



INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus

Mathématique

Marc Girondot

Nouvelle méthodologie d'analyse bayésienne du report des voix et de l'abstention lors du deuxième tour des élections

Volume 361 (2023), p. 243-256

<https://doi.org/10.5802/crmath.401>



This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



Les Comptes Rendus. Mathématique sont membres du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 1778-3569



Probabilités, Statistiques / *Probability theory, Statistics*

Nouvelle méthodologie d'analyse bayésienne du report des voix et de l'abstention lors du deuxième tour des élections

Marc Girondot*, ^a

^a Laboratoire Écologie, Systématique et Évolution, Équipe de Processus Écologiques
et Pressions Anthropiques, CNRS, AgroParisTech et Université Paris-Saclay,
UMR 8079, 91405 Orsay Cedex, France

Courriel: marc.girondot@universite-paris-saclay.fr

Résumé. Le mécanisme de report des voix lors des élections à deux tours est une grande inconnue : l'électeur suit-il une consigne de vote, lorsqu'il y en a une, ou bien développe-t-il une stratégie personnelle de vote ? L'information issue des dépouillements des premiers et seconds tours est utilisée pour modéliser le comportement des électeurs comme un problème inverse résolu par une approche bayésienne. Il est alors possible d'estimer les flux de votes entre le premier et le second tour avec une très bonne précision afin de fournir les bases objectives qui serviront à une analyse politique ou sociologique objective.

Abstract. The mechanism for deferring votes in two-round elections is a big unknown: does the voter follow a voting instruction, when there is one, or does he develop a personal voting strategy? The information from the counts of the first and second rounds is used to model voter behavior as an inverse problem solved using a Bayesian approach. It is then possible to estimate the flow of votes between the first and the second round with very good precision and to provide the objective bases which will serve for an objective political or sociological analysis.

Manuscrit reçu le 1^{er} septembre 2021, accepté le 26 juin 2022.

Introduction

Le vote à bulletin secret date de la Grèce et de la Rome antique même si ses modalités furent très différentes de ce qui est fait actuellement [16]. Son introduction eut lieu en 1795 en France avec la Constitution de l'an III qui prévoyait dans son article 31 « Toutes les élections se font au scrutin secret ». Le secret du vote fut aboli lors du passage aux constitutions napoléoniennes. Mais en 1848 la constitution de la II^e République réintroduisit le vote secret dans son article 24 « Le suffrage est direct et universel. Le scrutin est secret » et depuis cette date cette modalité du vote n'a plus été remise en cause en France. Le vote à bulletin secret est une méthode de vote dans

* Auteur correspondant.

laquelle le choix du votant à une élection ou à un référendum est anonyme, tant au moment du vote, qu'après celui-ci. L'enjeu principal est d'assurer que l'électeur exprime un choix profond en empêchant toute tentative d'influence que ce soit par intimidation ou subornation [11]. Ce qui se passe dans l'isoloir reste donc souvent un mystère et l'interprétation du résultat des élections peut être parfois très délicate. Une voie qui n'a pas été explorée à notre connaissance pour cerner les motivations des électeurs concerne l'information que l'on peut tirer d'une élection à deux tours. En effet, une partie importante du corps électoral est présent à la fois dans les deux tours ce qui peut constituer une source d'information pour comprendre les motivations des votes. Le modèle qui est présenté ici s'applique aux élections à deux tours telles que pratiquées en France, mais la généralité du modèle fait qu'il pourrait sans doute s'appliquer dans d'autres contextes électoraux étrangers.

Les élections en France sont gérées par le code électoral qui regroupe les dispositions législatives et réglementaires relatives aux élections politiques, c'est-à-dire à l'élection présidentielle (scrutin uninominal majoritaire à deux tours), l'élection des sénateurs (scrutin majoritaire à deux tours ou scrutin de liste à un seul tour selon le nombre de postes à pourvoir), des députés (scrutin majoritaire à deux tours), des conseillers régionaux (scrutin de liste à deux tours), des conseillers départementaux (scrutin majoritaire uninominal à deux tours), des représentants au parlement européen (scrutin de liste à la représentation proportionnelle) et des conseillers municipaux (scrutin majoritaire à deux tours ou scrutin de liste à deux tours selon la taille de la commune) [1]. Nous nous intéresserons ici aux élections comportant deux tours et prendrons en exemple les élections municipales qui permettent de désigner les conseils municipaux des 34 968 communes françaises (donnée au 1er mars 2019) qui à leur tour élisent en leur sein leur maire qui préside le conseil municipal et les maires-adjoints. Il est souvent coutume de désigner cette élection comme un scrutin local mais il prend un sens national quand on prend en compte que les grands électeurs du Sénat sont déterminés à 95 % par les élections municipales.

Le mode de scrutin applicable aux élections municipales dépend de la taille de la commune. La réforme introduite par la loi n° 2013-403 du 17 mai 2013 définit deux modes de scrutin très différents selon la taille de la commune. Le scrutin est majoritaire plurinominal avec panachage pour les communes de moins de 1 000 habitants. Les suffrages sont alors comptabilisés individuellement, et le panachage est autorisé : les électeurs ont le droit de ne pas respecter les listes de candidats, en votant pour des candidats issus de listes différentes. Pour les communes de plus de 2000 habitants, l'élection des conseillers municipaux se déroule selon un scrutin de liste à deux tours avec représentation proportionnelle : les candidats se présentent en listes. Lors du vote, on ne peut faire ni adjonction, ni suppression, ni modification de l'ordre de présentation des listes. L'élection peut se limiter à un seul tour en cas de majorité absolue, ou donner lieu à un second tour, auquel cas :

- Les listes qui ont obtenu au moins 10 % des suffrages exprimés peuvent s'y maintenir (mais sans possibilité de modifier l'ordre ou les membres de la liste s'il n'y a pas de fusion avec une autre liste), avec un minimum de deux ;
- Les candidats d'une liste qui a obtenu plus de 5 %, sans pouvoir se maintenir en tant que tels peuvent rejoindre une autre liste, ce qui peut amener à modifier l'ordre de présentation des candidats.

Nous nous intéresserons particulièrement à ce cas de figure : quel est le comportement des électeurs lorsque des listes fusionnent au second tour : suivent-ils la décision de leur liste de choix du premier tour ou bien développent-ils une stratégie propre soit de vote vers une autre liste soit d'abstention. Par ailleurs, on peut aussi se demander quel est le comportement des électeurs qui s'étaient abstenus ou avaient voté blanc lors du premier tour.

Pour répondre à ces questions, la méthode des sondages est la méthode généralement privi-

TABLEAU 1. Expression des probabilités de vote C quand il y a I listes. v est la probabilité qu’un électeur vote et donc $1 - v$ est la probabilité qu’il s’abstienne. u et c représentent respectivement la probabilité qu’un électeur ayant voté ait effectué un vote nul et blanc. d_A à d_I représentent la probabilité qu’un électeur ayant voté ait choisi un vote respectivement pour les listes A à I . Comme $u + c + d_A + \dots + d_I = 1$, il s’ensuit que $I + 1$ paramètres sont nécessaires pour décrire le comportement d’une personne ayant voté, le $(I + 2)^{\text{ème}}$ se déduisant facilement car la somme des probabilités doit faire 1. Par ailleurs, un paramètre doit être rajouté pour décrire le taux d’abstention $1 - v$. Il y a donc besoin de $I + 2$ paramètres pour décrire le comportement des électeurs. Ces expressions peuvent être utilisées pour décrire les comportement lors du premier tour avec des paramètres $I_1, v_1, u_1, b_1, d_{1.1}, \dots, d_{1.I_1}$ ou du second tour avec des paramètres $I_2, v_2, u_2, b_2, d_{2.1}, \dots, d_{2.I_2}$. Par ailleurs, ces expressions peuvent être aussi utilisées à l’échelle de chacun des bureaux de vote.

| | | | | | | | |
|--------------|------------|--------------|--------------|----------------|----------------|-----|----------------|
| C | Abstention | Nul | Blanc | Liste A | Liste B | ... | Liste I |
| Probabilités | $1 - v$ | v | | | | | |
| Probabilités | $1 - v$ | $v \times u$ | $v \times c$ | $v \times d_A$ | $v \times d_B$ | ... | $v \times d_I$ |

légée. Des données issues de telles analyses ont permis par exemple de déterminer l’origine sociale ou familiale du vote [25]. Mais cette solution a un coût financier non négligeable qui la rend inutilisable dans un contexte d’élection locale. Par ailleurs, cette méthode des sondages fait l’hypothèse que les réponses représentent réellement ce qu’a fait l’électeur ce qui n’est pas toujours le cas et nécessite des ajustements *a posteriori* [7] qui sont souvent mal documentés [17]. Une alternative présentée dans cette publication traite l’élection comme un problème inverse. Un problème inverse est une situation dans laquelle on tente de déterminer les causes d’un phénomène à partir des observations expérimentales de ses effets [19].

La résolution du problème inverse passe par une étape initiale de modélisation du phénomène, dite problème direct qui décrit comment les paramètres du modèle se traduisent en effets observables expérimentalement. Ensuite, à partir des mesures obtenues sur le phénomène réel, la démarche va consister à approximer au mieux les paramètres qui permettent de rendre compte de ces mesures [24]. Cette résolution peut se faire par simulation numérique ou de façon analytique. La résolution mathématique est rendue difficile par le fait les seules observations expérimentales ne suffisent pas généralement à déterminer parfaitement tous les paramètres du modèle.

1. Modélisation du comportement de vote

1.1. Formulation du problème direct

Soit N le nombre d’électeurs (par convention, on utilisera des lettres capitales pour désigner des effectifs et des lettres minuscules pour désigner des proportions ou des probabilités), N_1 et N_2 étant respectivement le nombre d’électeurs inscrits sur les listes électorales au premier et deuxième tour. Lors du premier tour, les N_1 électeurs ont pu avoir les comportements suivants : abstention, vote nul ou annulé, vote blanc ou enveloppe vide ou bien vote pour une des I_1 listes se présentant au scrutin du premier tour. On appellera C_1 les comportements possibles d’un électeur au premier tour. Le comportement des N_1 électeurs peut donc être décrit avec $I_1 + 2$ paramètres (tableau 1). Ces paramètres peuvent être estimés directement à partir des résultats du vote puisque ce sont les données issues du dépouillement.

Au second tour, I_2 listes vont être en compétition avec $I_2 \leq I_1$ suite à l’élimination de certaines listes et à la fusion potentielle de listes. Le comportement des électeurs au deuxième tour est

appelé C_2 . C_2 est dépendant de C_1 , c'est-à-dire que le comportement de vote d'un électeur au deuxième tour n'est pas indépendant de son comportement au premier tour. Le modèle décrivant le comportement des électeurs au second tour doit donc prendre en compte le comportement des électeurs au premier tour. Ainsi cela peut se décrire comme : quelle est la probabilité d'un comportement C_2 pour un électeur ayant eu un comportement C_1 au premier tour.

Le comportement des électeurs au deuxième tour en fonction de leur comportement au premier peut être décrit par une matrice $(I_2 + 3) \times (I_1 + 3)$ et il y a $I_1 + 3$ et $I_2 + 3$ informations issues du dépouillement respectivement au premier et second tour : nombre d'abstentions, de nuls, de blancs et de votes pour chacune des I_1 et I_2 listes (tableau 2). Il faut noter que toutes les informations du tableau 2 ne sont pas indépendantes puisque dans la formulation présentée la somme de chaque ligne fait 1. Dans cette formulation, on modélise le comportement C_1 d'un électeur ayant eu un comportement C_2 au second tour. Il y a donc $(I_1 + 2) \times (I_2 + 3)$ paramètres indépendants. Le nombre d'informations indépendantes sera alors $I_1 + 2$ puisque la somme est connue : c'est le nombre d'inscrits sur la liste électorale lors du premier tour. On aurait pu présenter le tableau avec la somme des colonnes faisant 1 signifiant que l'on modélise d'un comportement C_2 pour un électeur ayant eu un comportement C_1 au premier tour. Alors le nombre de paramètres indépendant aurait été de $(I_1 + 3) \times (I_2 + 2)$. Le nombre d'informations indépendantes sera alors $I_2 + 2$ puisque la somme est connue : c'est le nombre d'inscrits sur la liste électorale lors du second tour. Comme $I_2 \leq I_1$, la première formulation, qui est celle présentée dans le tableau 2, permet de minimiser le nombre de paramètres indépendants permettant d'expliquer le vote, sans perte de généralité.

Le ratio du nombre d'informations indépendantes disponible sur le nombre de paramètres indépendants à estimer est de $\frac{I_1 + 2}{(I_1 + 2) \times (I_2 + 3)} = (I_2 + 3)^{-1}$. Comme ce ratio est inférieur à 0, le modèle n'est pas identifiable et une infinité de solutions donneront des résultats équivalents.

La solution est d'analyser les résultats de chaque bureau de vote (NB) dans un modèle global en assumant que le comportement moyen d'un électeur au deuxième tour est dépendant de son comportement au premier tour mais est indépendant de son bureau de vote (on verra dans la discussion que cette hypothèse forte peut être relâchée). Le nombre d'informations indépendantes issues du dépouillement sera donc $NB \times (I_1 + 2)$ alors que le nombre de paramètres indépendants à ajuster sera toujours $(I_1 + 2) \times (I_2 + 3)$. Le ratio informations/paramètres est égal $NB \times (I_2 + 3)^{-1}$. Donc, dès que le nombre de bureaux de vote (NB) est supérieur au nombre de listes au deuxième tour plus 3 ($= I_2 + 3$), on dispose d'assez d'informations indépendantes pour estimer les paramètres du modèle.

Une autre information peut être relativement facilement collectée à partir des registres de signatures des bureaux : combien de votants au premier tour n'ont pas voté au second et inversement combien de votants au second tour n'avaient pas voté au premier. On verra que cette information se révèle importante pour limiter les possibilités numériques des solutions. Idéalement, cette information devrait être obtenue pour chacun des bureaux de vote indépendamment.

1.2. Estimation de la distribution des paramètres

Dans le modèle direct présenté dans le tableau 2, le nombre de votes C_2 est connu. Par contre, les valeurs prédites pour les votes du premier tour dépendent de la valeur des paramètres p_* . Par exemple, le nombre de votes prédits pour le premier tour (appelé M) des abstentions (M_{S_1}), des bulletins nuls ou annulés (M_{U_1}), blancs ou enveloppes vides (M_{C_1}) ou bien pour chacune des I_1 listes ($M_{D_{A_1}}$ à M_{D_1}) dans un bureau de vote sont indiqués dans le tableau 2.

Les valeurs observées pour ce bureau de vote étant respectivement N_{S_1} , N_{U_1} , N_{C_1} , N_{D_A} , N_{D_B} , ..., N_{D_1} . On appellera m_{S_1} , m_{U_1} , m_{C_1} , m_{D_A} , m_{D_B} , ..., m_{D_1} les proportions prédites dans le modèle des votes au premier tour avec $m_{x_1} = M_{x_1} / \sum M_{*1}$.

TABLEAU 2. Expression de la probabilité de vote au premier tour dépendant de la probabilité de vote au second tour quand il y a I_1 listes (A_1, B_1, \dots, I_1) au premier tour et I_2 listes au second (Z_2, Y_2, \dots, I_2). La somme de chaque ligne est égale à 1 ce qui revient à dire qu'il y a besoin de $(I_1 + 2) \times (I_2 + 3)$ probabilités pour décrire le choix des électeurs au premier tour conditionnellement à leur choix au second tour. Le nombre d'observations disponibles est $I_1 + 3$ (ligne Votes) mais comme on connaît le nombre d'électeurs inscrits pour le premier tour, il suffit de $I_1 + 2$ observations indépendantes pour décrire l'ensemble des choix C_1 . Un caractère * dans un nom de variable indique que toutes les possibilités doivent être prises en compte. La case grisée correspond aux électeurs s'étant abstenus aux deux tours, les cases vertes correspondent aux électeurs qui ont voté au premier tour mais pas au second (et inversement pour les bleues) et les cases à fond jaune correspondent aux électeurs qui ont voté aux deux tours. Les votes prédits M_* correspondent aux sommes des colonnes correspondantes. Il faut noter que le modèle n'implique pas $N_1 = N_2$

| $C_2 \backslash C_1$ | | Abstentions ₁ | Nuls ₁ | Blancs ₁ | Liste A_1 | Liste B_1 | ... | Liste I_1 |
|--------------------------|---------------|--------------------------------------|--------------------------------------|--------------------------------------|--|--|-----|---|
| | Votes | N_{S_1} | N_{U_1} | N_{C_1} | $N_{D_{1.A}}$ | $N_{D_{1.B}}$ | | $N_{D_{1.I_1}}$ |
| Abstentions ₂ | N_{S_2} | $p_{S_1 S_2} \times N_{S_2}$ | $p_{U_1 S_2} \times N_{S_2}$ | $p_{C_1 S_2} \times N_{S_2}$ | $p_{D_{1.A} S_2} \times N_{S_2}$ | $p_{D_{1.B} S_2} \times N_{S_2}$ | | $(1 - \sum p_{* S_2}) \times N_{S_2}$ |
| Nuls ₂ | N_{U_2} | $p_{S_1 U_2} \times N_{U_2}$ | $p_{U_1 U_2} \times N_{U_2}$ | $p_{C_1 U_2} \times N_{U_2}$ | $p_{D_{1.A} U_2} \times N_{U_2}$ | $p_{D_{1.B} U_2} \times N_{U_2}$ | | $(1 - \sum p_{* U_2}) \times N_{U_2}$ |
| Blanc ₂ | N_{C_2} | $p_{S_1 C_2} \times N_{C_2}$ | $p_{U_1 C_2} \times N_{C_2}$ | $p_{C_1 C_2} \times N_{C_2}$ | $p_{D_{1.A} C_2} \times N_{C_2}$ | $p_{D_{1.B} C_2} \times N_{C_2}$ | | $(1 - \sum p_{* C_2}) \times N_{C_2}$ |
| Liste Z_2 | $N_{D_{Z_2}}$ | $p_{S_1 D_{Z_2}} \times N_{D_{Z_2}}$ | $p_{U_1 D_{Z_2}} \times N_{D_{Z_2}}$ | $p_{C_1 D_{Z_2}} \times N_{D_{Z_2}}$ | $p_{D_{1.A} D_{Z_2}} \times N_{D_{Z_2}}$ | $p_{D_{1.B} D_{Z_2}} \times N_{D_{Z_2}}$ | | $(1 - \sum p_{* D_{Z_2}}) \times N_{D_{Z_2}}$ |
| Liste Y_2 | $N_{D_{Y_2}}$ | $p_{S_1 D_{Y_2}} \times N_{D_{Y_2}}$ | $p_{U_1 D_{Y_2}} \times N_{D_{Y_2}}$ | $p_{C_1 D_{Y_2}} \times N_{D_{Y_2}}$ | $p_{D_{1.A} D_{Y_2}} \times N_{D_{Y_2}}$ | $p_{D_{1.B} D_{Y_2}} \times N_{D_{Y_2}}$ | | $(1 - \sum p_{* D_{Y_2}}) \times N_{D_{Y_2}}$ |
| ... | | | | | | | | |
| Liste I_2 | $N_{D_{I_2}}$ | $p_{S_1 D_{I_2}} \times N_{D_{I_2}}$ | $p_{U_1 D_{I_2}} \times N_{D_{I_2}}$ | $p_{C_1 D_{I_2}} \times N_{D_{I_2}}$ | $p_{D_{1.A} D_{I_2}} \times N_{D_{I_2}}$ | $p_{D_{1.B} D_{I_2}} \times N_{D_{I_2}}$ | | $(1 - \sum p_{* D_{I_2}}) \times N_{D_{I_2}}$ |
| Votes prédits | | $M_{S_1} = \sum \uparrow$ | $M_{U_1} = \sum \uparrow$ | $M_{C_1} = \sum \uparrow$ | $M_{D_{1.A}} = \sum \uparrow$ | $M_{D_{1.B}} = \sum \uparrow$ | | $M_{D_{1.I_1}} = \sum \uparrow$ |

La formulation du modèle peut sembler contre-intuitive puisque finalement ici on cherche les paramètres p_* qui permettent de prédire les comportements de vote au premier tour C_1 connaissant les comportements de vote au deuxième tour C_2 . Le sens de cette formulation permet de gagner de l'information car, par le jeu d'élimination et de fusion des listes, on a $I_1 \geq I_2$, on a donc plus d'informations au premier tour qu'au deuxième s'il y a moins de listes au deuxième tour.

La vraisemblance (L , pour likelihood) de ces données observées dans ce modèle pour ce bureau de vote i peut être déterminée grâce à une distribution multinomiale [20] à une constante de proportionnalité près :

$$L_i \# m_{S_1}^{N_{S_1}} \times m_{U_1}^{N_{U_1}} \times m_{C_1}^{N_{C_1}} \times m_{D_{1,A}}^{N_{D_{1,A}}} \times m_{D_{1,B}}^{N_{D_{1,B}}} \times \dots \times m_{D_{1,I_1}}^{N_{D_{1,I_1}}}$$

La vraisemblance à l'échelle de tous les bureaux de vote (L) est alors proportionnelle au produit de la vraisemblance de chacun des bureaux de vote :

$$L \# \prod_{i=1}^{NB} L_i$$

Pour faciliter les calculs, l'opposé du logarithme de la vraisemblance sera utilisé. Lors de la recherche des valeurs des paramètres, il est important que la somme des probabilités d'une ligne du tableau 2 fasse 1. Pour s'assurer que c'est bien toujours le cas, des probabilités conditionnelles sont implémentées dans le modèle informatique sans aucune conséquence pour la généralité du modèle présenté ici.

Pour déterminer les paramètres décrivant le comportement des électeurs, une approche utilisant le maximum de vraisemblance [5] et une approche bayésienne utilisant l'algorithme de Metropolis-Hastings [15] seront utilisées. Le maximum de vraisemblance sera utilisé pour déterminer une valeur initiale des paramètres. L'algorithme Metropolis-Hastings est une méthode de Monte Carlo avec chaîne de Markov (MCMC) pour obtenir une séquence d'échantillons aléatoires dont la distribution stationnaire est la distribution recherchée.

Au cours du processus d'itération, une chaîne de Markov est construite en utilisant les valeurs actuelles de paramètres π_t (π_t représente le jeu de paramètres à l'itération t) sur lesquelles une fonction aléatoire gaussienne définie par les écart-types s appliqués à chacun des paramètres, $\pi_{t+1} = N(\pi_t, s)$. Cette étape correspond au processus de Monte-Carlo. La valeur des écarts types s sont un compromis entre deux contraintes : si les valeurs sont trop élevées, les nouvelles valeurs π_{t+1} peuvent donner des résultats très loin de la solution optimale, tandis que si elles sont trop faibles, le modèle peut être piégé dans des minimums locaux. Un modèle de distribution adaptative permettant de changer les valeurs s au cours des itérations [23], tel qu'implémentée dans le package R HelpersMG [9], garantira que le taux d'acceptation soit proche de 0,234, qui est le taux d'acceptation optimal [22]. Tous les priors des paramètres sont choisis dans une distribution uniforme comprise entre 0 et 1 pour ne faire aucune hypothèse forte sur les valeurs des paramètres. Comme le point de départ de la chaîne de Markov est l'estimation des paramètres obtenue par maximum de vraisemblance, aucune phase de *burn-in* ne sera implémentée et 10 000 itérations sont utilisées pour établir la distribution des paramètres.

Le code R [18] utilisé pour effectuer ces estimations est disponible auprès de l'auteur.

2. Application de la méthode sur l'élection municipale de 2020

Le premier tour de l'élection municipale s'est tenu le 15 mars 2020, en début de pandémie de maladie à coronavirus (SARS-CoV-2). Le taux d'abstention a été de 55,34 % des électeurs inscrits à l'échelle nationale. En raison de la pandémie, le second tour, initialement prévu le 22 mars, est

d'abord reporté *sine die*, puis au 28 juin suivant. L'abstention au second tour est de 58,4 % des électeurs inscrits à l'échelle nationale.

Chaville (Hauts-de-Seine), Lavandou (Var) et Choisy-le-Roi (Val-de-Marne) sont trois communes ayant vu un rapprochement au second tour d'une liste soutenue par EELV (Europe Écologie Les Verts) avec une liste du centre ou de droite. Cette configuration était originale puisqu'en général une fusion des listes soutenues par EELV et des listes de gauche était observée. Le comportement des électeurs ayant porté leur voix sur les listes soutenues par EELV au premier tour était inconnu comme l'analyse France 3 le 16 juin 2020 : « Une élection qui s'annonce donc bien complexe, alors que personne ne peut prédire si les électeurs suivront les consignes de vote des responsables politiques. » [8].

Le contexte électoral Chavillois et Choisyien a été analysé grâce à une technologie d'intelligence artificielle (IA) et alors que l'IA a bien prédit la victoire de la liste de rassemblement du centre, de la droite et d'EELV à Chaville, elle a en revanche échoué à prédire celle du même rassemblement à Choisy-le-Roi [14]. Ceci démontre la difficulté d'anticiper le comportement des électeurs pour des rassemblements inédits même avec les technologies les plus récentes. En conséquence, nous concentrerons notre analyse sur le contexte de Choisy-le-Roi pour tenter de comprendre les modalités de report de voix. En effet les prédictions ont été rendues encore plus compliquées en prenant en compte que le député LREM ne soutient pas la candidate LREM au second tour : « Autre curiosité : le député LREM de la circonscription, Jean-François Mbaye, a lui aussi apporté son soutien à Tonino Panetta car "favorable à un basculement de la ville", rapporte le média local 94 Citoyens, et non à la candidate LREM Nathalie Lemoine lui reprochant de ne pas avoir fusionné sa liste avec le candidat Divers droite. » [2].

2.1. Les élections municipales de 2020 à Choisy-le-Roi (Val-de-Marne)

Choisy-le-Roi est une commune du Val-de-Marne ayant 45 331 habitants en 2017 [19] et 21 207 puis 21 251 électeurs inscrits sur les listes électorales respectivement au premier et deuxième tour de l'élection municipale de 2020. Le conseil municipal est présidé depuis 2014 par Mr Didier Guillaume qui a été élu avec une étiquette PCF (Parti Communiste Français) et se représentait en 2020 comme tête de liste divers gauche. Un total de 6 listes se présentaient au premier tour : Ali Id Elouali – Choisy demain (CD) (EELV), Josefa Torres – Lutte ouvrière – Faire entendre le camp des travailleurs (LO), Robin Albert – Choisyens à vous la main (CEVLM) (divers), Nathalie Lemoine – Choisy C vous (CCV) (LREM), Tonino Panetta – Gagner ensemble (GE) (Divers droite) et Didier Guillaume – Choisy en commun (CEC) (PCF, divers gauche).

À l'issue du premier tour, 4 listes étaient en position de se maintenir (CCV 10,70 % ; CD 15,44 % ; CEC 29,48 % et GE 33,19 %). Les listes CD et GE ont fusionné et le second tour s'est joué avec 3 listes : Nathalie Lemoine – Choisy C vous (CCV) (Le mouvement radical et l'alliance centriste, LREM), Tonino Panetta – Gagner ensemble (GE) (Divers droite, EELV) et Didier Guillaume – Choisy en commun (CEC) (PCF, divers gauche). Les résultats du second tour sont les suivants : GE 4531 voix (21,32 % ; 55,31 %) ; CEC 3074 voix (14,47 % ; 37,52 %) ; CCV 587 voix (2,76 % ; 7,17 %) (le premier pourcentage représente le pourcentage de vote par rapport aux inscrits et le second par rapport aux suffrages exprimés). Le conseil municipal est donc composé de 34 élus de la liste GE, 8 élus de la liste CEC et 1 élu de la liste CCV. Mr Tonino Panetta a été élu maire le 4 juillet 2020.

2.2. Résultats de la modélisation du comportement des électeurs choisyiens

L'ajustement du modèle par maximum de vraisemblance puis par une approche bayésienne conduit à une matrice de 56 combinaisons de votes (tableaux 3). L'analyse a été effectuée d'abord

TABLEAU 3. (A) Prédiction par méthode bayésienne des modalités de répartition des voix au second tour (lignes) en fonction des choix des électeurs au premier tour (colonnes) pour un pourcentage de 90 % d'électeurs n'ayant pas voté au second tour qui n'avaient pas voté au premier tour non plus. Le nombre entre parenthèses est l'erreur standard corrigée pour l'autocorrélation en prenant une valeur toutes les 10 itérations ([21]). (B) Sensibilité des résultats mesurée en nombre de voix de différence par % d'électeurs n'ayant pas voté au second tour qui n'avaient pas voté au premier tour, mesurés pour 89 % et 91 %. Les cases grisées correspondent à des électeurs qui ont voté pour la même liste ou la liste de fusion au premier et deuxième tour.

| (A) : Prédictions bayésiennes | | Premier tour | | | | | | | | |
|-------------------------------|-------------|--------------------|------------------|-----------------|-------------------|-------------------|-------------------|------------------|------------------|-----------------|
| | | Abstentions | Nuls | Blancs | GE | CEC | CD | CCV | CAVLM | LO |
| Second tour | Abstentions | 11549,98 (0,98) | 58,68 (1,46) | 32,96 (0,78) | 45,44 (1,37) | 559,37 (6,01) | 34,09 (0,97) | 136,07 (2,98) | 278,01 (3,41) | 55,39 (1,21) |
| | Nuls | 58,65 (1,32) | 1,70 (0,12) | 0,69 (0,05) | 23,77 (0,70) | 38,56 (1,03) | 14,60 (0,54) | 7,42 (0,36) | 2,96 (0,17) | 0,65 (0,05) |
| | Blancs | 78,98 (1,44) | 3,24 (0,20) | 4,36 (0,29) | 30,13 (0,86) | 20,00 (0,65) | 11,68 (0,50) | 4,92 (0,24) | 3,25 (0,17) | 3,44 (0,24) |
| | GE | 268,99 (4,38) | 105,00 (1,76) | 20,30 (0,64) | 2378,78 (4,08) | 268,45 (4,21) | 1071,76 (2,41) | 364,82 (4,39) | 34,78 (1,01) | 17,10 (0,62) |
| | CEC | 891,17 (3,88) | 53,52 (1,25) | 34,57 (0,84) | 48,99 (1,38) | 1411,57 (4,98) | 66,36 (1,61) | 85,94 (2,02) | 426,68 (3,36) | 55,21 (1,18) |
| | CCV | 67,88 (1,86) | 12,45 (0,56) | 2,66 (0,12) | 134,70 (3,08) | 61,14 (1,61) | 39,87 (1,08) | 253,92 (3,31) | 10,75 (0,40) | 3,63 (0,19) |

| (B) : Sensibilité | | Premier tour | | | | | | | | |
|-------------------|-------------|--------------|-------|--------|------|--------|------|--------|--------|-------|
| | | Abstentions | Nuls | Blancs | GE | CEC | CD | CCV | CAVLM | LO |
| Second tour | Abstentions | 121.56 | -5.11 | -0.59 | 0 | -74.23 | 0 | -23.94 | -14.65 | -3.04 |
| | Nuls | 0 | 0.11 | 0.17 | 0 | 0.2 | -0.4 | -0.07 | 0 | 0 |
| | Blancs | -1.03 | 0.41 | 0.36 | 0 | -0.08 | 0.24 | 0 | 0 | 0.1 |
| | GE | -46.67 | 0 | 0 | 3.38 | 27.65 | 1.25 | 14.39 | 0 | 0 |
| | CEC | -84.15 | 4.19 | 0.11 | 0 | 49.22 | 0.6 | 11.6 | 15.34 | 3.1 |
| | CCV | 0 | 0.64 | 0 | 0.32 | 0 | 0 | -0.96 | 0 | 0 |

sans prendre en compte l'information correspondant à la proportion d'électeurs s'étant abstenue au premier ou au second tour, et ensuite en postulant que 85, 89, 90, 91 et 95 % des électeurs qui se sont abstenus au second tour s'étaient déjà abstenus au premier tour.

Lorsque la proportion d'électeurs s'étant abstenus au premier ou au second tour est laissée sans contrainte, l'ajustement est très bon ($r^2 = 0,996$) mais l'analyse fine du comportement des électeurs tel qu'il est modélisé est incohérente : le meilleur ajustement numérique est obtenu lorsque les abstentionnistes du premier tour votent au second et les abstentionnistes du second avaient tous votés au premier. Le comportement du modèle non-contraint est juste d'un point de vue numérique mais n'a à peu près aucun intérêt d'un point de vue sociologie du vote.

Au contraire, lorsqu'on contraint à ce que 90 % des électeurs qui se sont abstenus au second tour s'étaient déjà abstenus au premier tour, on parvient toujours à très bien prédire le comportement des électeurs ($r^2 = 0,997$, figure 1) et les résultats prennent alors plus de sens (tableau 3A) et sont relativement stables tant que le % des électeurs qui se sont abstenus au second tour s'étaient déjà abstenus au premier tour reste élevé (tableau 3B).

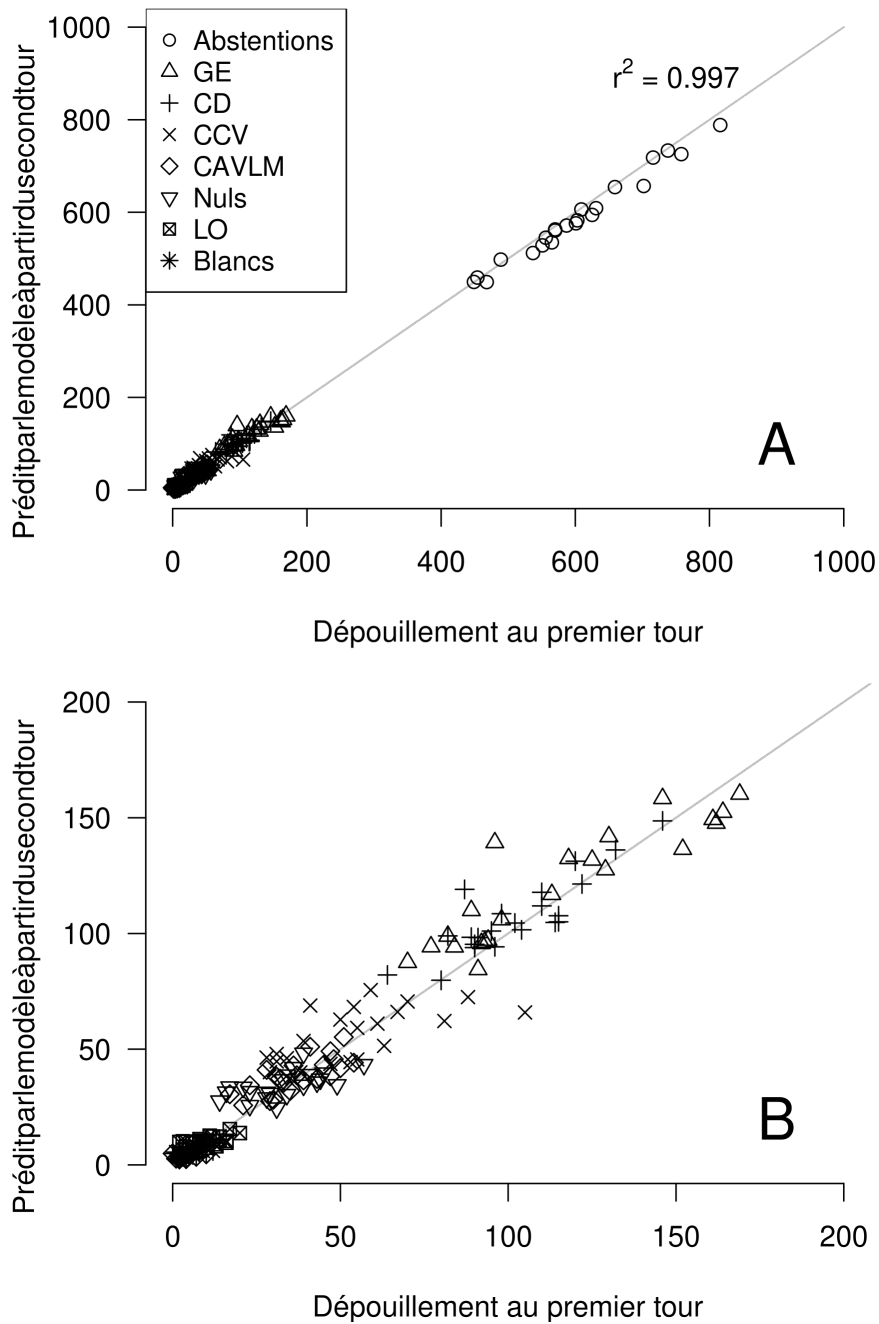


FIGURE 1. Prédications du nombre de voix au premier tour en fonction du nombre de voix au second tour de l'élection municipale à Choisy-le-Roi en 2020 (A ensemble des données, B agrandissement sur la partie correspondant aux non-abstentionnistes). Les abréviations du nom des listes sont données dans le texte.

On peut remarquer que les résultats sont conformes à la logique puisqu'en majorité le choix des électeurs du second tour est en correspondance avec leur choix au premier tour. Ainsi les électeurs ayant voté pour les listes GE ou CD au premier tour ont voté en très grande majorité pour la liste de fusion GE au second tour tandis que les électeurs ayant voté pour la liste CEC ou la liste CCV ont aussi suivi leur liste. L'erreur standard de ces prédictions, qui mesure comment cette moyenne est connue, apparaît relativement faible (nombres entre parenthèses du tableau 3A) et implique que l'on peut chercher à interpréter ces résultats avec une certaine confiance.

2.3. *Analyse du report de voix entre le premier et le second tour*

À Choisy-le-Roi, la liste de fusion (GE) a réuni au second tour plus d'électeurs que la somme de ce qu'avaient réuni les deux listes participant à la fusion (CD, GE) au premier tour (4531 contre 3714). Cependant, pour éviter une analyse erronée, il faut prendre en compte le taux d'abstention différent au deuxième tour par rapport au premier (63,98 % vs. 61,45 % à Choisy-le-Roi) et le fait que deux listes n'ont pu se maintenir au second tour et donc constituaient un réservoir de voix disponibles pour les listes présentes au second tour.

On observe que la liste de fusion (GE) a obtenu un très bon report des voix à la fois de la liste GE (droite) et CD (Écologistes) ; les électeurs ont donc suivi leur tête de liste du premier tour. Cette observation contredit les prédictions qui avaient été faites par Mr Fabien Guillaud-Bataille, secrétaire départemental du PCF Val-de-Marne, qui s'était déclaré « dubitatif sur le report des voix écologistes sur le candidat LR » [2]. On peut noter aussi que les électeurs écologistes n'ont pas suivi les consignes d'EELV Val-de-Marne qui avait suspendu les membres du parti apparaissant sur la liste de fusion [2]. Par contre, les électeurs de la liste CEC (gauche) ont eu un comportement différent : on note une abstention importante et même une fraction importante de ces électeurs qui ont voté pour la liste de fusion GE ce qui peut sembler de prime abord surprenant. Mais en prenant en compte que dans la liste de fusion GE, la composante écologiste est souvent classée à gauche, on voit que la frontière entre les listes a été poreuse. Le comportement des électeurs ayant voté pour la liste CCV apparaît au contraire beaucoup plus labile. Une majorité d'entre eux auraient même voté pour la liste de fusion GE comme le demandait le député LREM Jean-François Mbaye [8] et un nombre important se serait abstenu (tableau 3A). Les électeurs de la liste CAVLM (liste citoyenne) ont rejoint majoritairement la liste CEC (gauche) mais 1/3 d'entre eux se seraient abstenus. Le comportement des électeurs ayant voté LO au premier tour est plus difficile à identifier en raison de leur faible nombre. On note cependant une certaine logique puisqu'ils se répartissent en égale proportion entre un vote pour la liste CEC (gauche) et l'abstention.

Les abstentionnistes du premier tour, quand ils ont voté au second tour, ont choisi majoritairement la liste CEC (gauche) et dans une moindre mesure la liste GE (droite et écologistes). Il convient de se rappeler que l'abstention est une expression politique [4].

L'analyse fine des raisons de ces reports ou non-reports de voix rentre dans le cadre de l'analyse politique et sociologique de ces comportements. En l'absence d'enquête sur le terrain ou d'interview des électeurs, il est préférable, pour l'auteur de cette étude, de rester au niveau de la description objective du comportement électoral.

3. **Difficultés rencontrées et extensions possibles du modèle**

Le modèle présenté ici est applicable à tout type d'élection à deux tours comme ici les élections municipales en France, mais aussi l'élection présidentielle (scrutin uninominal majoritaire à deux tours), l'élection des sénateurs (scrutin majoritaire à deux tours ou scrutin de liste à un seul tour selon le nombre de postes à pourvoir), des députés (scrutin majoritaire à deux tours), des conseillers régionaux (scrutin de liste à deux tours), des conseillers départementaux (scrutin

majoritaire uninominal à deux tours). Si de nombreux bureaux de vote doivent être analysés, le temps de calcul peut être assez long. Pour résoudre cette difficulté, une augmentation du nombre de cœurs utilisés pour les calculs est envisageable car l'algorithme est trivialement parallélisable (« embarrassingly parallel problems ») [12] puisque les bureaux de vote sont analysés indépendamment les uns des autres lors du calcul de la vraisemblance.

Le modèle présenté ici ne prend en compte qu'une matrice de transition C_1/C_2 mais il est tout à fait possible d'en définir plusieurs qui peuvent s'appliquer différemment selon les caractéristiques sociologiques des populations des bureaux de vote. On peut alors décider de la pertinence de la prise en compte des différences entre bureau de vote en utilisant une statistique de type Akaike Information Criterion (AIC) qui est une mesure de la qualité d'un modèle pénalisée par sa complexité [3].

Le plus gros problème concerne la capacité à limiter la possibilité pour le modèle à trop bien expliquer les données en affectant sur les abstentions des premiers et seconds tours, des votes à l'autre tour. En effet, dans les cas où le nombre d'abstentions est important, la meilleure solution numérique consiste à postuler que les abstentionnistes du premier tour ne sont pas les mêmes que ceux du deuxième. Bien que valide d'un point de vue numérique, cette solution n'est pas satisfaisante car une forte proportion des abstentionnistes du premier tour seront aussi des abstentionnistes au second tour. Il est aisé d'évaluer cette proportion en comparant les signatures sur les listes d'émargement dans les bureaux de votes au premier et second tour. On peut alors dénombrer les électeurs inscrits aux deux tours et ayant voté à l'un ou à l'autre des tours. Cette valeur peut alors être utilisée pour contraindre le modèle qui fournira des estimations plus plausibles du comportement de vote des électeurs comme nous l'avons vu précédemment. Si cette valeur ne peut pas être obtenue (voir la partie discussion en ce qui concerne la possibilité légale d'obtenir l'accès aux registres d'émargement), on est obligé de choisir une gamme de valeurs, par exemple 85 %, 89 %, 90 %, 91 % et 95 %, correspondant à la probabilité qu'un électeur n'ayant pas voté au premier tour n'ait pas voté non plus au second, et sur cette base on peut effectuer une analyse de sensibilité [10]. C'est la démarche qui a été utilisée ici (tableau 3B) puisque l'accès aux registres d'émargement n'a pas été possible.

Généralement, les listes électorales du premier et du second tour sont les mêmes puisqu'il n'y a qu'une semaine entre les tours ce qui ne permet pas de réunir une commission de révision des listes électorales. S'il y a des décès entre les deux tours, ils seront comptés comme des abstentions au second tour. Ceci n'entraîne pas de biais si le décès est indépendant du vote au premier tour. Cette hypothèse est cependant probablement fautive en raison de deux phénomènes liés : la probabilité de décès et le choix lors du vote qui ne sont pas indépendants de l'âge. Cependant cet effet est extrêmement marginal puisque, à l'échelle du territoire français, le nombre moyen de décès journalier d'individus de plus de 24 ans est de 1840 personnes (données du 15 mars au 28 juin 2020 [13]), soit 12 884 décès entre deux tours séparés de 7 jours. En considérant que le taux d'inscription sur les listes électorales est de 93 % [6], la probabilité de décès d'un électeur inscrit sur les listes électorales est de $2,5 \cdot 10^{-4}$ ($0,93 \times 12884/47,1 \cdot 10^6$). Le nombre moyen de décès entre deux tours séparés d'une semaine est donc de 0,23 électeur par bureau de vote (en considérant une moyenne de 900 électeurs par bureau de vote). Pour le cas particulier de l'élection municipale de 2020, 106 jours ont séparé le premier du second tour ce qui conduit à une estimation moyenne de 3,46 décès par bureau de vote. Cette valeur, bien que plus forte que ce qui est habituellement noté en raison du surcroît de mortalité lié à la pandémie et à la durée inhabituelle entre les deux tours, est peu susceptible d'avoir influencé les résultats de l'analyse de l'élection présentés ici.

Les listes électorales entre les deux tours de l'élection municipale de 2020 ont été révisées à Choisy-le-Roi de façon automatique par l'ajout d'électeurs ayant atteint 18 ans et par le retrait de personnes décédées. Par rapport au modèle présenté dans le tableau 2, on pourrait rajouter

une colonne et une ligne avec respectivement les électeurs qui seront retirés et rajoutés des listes électorales entre les deux tours des élections. Cependant ceci n'aurait de sens que si :

- (1) le nombre d'inscrits impacté est important et
- (2) on a des raisons de penser que les électeurs retirés ou rajoutés ne sont pas un échantillon aléatoire de ceux présents au premier tour.

Concernant le point (1), le différentiel du nombre d'électeurs est très faible (+44 inscrits sur 22 251 inscrits au second tour) et il n'y a aucune raison de penser que le choix de ces électeurs ait été différent des autres électeurs choisyens. En conséquence, le modèle utilisé ici ne différencie pas les électeurs inscrits aux deux tours de ceux ayant été intégrés ou étant sortis des listes électorales entre les deux tours.

4. Conclusions

La compréhension du choix des électeurs est un élément très important de l'analyse du ressenti dans le domaine de la politique. Au niveau national, les enquêtes d'opinion permettent d'avoir accès à une information, certes potentiellement biaisée [7], mais au moins existante. À une échelle plus fine, il n'existe pas d'outil simple d'accès permettant d'interpréter le choix des électeurs. La méthodologie présentée ici permet d'établir les modalités de choix des électeurs entre le premier et le second tour d'une élection mais aussi l'incertitude sur ces modalités de choix. Ainsi, l'analyse des résultats de l'élection municipale des 15 mars et 28 juin 2020 à Choisy-le-Roi éclaire d'un jour nouveau le comportement des électeurs dans une configuration inédite : la fusion d'une liste écologiste avec une liste de droite.

L'article L68 de la loi n°2013-403 du 17 mai 2013 précise que les listes d'émargement déposées à la préfecture ou à la sous-préfecture sont communiquées à tout électeur requérant pendant un délai de dix jours à compter de l'élection. Une demande d'avoir accès à ces listes d'émargement après cette période de 10 jours dans le cadre du travail présenté ici s'est vu notifier d'un refus sur la base à la fois de l'article L68 de la loi n°2013-403 et de l'avis 20082653 (séance du 03/07/2008) de la commission d'accès aux documents administratifs sur les motifs suivants : « La commission estime que ces dispositions particulières font obstacle à l'application de la loi du 17 juillet 1978 jusqu'à l'expiration du délai de dix jours à compter de l'élection. Passé ce délai, ces documents administratifs ne sont, en tout état de cause, pas communicables sur le fondement de la loi de 1978 dès lors qu'elles révèlent le choix d'électeurs nommément désignés de se rendre ou non aux urnes, choix qui relève du secret de la vie privée. ». Sur la base d'informations obtenues à la Préfecture des Hauts-de-Seine (Direction de la Citoyenneté et de la Légimité – Bureau de la Réglementation Générale et des Élections), les listes d'émargement ne sont pas conservées, même sous forme numérique. On peut regretter cette situation car, on le voit ici, l'information disponible sur des listes d'émargement est indispensable pour faire une analyse correcte des modalités de vote. Cependant le stockage sous forme papier des 100 000 listes d'émargement correspondant aux deux tours des élections municipales (estimation sur la base d'environ 900 électeurs par bureau de vote) aurait un coût certain pour l'état. Une solution de numérisation des listes d'émargement pourrait sembler, dans un premier temps, plus adéquate, mais la maintenance de ces données aurait un coût énergétique dont l'intérêt peut sembler disproportionné par rapport aux enjeux. Rallonger le délai de consultation des listes d'émargement pourrait être une solution intermédiaire intéressante.

En conclusion, le comportement de report des voix lors du second tour a pu être modélisé avec une bonne fiabilité et a permis d'avoir des enseignements qui étaient inaccessibles sans cet outil :

- (1) les électeurs écologistes n'ont pas suivi les consignes du parti EELV qui a suspendu le tête de liste de la liste Choisy Demain (CD) suite à la fusion de cette liste avec la liste divers

droite Gagner Ensemble (GE) ; le report des voix des électeurs de CD vers la liste de fusion GE a été quasi-total,

- (2) l’alliance avec les écologistes n’a pas fait perdre de voix des électeurs GE du premier tour vers la liste de fusion GE, et
- (3) le maintien de la candidate LREM contre l’avis du député LREM de la circonscription explique sans doute la forte dispersion des voix des électeurs de la liste Choisy C’est Vous (CCV) qui ont voté majoritairement pour la liste de fusion divers droite-écologistes (GE).

Le comportement des électeurs de Choisy-le-Roi tel qu’il est décrit ici correspond à un contexte local et n’est sans doute pas généralisable à toutes les situations similaires.

Remerciements

L’auteur tient à remercier Mr Adlane Ait Djaoud du Service Accueil Famille et Citoyenneté à la mairie de la ville de Chaville, l’adjoint au directeur du Bureau de la Réglementation Générale et des Élections (Direction de la Citoyenneté et de la Légalité) de la Préfecture des Hauts-de-Seine et le service « Populations » de la Mairie de Choisy-le-Roi pour leur aide à l’obtention d’informations utilisées dans cette publication.

Références

- [1] « Code électoral, Version consolidée au 30 juin 2020 », 2020, <https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006070239>, (page consultée le 10 août 2020).
- [2] 94Citoyens. com, « Municipales à Choisy-le-Roi : fusion Panetta (LR) – Id Elouali (EELV) contre Guillaume (PCF) », 2020.
- [3] H. Akaike, « A new look at the statistical model identification », *IEEE Trans. Autom. Control* **19** (1974), p. 716-723.
- [4] C. Braconnier, J.-Y. Dormagen, « Ce que s’abstenir veut dire : Une autre forme d’expression politique », *Le Monde diplomatique* **Mai** (2014), p. 1-5.
- [5] R. H. Byrd, P. Lu, J. Nocedal, C. Zhu, « A limited memory algorithm for bound constrained optimization », *SIAM J. Sci. Comput.* **16** (1995), p. 1190-1208.
- [6] M. Demotes-Mainard, « 47,1 millions d’électeurs présents en avril 2019 sur les listes électorales françaises », 2020, Programme Elire, INSEE.
- [7] C. Flament, C. Guimelli, J.-C. Abric, « Effets de masquage dans l’expression d’une représentation sociale », *Les Cahiers Internationaux de Psychologie Sociale* **69** (2006), p. 15-31.
- [8] France 3, « Municipales 2020 : les surprenantes alliances à Choisy-le-Roi », 2020, <https://france3-regions.francetvinfo.fr/paris-ile-de-france/val-de-marne/municipales-2020-surprenantes-alliances-choisy-roi-1841372.html>.
- [9] M. Girondot, « HelpersMG : Tools for Environmental Analyses, Ecotoxicology and Various R Functions », 2021, The Comprehensive R Archive Network.
- [10] D. M. Hamby, « A review of techniques for parameter sensitivity analysis of environmental models », *Environmental Monitoring and Assessment* **32** (1994), p. 135-154.
- [11] J. C. Heckelman, « The effect of the secret ballot on voter turnout rates », *Public Choice* **82** (1995), p. 107-124.
- [12] M. Herlihy, N. Shavit, *The art of multiprocessor programming*, revised first éd., Elsevier, 2012.
- [13] INSEE, « Nombre de décès quotidiens. France, régions et départements », 2020, INSEE.
- [14] linternute, « ElectionLab. Une intelligence artificielle peut-elle prédire les résultats des municipales 2020? », 2020, <https://www.election-lab.fr/>.
- [15] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, « Equations of state calculations by fast computing machines », *J. Chem. Phys.* **21** (1953), p. 1087-1092.
- [16] C. Nicolet, « Cicéron, Platon et le vote secret », *Historia : Zeitschrift für Alte Geschichte* **19** (1970), p. 36-66.
- [17] H. Portelli, J.-P. Sueur, « Sondages et démocratie : pour une législation plus respectueuse de la sincérité du débat politique », 2010, Rapport d’information n° 54 (2010-2011) fait au nom de la commission des lois, déposé le 20 octobre 2010, Paris, France.
- [18] R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, 2021.
- [19] A. G. Ramm, *Inverse problem : Mathematical and analytical techniques with applications to engineering*, Springer, 2005.

- [20] C. R. Rao, « Maximum likelihood estimation for the multinomial distribution », *Sankhyā, Ser. A* **18** (1957), p. 139-148.
- [21] G. O. Roberts, « Markov chain concepts related to sampling algorithms », in *Markov Chain Monte Carlo in Practice*, Chapman & Hall, 1996, p. 45-58.
- [22] G. O. Roberts, J. S. Rosenthal, « Optimal scaling for various Metropolis-Hastings algorithms », *Stat. Sci.* **16** (2001), p. 351-367.
- [23] J. S. Rosenthal, « Optimal proposal distributions and adaptive MCMC », in *MCMC Handbook*, Chapman & Hall/CRC, 2011, p. 93-112.
- [24] A. Tarantola, *Inverse problem theory and model parameter estimation*, Society for Industrial and Applied Mathematics, 2005.
- [25] M. Visser, « Voting : A Behavioral Analysis », *Behavior and Social Issues* **6** (1996), p. 23-34.