculés à partir de la base de données BDPORC. Les variables de conjoncture sont des données IFIP (prix d'aliment IFIP, indicateur de marge des élevages naisseur-engraisseur) ou de FranceAgriMer (prix du porc classes SE).

Un modèle de régression linéaire a été utilisé pour modéliser et prédire la production de porcs charcutiers. Pour les données BDPORC, un historique de seulement 60 mois était disponible, et des retards de 15 mois sont utilisés dans le modèle pour certaines variables. Le nombre de variables à introduire dans le modèle ne peut donc dépasser 45.

Une procédure de type pas à pas à partir du Critère d'Information d'Akaike (AIC) a été utilisée pour réduire le nombre de variables significatives (fonction R *stepAIC()*).

Les résultats de ce modèle sont prometteurs et permettent de mieux appréhender les fluctuations intra-annuelles de la production porcine que ne le font les modèles actuels.

Toutefois, il est nécessaire de consolider ce modèle avec un historique plus long, en fiabilisant les données et de tester d'autres types de méthodes de prévision.

2.2 Des méthodes de prévision simples voire simplistes pour le modèle bovin

2.2.1 Du modèle démographique à la prévision

Afin de prévoir la population à l'instant $t+k$ et connaissant son état à l'instant t , il faut proposer une prévision pour chacun des paramètres démographiques pour les périodes allant de $[t, t+1[$ à $[t+k-1, t+k[$, et pour chacune des catégories de population concernée (Figure 5).

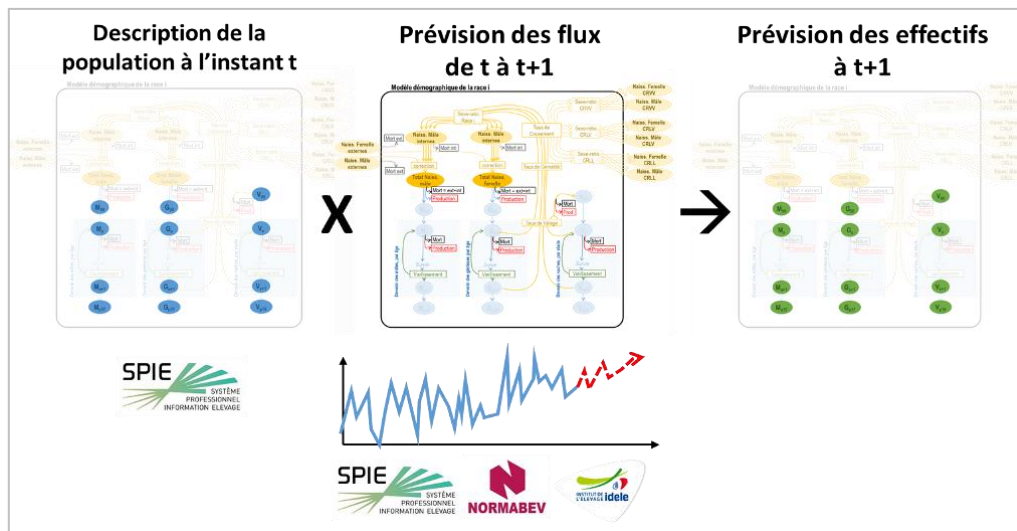


Figure 5 : Schéma de la mécanique de prévision des effectifs et des flux. Source Idele.

L'enjeu de la prévision se pose ainsi tout à fait différemment dans le cas du modèle bovin, étant donné le nombre de séries à prévoir : pour chaque périmètre et chacun des 11 groupes raciaux, une prévision pour 50 flux, dont chacun se décline pour part ou totalité des 90 catégories de bovins décrites précédemment doit être proposée. Par conséquent, il était exclu de pouvoir proposer des méthodes optimisées manuellement pour la prévision de chacune des séries dans le temps imparti du projet.

2.2.2 Des méthodes de prévision des flux simples, mais évolutives

En accord avec les utilisateurs, il a été opté pour l'utilisation d'une médiane saisonnière : la série est prédite en s'appuyant sur la médiane de ses valeurs passées 1, 2 et 3 ans auparavant. Cette méthode permet de proposer l'application d'une situation « normale », ce qui s'est avéré particulièrement utile dans le cas des premières sollicitations de l'outil. Cette médiane peut s'appliquer soit à l'effectif en nombre absolu, soit au taux, en rapportant le flux à la population présente en début de période. La nuance est particulièrement importante pour les prévisions de production. Une prévision s'appliquant au taux suppose un prélèvement normal des animaux, et donc un lien direct avec les effectifs disponibles en ferme, plutôt caractéristique d'un flux poussé par l'offre. Une prévision appliquée à l'effectif suppose davantage une certaine stabilité de la demande, et serait plus proche de l'hypothèse d'un flux tiré par la demande.

Conscients des limites de cette approche, le programme a été organisé pour permettre l'appel à de nouvelles méthodes de prévision, indépendamment pour chacune des séries, afin de pouvoir améliorer le modèle au gré des analyses futures. Une méthode de prévision est ainsi une fonction, qui prend en entrée la série à prévoir et quelques autres paramètres (taux ou effectif, profondeur, éventuelles séries externes capables d'améliorer la prévision) et fournit en sortie la prévision de la série pour les mois suivants.

Par ailleurs, différents scénarii de prévision sont possibles, une des demandes majeures des utilisateurs, pour pouvoir tester les effets d'un élément perturbateur (type épizootie, fermeture de marché à l'export) sur la dynamique des cheptels et des productions.

Un scénario est structuré dans l'outil comme un jeu complet des méthodes de prévision qui s'appliquent à chacun des flux démographiques. Le scénario tendanciel est le scénario de prévision estimé comme le meilleur rapport précision – complexité, dans l'état des connaissances du moment. Comme expliqué précédemment, il est amené à évoluer au gré des investigations. Cependant définir un scénario alternatif pertinent n'a pas été réalisable dans les temps du projet, mais l'application permet déjà leur implémentation future.

2.2.3 Évaluation de la qualité du modèle de prévision et premiers résultats

La stratégie d'amélioration continue de la qualité des prévisions de chacune des séries du modèle exige de disposer d'outil d'analyse des erreurs de prévisions. Par ailleurs, la possibilité d'explorer les erreurs passées du modèle était une demande des utilisateurs finaux.

Une analyse systématique de l'historique des prévisions par rétro-simulation a été implémentée : pour le scénario tendanciel, la prévision est lancée pour chacune des dates antérieures disposant d'une profondeur historique suffisante. Dans la version actuelle, la prévision peut être lancée à compter du 1^{er} janvier 2013. L'ensemble des séries est alors prévu (flux entrants et sortants, population résultante) pour les horizons $h=1$ à 12 mois, du 1^{er} janvier 2013 jusqu'à la dernière date disponible, en s'appuyant à chaque fois sur les seules données historiques antérieures à la date de prévision. De cette façon, il est possible d'évaluer *a posteriori* la qualité de la prévision d'un scénario donné depuis 2013, par un indicateur type MAEP (Moyenne des Écarts Absolus en Pourcentage), ou encore visuellement par analyse de courbe (rétro-simulation (Figure 7)).

2.2.4 Qualité de la prévision : l'exemple de la population charolaise

La qualité du modèle pour un périmètre donné repose sur la qualité de la prévision de chacun des flux, appliqués à chacune des catégories de population, pour chaque race.

Dans le présent article, il a été choisi de représenter les résultats du modèle concernant la prévision de la population charolaise globale, et des productions associées, pour un horizon de 4 mois. La Figure 6a présente la MAEP obtenue pour une prévision de la population de mâles à horizon 4 mois, par classe d'âge de 0 à 23 mois. Il ressort que l'erreur augmente (i) à mesure que l'on s'approche des pics de production de brouards (autour de 12 mois) et de jeunes bovins (19 mois), et (ii) pour les catégories les plus jeunes, qui s'appuient sur la prévision de naissance. Ce graphe illustre combien l'amélioration de la qualité de la prévision des effectifs repose ainsi directement sur l'amélioration de la qualité de la prévision des flux eux-mêmes.

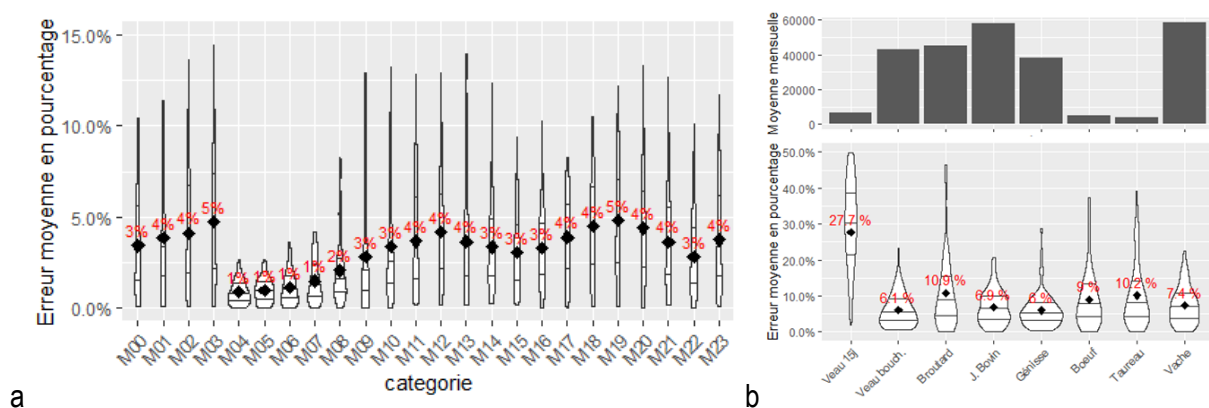


Figure 6 : a) Dispersion et moyenne de l'erreur absolue réalisée pour la prévision à 4 mois des effectifs de mâles âgés de 0 à 23 mois. Source Idele. b) Partie haute : Production mensuelle moyenne sur la période, en têtes. Partie basse : dispersion et moyenne de l'erreur absolue en pourcentage, par type de production.

Les diagrammes en violon représentent la dispersion et la moyenne de l'erreur absolue en pourcentage réalisées pour chaque type de production (Figure 6b). Il apparaît que la prévision réalisée par la simple application de médiane saisonnière offre déjà pour la prévision des sorties de jeunes bovins des performances comparables au modèle alors en production, qui reposait sur des méthodes d'analyse des séries chronologiques (procédure SAS State Space). Les résultats semblent également prometteurs pour les génisses, les vaches et les veaux de boucherie. L'erreur est plus importante pour les exportations de brouards, ou pour les productions plus rares (boeufs, taureaux, et surtout exportations de veaux de 15 jours, atypique en élevage allaitant).

3. L'outil de visualisation Modemo Bovins et premières utilisations concrètes

3.1 Un outil de visualisation dynamique R-shiny

3.1.1 Trois représentations dynamiques sélectionnées

Le développement de l'outil en ligne Modemo'Bovins répond à l'enjeu de la mise à disposition et de la visualisation des données du modèle. Nous avons dissocié le modèle de prévision démographique lui-même, gourmand en temps de calcul, de l'outil de visualisation, exigeant au contraire une grande réactivité. Nous avons garanti une connexion possible entre les deux, en programmant le premier en R, et le second en R-shiny. Le choix de R-shiny permettait par ailleurs la portabilité de l'application, et donc l'hébergement ultérieur sur le serveur Idele, ou à plus long terme sur un futur infocentre de la profession agricole. L'INRA-ODR a par ailleurs partagé ses compétences en la matière, pour faciliter la maintenance et les évolutions futures de l'application chez Idele.

Chaque mois, le modèle de prévision est lancé suite à la mise à jour des données d'identification animale SPIE-BDNI et Normabev. Les données de sortie du modèle sont alors chargées en entrée de l'application de visualisation. L'utilisateur s'identifie, et affiche les données qui l'intéressent selon son propre paramétrage : périmètre géographique, races (avec regroupements libres possibles), indicateur (population, naissances, production...), catégories de bovins, horizon de prévision (limité à 12 mois).

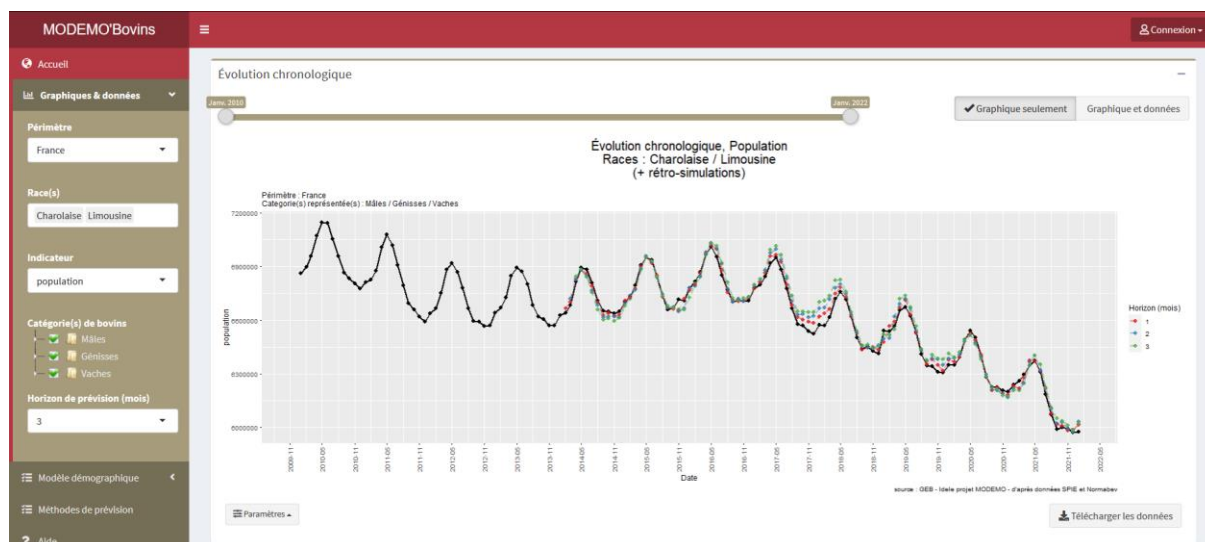


Figure 7 : Prise d'écran de l'outil Modemo'Bovins en février 2022. Évolution chronologique de la population des races charolaises et limousines en France métropolitaine, et des rétro-simulations pour les horizons 1 à 3 mois.

L'utilisateur a alors accès à trois représentations dynamiques, sélectionnées avec les utilisateurs du comité de pilotage, et chacune accompagnée d'un tableau de données.

- « **Évolution chronologique** », permet une analyse sur temps long depuis 2010. Cette représentation permet d'afficher l'historique de la série, les prévisions, l'historique des prévisions

(rétro-simulations), ainsi que des outils d'analyse des séries chronologiques : moyennes mobiles, médianes saisonnières des années précédentes. Une représentation atypique de l'historique des prévisions a été choisie : plutôt que d'afficher chaque prévision d'horizon 12 mois par date de prévision passée, l'affichage de l'historique des prévisions est réalisé par horizon (de 1 à 12 mois), afin d'avoir des graphes plus lisibles (Figure 7).

- « **Comparaison par campagne** », permet de mettre en regard l'année en cours avec quelques années passées choisies, selon la fenêtre la plus adaptée à l'indicateur (juillet à juin, par exemple).
- « **Représentation mensuelle par catégorie** », permet de comparer l'évolution du profil des animaux d'une année à l'autre. Cette représentation permet d'afficher la valeur de l'indicateur pour un mois donné par catégorie de bovins, en comparant plusieurs années. Ceci permet par exemple de comparer le profil d'âge des jeunes bovins produits en mai 2020 par rapport à 2019.

3.1.2 Gestion des profils utilisateurs et droits d'accès

Les données d'identification animale SPIE-BDNI et Normabev sont des données réglementaires, dont l'usage est soumis à conventionnement avec les organismes gestionnaires de données SPIE et Normabev, et à l'application des règles de protection des données individuelles. L'accès aux données du modèle démographique est donc soumis à conventionnement préalable. L'application prévoit une procédure d'identification, et différents profils utilisateurs pour lesquels chacun des objets de l'application peut être activé ou non (bloc graphique, périmètre, prévisions, téléchargement, etc...)

3.2 Un outil de prévision rapidement sollicité par les acteurs des filières

L'outil Modemo'Bovins a rapidement été mobilisé pour répondre à différentes interrogations des filières. En particulier, la crise de la Covid-19 a révélé usage pertinent de Modemo'Bovins. En affectant si brutalement l'ensemble de la filière alimentaire, cette crise a interrogé en particulier la filière viande : modifications de la consommation, perturbation des outils et de la logistique... Des inquiétudes grandissantes sont remontées aux organisations nationales d'Interbev concernant des difficultés d'écoulement des jeunes bovins, qui auraient constitué ainsi des stocks importants en ferme, pénalisant directement le marché et le prix payé au producteur. L'interprofession a donc demandé à Idele d'objectiver les « surstocks » présents en ferme du fait de cette crise, pour évaluer l'ampleur du problème, et pouvoir dimensionner des solutions de désengorgement du marché national.

Modemo'Bovins a été mobilisé pour proposer une estimation des sorties normalement attendues tout au long de l'année 2020, en fonction des effectifs de mâles présents en ferme en début de crise (1^{er} mars). La différence cumulée avec les sorties hebdomadaires réelles a permis de proposer une estimation de ce surstock et de son évolution tout au long de la crise. Cet indicateur a été décliné à l'échelle des régions administratives pour éclairer des inquiétudes locales. La méthode et l'estimation ont été systématiquement confrontées au ressenti terrain des opérateurs avant diffusion.

Les réflexions menées à l'occasion de cette crise avec Interbev ont conduit à proposer de façon pérenne un outil de suivi de la dynamique d'écoulement des jeunes bovins, qui est depuis 2021 actualisé chaque quinzaine. Cet outil compare l'écoulement réel des jeunes bovins avec une modélisation réalisée en début d'année en fonction des effectifs de mâles présents en ferme, auxquels on applique un taux de prélèvement « normal », grâce à l'outil Modemo'Bovins (médiane saisonnière sur 4 ans). Cet indicateur a permis de mettre en évidence tout au long de 2021 des prélèvements anormalement élevés réalisés en ferme du fait du manque d'offre : les sorties ont été largement supérieures à la modélisation jusqu'à l'automne, avec des abattages d'animaux plus jeunes que la normale. Ce rythme de sortie s'est ensuite brutalement contracté en fin d'année, faute d'offre, tandis que les prix se sont envolés (Idele, 2022). Encore une fois, l'indicateur est régulièrement confronté aux éleveurs et aux acteurs des filières pour s'assurer de sa cohérence avec les remontées du terrain.

Conclusion

Le projet Modemo a été l'occasion d'un travail méthodologique de fond qui a permis à l'IFIP et à l'Idele de réinvestir sur la question de la prévision démographique, et dans les deux cas de proposer des modèles avec des résultats prometteurs et des utilisations attendues par les filières. Les méthodes de prévisions utilisées par l'Idele et l'IFIP sont amenées à évoluer en continu, au gré des approfondissements menés par les deux instituts.

L'outil Modemo'bovins, développé par l'INRAE-ODR et l'Idele, a été d'emblée pensé pour accueillir ces évolutions de façon transparente, en y présentant l'historique et le détail des méthodes des prévisions. C'est un outil d'ores et déjà utile et utilisé par l'Idele dans sa version actuelle pour répondre avec réactivité aux demandes d'intérêt général des filières, en permettant d'explorer plus aisément les bases de données d'identification. Sa vocation est de s'ouvrir à un public plus large, ce qui pose des questions de droits d'accès aux données, qui sont actuellement en réflexion avec les gestionnaires de données.

Le modèle actuel peut être décliné sur de nouveaux périmètres (cheptels en agriculture biologique, adhérents d'une coopérative), avec des scénarios alternatifs, sans modification structurante. Deux évolutions plus profondes du modèle sont actuellement prévues à l'échelle du modèle démographique : (i) prise en compte de l'âge des vaches, qui impacte fortement leur valorisation commerciale, et (ii) segmentation des mâles selon le type d'atelier où il se trouve, ce qui permettrait de mieux anticiper leur destination (travaux en cours avec Interbev sur l'estimation des mises en place de jeunes bovins à l'engraissement). L'outil de visualisation, quant à lui, a été prévu pour pouvoir intégrer les évolutions ci-dessus, anticipées dans le cadre du projet. Il ne nécessitera donc que des évolutions marginales.

Les principales évolutions attendues concernent l'amélioration des méthodes de prévision, pour le modèle porcin comme pour le modèle bovin. En effet, si l'utilisation d'une médiane saisonnière permet de proposer une prévision pertinente dans un scénario normal, pour un indicateur stable, elle n'est pas adaptée pour un indicateur en hausse tendancielle, et accuse alors un retard systématique. Des méthodes de prévision du type de celles développées par l'IFIP, et utilisées dans le cadre d'autres travaux de prévision, permettront de lever cette difficulté. l'Idele et l'IFIP poursuivent leur collaboration en ce sens, pour continuer le travail méthodologique commun initié dans le cadre du projet.

Références bibliographiques

- Caswell H., 2001 Matrix population models - construction, analysis, and interpretation. Sinauer, Sunderland, MA., pp 722
- Charpentier A., 2012. Modèles de prévision. Séries Temporelles. UAM, ACT6420, 15 mai 2012
- Idele, Webzine Tendances Lait et Viandes – site internet <https://www.tendances-lait-viande.fr/>, 2022
- Idele, Résumé de l'étude sexage et croisement pour Interbev, 2018
- Idele, Dossier annuel bovins viande 2021, Dossier Economie de l'élevage N°526, Janvier 2022
- Roguet C., 2019. Les élevages de porcs en France en 2018 : 10 000 sites, la moitié avec truies. Baromètre Porc de l'ifip, juin 2019, synthèse p.8
- Roguet C., Lécuyer B., Busnel P., 2018. Création d'un observatoire des structures d'élevage de porcs en France. Etudes économiques. 43 p

Cet article est publié sous la licence Creative Commons (CC BY-NC-ND 3.0)



<https://creativecommons.org/licenses/by-nc-nd/3.0/fr/>

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « Innovations Agronomiques », la date de sa publication, et son URL)