

UNIVERSITÉ DU QUÉBEC

INSTITUT NATIONAL DE LA RECHERCHE SCIENTIFIQUE

CENTRE EAU TERRE ENVIRONNEMENT

**Évaluer et projeter l'effet de la météorologie sur les décès maladies
cardiovasculaires les plus mortelles au Québec**

Rapport d'examen doctoral

Programme de Ph.D. en Sciences de l'eau – EAU804

Fatima Laanaya

Jury d'évaluation

Examineur externe	Lynne Moore Université Laval
Examineur interne	Pierre Gosselin, INSPQ, INRS
Co-directeur de recherche	Belkacem Abdous Université Laval
Directeur de recherche	Fateh Chebana INRS

20 juin 2014

TABLE DES MATIÈRES

1. Mise en contexte	3
2. L'état de l'art.....	4
2.1. Les variables confondantes et les effets modificateurs de la météorologie	5
2.1.1. Les facteurs de confusion mesurés.....	5
2.1.2. Les facteurs de confusion non mesurés.....	8
2.2. Importance des variables à délais	8
2.3. Méthodes statistiques utilisées.....	9
3. Problématique.....	11
3.1. Les aspects de la relation météo/décès.....	11
3.2. La stratification	12
3.3. Données sanitaires agrégées	12
4. Pertinence du projet.....	14
4.1. But du projet.....	14
4.2. Contribution à l'action 21-volet santé et à la santé publique.....	14
4.3. Objectifs du projet	15
4.3.1. Partie A : Procédure méthodologique	15
4.3.2. Partie B : GEE-Modèle multiniveau	15
4.3.3. Partie C : Données sanitaires agrégées.....	16
5. Méthodologie.....	17
5.1. Variables étudiées.....	17
5.1.1. Les variables sanitaires et individuelles	18
5.1.2. Variables environnementales	19
5.1.3. Variables Contextuelles.....	24

5.2.	Constitution des banques de données	25
5.3.	Méthodologies statistiques	26
5.3.1.	Partie A : Procédure méthodologique	26
5.3.2.	Partie B : GEE-Modèle multiniveau	29
5.3.3.	Partie C : Données sanitaires agrégées.....	31
6.	Apport à la recherche	35
6.1.	Volet épidémiologique.....	35
6.2.	Volet méthodologique	36
7.	Implémentation.....	38
8.	Étapes de réalisations et avancement	39
8.1.	Étapes de réalisation	39
8.2.	Travaux effectués et résultats.....	41
8.3.	Travaux à venir	49
	Bibliographie.....	51
	Annexe	61

1. Mise en contexte

Au début des années 90, l'incidence des changements climatiques (CC) sur la santé humaine n'était pas prise en compte dans la littérature publiée. Cela s'explique aussi par le peu de références dans le premier rapport (GIEC, 1990). Les années 2000 ont connu le lancement de plusieurs recherches sur cette incidence (GIEC, 2001) et surtout avec le développement des instruments météorologiques et des outils statistiques. Ces recherches ont mis en évidence cette incidence sur la santé humaine (p. ex. Braga et al., 2002; Giroux, 2011b).

Le climat du Québec a connu des changements significatifs au cours des dernières décennies (Ouranos, 2010). En l'occurrence, les températures journalières moyennes ont augmenté de 0,2°C à 0,4°C par décennie depuis 1960 dans le sud de la province, incluant une hausse de la fréquence de fortes précipitations. Selon les projections annoncées par le GIEC (GIEC, 2013), ce réchauffement s'accroîtra selon la saison et le lieu géographique. D'ici 2050, le climat du Québec n'y fera pas exception (Ouranos, 2010). L'effet de ce réchauffement sera ressenti particulièrement en hiver (cela étant dit, on aura encore des vagues de froid intense), mais également en été avec une hausse de la température moyenne de 1,9 °C à 3,0 °C au Sud et de 1,6 °C à 2,8 °C au Nord, par rapport au climat de 1961-1990. Il faut noter que cette croissance n'a pas touché que la température moyenne mais aussi la fréquence et l'intensité des températures extrêmes (GIEC, 2013; WHO, 2009). Une adaptation de la santé publique aux CC est une question nécessaire, inévitable pour remédier aux effets néfastes de ces derniers sur la santé au cours des prochaines décennies (Huang et al., 2011).

L'étude de la relation entre la mortalité ou la morbidité et la météorologie dans un contexte de CC demeure, néanmoins, un champ de recherche relativement récent (WHO, 2009), voir même un axe principal de recherche en santé publique à cause de ces CC et leurs impacts. L'intérêt à l'évaluation de cette relation a fortement augmenté avec ces CC et augmentera dans le futur (McMichael et al., 2006).

Au Québec, la mortalité par maladies cardiovasculaires (MCV) diminue depuis quelques années. Malgré cela, elle demeure la deuxième cause de décès derrière les tumeurs, avec un nombre s'élevant à 14 342 en 2012, soit 24,2 % des décès répertoriés par l'Institut de la statistique du Québec (ISQ, 2012). Il est donc possible que la tendance à la baisse de la mortalité par MCV s'inverse dans un contexte de CC, d'autant qu'au Québec la part dans la population des 65 ans et plus ne cesse de croître, passant de 12 % en 2001 à 24 % en 2025, (ISQ, 2000), et que ce groupe de personnes s'avère plus vulnérable au réchauffement climatique que celui des 15 à 65 ans (Doyon et al., 2008).

L'existence d'un lien direct entre les taux de mortalité MCV et le stress thermique, chaleur ou froid, est avérée (Basu & Samet, 2002), (Besancenot, 2002). Comme le suggèrent diverses études dans différentes régions du monde (p. ex. Braga et al., 2002; Carder et al., 2005; Hori. A 2012; Ren et al., 2011), l'augmentation de fortes chaleurs estivales pourrait peser lourd sur la santé cardiovasculaire. De même pendant les jours extrêmement froids, il y a une forte augmentation de la mortalité MCV et surtout les arrêts cardiaques (Medina-Ramón et al., 2006). Des études européennes et américaines ont d'ailleurs documenté le lien entre le froid et la mortalité par MCV. Ces études sont recensées dans quelques revues systématiques dont celle de Bustinza et coll. (2010).

2. L'état de l'art

Les études des séries chronologiques sont populaires (Gosling et al., 2009) et largement utilisées pour étudier l'effet de l'exposition à court terme (p. ex. température) sur la santé (p. ex. nombre de décès) (Basu, 2009; Ye et al., 2012). Les données utilisées sont souvent journalières et parfois hebdomadaires ou mensuelles. L'utilisation des données non journalières rend difficile la détection de l'effet de l'exposition (Peng & Dominici, 2008). Pour mener ce genre d'étude il faut prendre en compte trois aspects : la confondance, les délais et la modélisation.

2.1. Les variables confondantes et les effets modificateurs de la météorologie

Il y a différentes sources de confusion, mesurées et non mesurées (Peng et al., 2006). Nous présentons dans cette section les variables confondantes en lien avec les MCV.

2.1.1. Les facteurs de confusion mesurés

Les facteurs de confusion mesurés les plus importants comprennent : les polluants, la défavorisation, le vieillissement et le genre.

✓ Pollution atmosphérique et MCV

La pollution atmosphérique est considérée comme étant problématique au Québec (Bouchard & Smargiassi, 2009), tout particulièrement en milieu urbain. Certaines études suggèrent même l'existence d'un effet synergique entre les polluants et les températures élevées (Kosatsky et coll., 2007; Kovats, 2008; Luber, 2008). Cet effet synergique influence la mortalité (Besancenot, 2002). En effet, la formation de l'ozone et de particules fines secondaires (les deux principales composantes du smog estival) est plus rapide en présence de températures élevées. Les niveaux de ces polluants risquent donc d'augmenter avec le réchauffement climatique (MDDEP, 2002b). Enfin, l'hiver, le smog est surtout constitué de particules fines générées par les activités anthropiques comme le chauffage au bois (Lebel G, 2012), car les températures froides et le faible rayonnement ultraviolet pendant l'hiver ne favorisent pas la formation d'ozone (MDDEP, 2002b). La pollution atmosphérique constitue un problème de santé publique dans plusieurs régions du monde, mais avec une ampleur qui varie géographiquement. Son association avec la mortalité est d'ailleurs bien documentée : en Amérique du Nord (p. ex. Goldberg et al., 2001; Middleton et al., 2008; Pope III et al., 2004; Pope III et al., 2009; Schwartz et al., 2002), en Europe (p. ex. Breitner et al., 2009; Hoek et al., 2002; Ibal-Mulli et al., 2004; Pelucchi et al., 2009; Pope et al., 2004) et en Asie (p. ex. Chen et al., 2011; Hori. A 2012; Park et al., 2011; Zhang et al., 2012).

Par ailleurs, la chaleur peut accroître l'effet négatif de certains polluants atmosphériques sur la santé cardiovasculaire (Portier, 2010). Il a été démontré qu'une augmentation de particules fines et de l'ozone, qui coïncident souvent avec les vagues de chaleur, peut avoir un effet direct sur le cœur et a été associée aux décès à la suite de maladies cardio-pulmonaires, de maladies circulatoires et de cardiopathie ischémique (Grahame & Schlesinger, 2010; Holguín et al., 2003) (voir Santé Canada, 2011). Ces polluants sont souvent contrôlés lors de l'examen de l'effet de la température ambiante sur la mortalité MCV (Gouveia et al., 2003a; Medina-Ramón & Schwartz, 2007; Rainham & Smoyer-Tomic, 2003).

✓ **Défavorisation et MCV**

Le lien entre les effets sanitaires attribués aux fortes chaleurs estivales et la défavorisation sociale ou matérielle¹ est très bien documenté, notamment dans les revues systématiques de la littérature scientifique (Tairou et al., 2010). Le lien entre défavorisation et froid hivernal est toutefois moins clair (Bustinza et al., 2010). Parmi les raisons pouvant expliquer cela, est le fait qu'au Québec les ressources naturelles permettent d'assurer un chauffage hydroélectrique pour un coût relativement peu élevé et qui s'avère (avant taxes) l'une des plus avantageux par rapport à plusieurs pays industrialisés (Filion & Lauzier, 2002). Il n'en demeure pas moins que les Québécois les moins fortunés cherchent à habiter un logis dont le prix est le plus abordable possible, ce qui peut signifier un logement mal isolé, plus difficile à chauffer et dont les frais de chauffage sont élevés (Bélanger et al., 2006).

Par ailleurs, le risque de décéder de MCV est étroitement lié à la défavorisation (Blais et al., 2012; INSPQ, 2009). La surmortalité des plus défavorisées est de 25 % pour la dimension matérielle et

¹ Au Québec, la défavorisation matérielle et sociale est généralement étudiée à l'aide d'un indice développé par l'INSPQ. Cet indice inclut six indicateurs, soit la proportion de personnes de 15 ans ou plus sans certificat ou diplôme d'études secondaires, le ratio emploi/population chez les 15 ans ou plus et le revenu moyen des personnes de 15 ans ou plus sous l'axe matériel, puis la proportion de personnes de 15 ans ou plus séparées, divorcées ou veuves, la proportion de familles monoparentales et la proportion de gens de 15 ans ou plus vivant seuls dans leur domicile sous l'axe social (INSPQ, 2009).

de 11 % pour la dimension sociale. Cependant, cet écart s'accroît pour atteindre 53 % lorsque l'on considère les personnes les plus favorisées (Q1Q1) et les plus défavorisées (Q5Q5) sur les deux dimensions de la défavorisation. Cela dit, il est également possible que la défavorisation touche davantage certains groupes d'âge, ou un genre particulier, comme suggéré par Reid et al. (2009) notamment, et que l'interaction entre ces facteurs influe différemment sur la relation MCV-météorologie, par rapport à si on les avait considérés indépendamment l'un de l'autre.

✓ **Vieillessement et MCV**

Coupler les CC et le vieillissement de la population, auquel sont confrontés la plupart des pays développés, augmentera le nombre d'individus atteints d'une MCV (IPCDC, ca. 2011). Divers changements physiologiques associés au vieillissement prédisposent les personnes âgées à des troubles de santé liés à la chaleur et notamment chez les personnes atteintes de MCV (Gauthier et al., 2005), dont la réduction de la thermorégulation (Kenney & Hodgson, 1987) et des seuils de sudation plus élevés que chez les plus jeunes (Santé Canada, 2011). Y contribue également le fait que les personnes âgées ressentiraient moins la soif et percevraient moins la chaleur que leurs cadets (Blum et al., 1998). L'effet de la température est plus prononcée chez les personnes âgées touchées par une des MCV (Gouveia et al., 2003b; Wang et al., 2014).

D'après la revue de littérature de Ye (2012), le risque de décès dû au froid l'hiver est également plus élevé chez les personnes âgées que chez les plus jeunes, probablement à cause de la réduction naturelle de leur sensibilité au froid (Carder et al., 2005). D'un autre côté, il est aussi possible que l'impact du vieillissement sur la relation MCV-météorologie soit plus important chez les plus défavorisés que chez les plus favorisés (voir, p. ex., Reid et al., 2009) et que l'interaction entre ces facteurs influe différemment sur la relation, par rapport à si on les avait considérés indépendamment l'un de l'autre.

✓ Genre et MCV

Les effets de la chaleur ou du froid selon le genre sont contradictoires, comme en témoignent les revues systématiques (p. ex., Bustinza et al., 2010; Tairou et al., 2010). À titre d'exemple, la surmortalité due à la chaleur serait masculine aux États-Unis et féminine en Europe (Besancenot, 2002). Malgré ces divergences, cette variable demeure d'intérêt. Cela dit, il est possible que les femmes de 65 ans (comme suggéré dans la revue de la littérature de Lundgren et Jonsson (2012), notamment) aient plus d'impacts sanitaires que les hommes du même âge ou que leurs cadets, et que l'interaction entre l'âge et le genre influe différemment sur la relation MCV-météorologie, par rapport à si on les avait considérés indépendamment l'un de l'autre.

2.1.2. Les facteurs de confusion non mesurés

Les facteurs de confusion non mesurés ce sont les facteurs pour lesquels nous n'avons pas de données. Ils produisent une variation saisonnière et une tendance dans les données de mortalité (Armstrong et al., 2011; Vutcovici et al., 2013). Par exemple, on observe plus de décès en hiver en partie à cause des épidémies (Battaieb et al., 2010; Peng et al., 2006). Pour contrôler ces variations et cette tendance, il faut inclure la variable « temps » dans le modèle. Un modèle sans ce contrôle prévoira moins de décès pour les hivers des climats futurs (Doyon et al., 2006). Cette approche a été adoptée systématiquement dans les études épidémiologiques environnementales sur des décès quotidiens (p. ex. Braga et al., 2001a; Conceicao et al., 2001; Jbilou & El Adlouni, 2012; Schwartz, 1994; Schwartz, 2000a; Wood, 2006).

2.2. Importance des variables à délais

Les effets retardés sont toujours observés entre l'exposition (p. ex. température, humidité, ozone) et les décès (Muggeo & Hajat, 2009). Les effets retardés de la température, sont plus courts en saison chaude qu'en saison froide (Ye et al., 2012). L'effet chaud est restreint à quelques jours avant le décès alors que l'effet froid persiste quelques semaines avant le jour de décès (Anderson

& Bell, 2009; Pattenden et al., 2003). Dans la revue de Gosling (2009), la totalité des décès liés à la chaleur est souvent associée à un effet retardé de moins de trois jours.

Des études ont opté pour des modèles avec des regroupements temporels de l'exposition et ses effets retardés jusqu'à certain délai (p. ex. moyenne mobile de 3 jours) (Hajat et al., 2002; O'Neill et al., 2003). Pattenden (2003) a noté que l'utilisation des moyennes mobiles dans le cas où un effet de déplacement de mortalité est observé, peut surestimer l'effet chaud et sous-estimer l'effet froid sur la mortalité et suggère de prendre des périodes plus longues (p. ex. 3 jours pour l'effet chaud et 21 jours pour l'effet froid).

Une autre alternative est d'utiliser les modèles à effets retardés distribués (DLM) combinés aux techniques de régression, p. ex. régression de poisson (Armstrong, 2006; O'Neill et al., 2003). Il s'agit d'introduire dans le modèle l'exposition avec tous ses effets retardés simultanément pour explorer la structure de ces effets. L'approche la plus flexible dans le cas des effets non linéaires est d'appliquer les modèles à effets retardés distribués non linéaires (Distributed Lag Non Linear Model, DLNM) qui utilisent des polynômes (Schwartz, 2000b) ou des fonctions Splines (Armstrong, 2006).

2.3. Méthodes statistiques utilisées

La plupart des études menées sur la relation mortalité-météorologie suggèrent des relations non linéaires, en forme de U, de V ou de J (Pattenden et al., 2003). Cette relation est traitée selon trois approches. La première est basée sur l'utilisation d'un modèle linéaire généralisé-*splines* naturelles (GLM-NS), la deuxième explore le GAM, alors que la troisième exploite les (GLM ou GAM) combinés au DLNM.

L'approche GLM-NS est une modélisation semi-paramétrique combinant les modèles linéaires généralisés de type GLM (Nelder & Wedderburn, 1972) et des *splines* cubiques naturelles (Dominici et al., 2002; Peng et al., 2006), soit des fonctions polynomiales cubiques par morceaux

avec un nombre de nœuds prédéterminé à des emplacements connus (Hastie & Tibshirani, 1990). Cette démarche a été très utilisée. Elle a été notamment appliquée par Carder et al. (2005) et Crabbe (2012). Les principales limites de cette démarche sont le nombre de nœuds et leurs emplacements qui peut affecter les estimations (p. ex. Wood, 2001).

Introduit par Hastie et Tibshirani (1986), le GAM est une extension du modèle GLM, mais contrairement à ce dernier, le GAM est une modélisation non paramétrique. De ce fait, il dispose de nombreux avantages par rapport au modèle GLM. Le GAM est également un outil d'une grande flexibilité dans la forme des relations et permet ainsi au modèle de s'ajuster à la tendance présente dans les données.

Avec ces deux premières approches, l'exposition est présentée par une moyenne mobile qui résume l'exposition et ses effets retardés (p. ex. moyenne mobile Lag0-3). Ce qui fournit parfois des résultats difficiles à spécifier et à interpréter.

Quant à la troisième approche, les DLNM (Gasparrini et al., 2010) sont utilisés pour identifier les retards d'une exposition (p. ex. température) qui ont un effet significatif sur la variable d'intérêt (ici décès MCV). L'avantage de l'utilisation des DLNM c'est qu'ils sont capables d'identifier la répartition de la mortalité à la fois sur une période de temps (ordre du retard) ainsi qu'à l'échelle de l'exposition, au lieu d'un seul retard mixte (p. ex. moyenne mobile Lag0-3) (Barnett et al., 2012; Huang et al., 2012; Lavigne et al., 2014).

3. Problématique

La problématique du présent projet est scindée en trois grandes parties : les aspects de la relation météo/décès, la stratification et les données sanitaires agrégées.

3.1. Les aspects de la relation météo/décès

Le long de la revue de littérature, les trois aspects (la confondance, les délais et la modélisation) étaient abordés différemment d'une étude à l'autre. Il n'y a pas de procédure globale qui permet de traiter d'une façon claire ces aspects.

Au niveau de la confondance, on peut noter que la plus part des études sont restreintes aux polluants, genre et l'âge. Cependant, ils existent d'autres facteurs qui peuvent perturber l'association entre la météo et les décès, comme la densité de population et l'indice de défavorisation (ID) qui sont rarement considérés.

En ce qui concerne les délais (lags), il n'y a pas une valeur fixe ou une seule méthode pour déterminer combien de jours de retards de l'effet de l'exposition il faut introduire dans le modèle (Carder et al., 2005; Muggeo & Hajat, 2009). Identifier le décalage approprié de l'exposition à inclure dans le modèle est un problème qui nécessite des connaissances du domaine : type de maladie (Ballester et al., 1997), zones géographiques (Conti et al., 2005; Pattenden et al., 2003) et le type de l'exposition.

D'autre part, le choix de la méthode de modélisation n'est également pas trivial et parfois subjectif. Cependant peu d'études ont analysé la sensibilité de leurs modèles. Celles qui l'ont fait, elles ont seulement évalué l'effet du changement des degrés de liberté associés à la fonction de lissage de temps, utilisée pour contrôler la tendance et la saisonnalité, sur l'estimation de l'effet des autres variables introduites dans le modèle (Armstrong et al., 2011; Goldberg et al., 2011; Liu et al., 2011).

3.2. La stratification

Selon la littérature, les études de la relation météorologie décès utilisent des données sanitaires agrégées tel que le nombre de décès par jour. Ainsi, des modèles de régression de Poisson sont utilisés. Or cette façon de procéder ne permet pas de prendre en compte les caractéristiques individuelles et contextuelles, comme le genre et l'ID, pour les introduire dans le modèle. Pour tenir compte de ces effets certaines études ont recouru à la stratification de la population selon chacune des modalités des variables individuelles et/ou contextuelles. Giroux (2011a), par exemple, a appliqué des modèles additifs généralisés (GAM, generalized additive model) à chaque modalité des variables catégorielles potentiellement confondantes considérées (p. ex. au groupe de femmes âgées de 65 ans et plus). Cela a généré un nombre considérable de résultats, un modèle par modalité et une fonction de l'effet pour chaque modalité d'une variable confondante donnée), a compliqué l'interprétation et la discussion des résultats, sans compter l'absence de mesure d'effet permettant d'évaluer la contribution des différences observées pour chacune de ces variables.

3.3. Données sanitaires agrégées

Les données sanitaires, comme variable d'intérêt sont agrégées au niveau des sujets (nombre de décès), provoquant ainsi une perte d'informations individuelles et contextuelles. Ce qui génère l'absence de mesure d'effet permettant d'évaluer la contribution pour chacune des variables individuelles et contextuelles. L'état de l'art de ce rapport montre l'importance de prendre en considération ces effets. En effet la distribution des MCV est inégale au niveau contextuel (p. ex. aire de diffusion (AD) défavorisée versus AD favorisée) et individuel (p.ex. personnes âgées). Aussi, l'influence du contexte ne doit pas être considérée comme opérant de façon identique sur tous les individus. Cette posture proche du déterminisme néglige les adaptations individuelles et la façon dont les habitants vivent, pratiquent et s'approprient l'espace (Vallée, 2009). Il s'agit de voir

dans quelle mesure les disparités spatiales sont imputables aux différents profils sociodémographiques (l'effet de composition) et aux différentes caractéristiques de leur lieu de résidence (l'effet du contexte) (Chaix & Chauvin, 2002). D'où la nécessité au recours à des données non agrégées, décès MCV (présence ou absence de l'évènement) à la place du nombre de décès, exploitées par des méthodes susceptibles de cerner les relations possibles entre l'individu et son contexte. Aussi, capables de tenir compte des niveaux d'agrégation emboîtés (individus, AD, régions socio-sanitaires (RSS), etc.) ainsi que de la structure hiérarchisée de l'information, c'est-à-dire de données qui proviennent de plusieurs unités d'analyse (Herjean, 2006). D'où l'importance de disposer d'outils permettant de saisir cette réalité complexe (Gauvin & Dassa, 2004).

Il est admis que l'âge joue un rôle majeur dans le risque de survenue et l'évolution des maladies, ainsi que dans l'évolution de la plupart des phénomènes socio-économiques. Cependant, la période d'observation traduit tous les évènements qui sont susceptibles d'affecter simultanément et de la même manière, toutes les cohortes quel que soit l'âge des individus de ces cohortes. Les changements survenant dans les méthodes de diagnostic, la mise en place ou l'évolution des plans de dépistage, de même que l'évolution de phénomènes naturels affectant simultanément tous les individus des cohortes quel que soit leur âge, peuvent avoir un impact sur la mesure de l'évènement, dès lors qu'ils en modifient le niveau (Sala, 2009). Cet impact se traduit alors par un effet période. En effet la variable cohorte représente généralement la cohorte de naissance. Dans l'idée que tous les individus d'une même cohorte sont exposés de la même manière au cours du temps, mais que toutes les cohortes ne sont pas exposées à un même niveau de risque. La cohorte traduit ainsi tous les évènements qui affectent de la même manière tous les individus d'une même cohorte, indépendamment de l'âge des individus et de la période d'observation (p.ex. évolution du cancer à travers les générations). D'où l'intérêt de faire appel à des outils permettant de cerner ces trois effets à la fois (âge, période et cohorte) (Carstensen, 2007).

4. Pertinence du projet

La pertinence du projet est présentée sous trois aspects : son but, sa contribution à la santé publique et ses objectifs.

4.1. But du projet

Le but ultime du projet est de soutenir la surveillance en santé publique, en améliorant les connaissances épidémiologiques et méthodologiques sur la mortalité pour cause de MCV dans un contexte de CC. Les outils développés dans le cadre du présent projet pourront servir à des fins de prévention pour les MCV et pour d'autres classes de diagnostics, ainsi qu'à des fins de surveillance des maladies chroniques. De façon générale, ils seront à même d'éclairer davantage les effets de la météorologie/climat sur la santé humaine et déterminer l'impact réel des variables explicatives sur les décès pour causes de MCV. Ils serviront surtout à soutenir les efforts des pouvoirs publics pour mettre en place des stratégies et plans d'actions pouvant contribuer à atténuer ces effets en adoptant des mesures préventives ou/et en instaurant des systèmes d'alerte et de surveillance. Les effets des CC sur la santé humaine peuvent être réduits en adoptant des mesures préventives, en améliorant les connaissances et en instaurant des systèmes d'alerte et de surveillance. Ainsi, un nombre important de ces décès pourraient être évités avec un avertissement adéquat et une réponse appropriée, notamment sur le plan clinique auprès des personnes atteintes de problèmes cardiovasculaires.

4.2. Contribution à l'action 21-volet santé et à la santé publique

Le projet fait partie du Programme de recherche en santé et changements climatiques 2011-2015, Études appliquées à la relation MCV-météorologie/climat au Québec – programme qui s'inscrit dans le Plan d'action 2006-2012 sur les changements climatiques (PACC), « Le Québec et les changements climatiques, un défi pour l'avenir », et plus précisément à l'Action 21.

Ces projets contribuent également au Programme de santé publique 2003-2012 élaboré par le gouvernement du Québec, dont l'un des objectifs est de réduire l'incidence, la mortalité et la morbidité dues à certaines maladies chroniques, incluant les MCV.

Enfin, ce projet s'enregistre au sein du Plan commun de surveillance (PCS, volet changements climatiques) et du Plan ministériel de surveillance multithématique (PMSM), sous l'aspect du développement d'indicateurs de surveillance applicables au domaine de la santé et de l'environnement (un avis favorable du Comité d'éthique de la santé publique dans le cadre de son examen du PCS et du PMSM en 2010).

4.3. Objectifs du projet

Le projet de thèse est divisé en trois parties :

4.3.1. Partie A : Procédure méthodologique

Cette partie vise à établir une procédure méthodologique globale qui prend en compte l'ensemble des aspects de la relation météo-santé (confondance, délais, modélisation), qui peut être appliquée aux MCV ou autres maladies (correspond à la problématique 3.1). De plus, une analyse de sensibilité globale du modèle qui permet de déterminer quelles sont les variables d'entrée du modèle qui contribuent le plus dans l'explication des décès, quelles sont celles qui n'ont pas d'influence et quelles sont celles qui interagissent au sein du modèle. Une autre application de cette partie, concerne les projections des variations futures de décès sur les horizons 2030, 2050 ou 2080.

4.3.2. Partie B : GEE-Modèle multiniveau

L'objectif de la partie B est de répondre à la problématique 3.2, il s'agit de recourir à des approches permettant de réduire le nombre de modèles puis proposer un pour l'ensemble de strates au lieu d'un modèle par strate (une seule écriture de modèle). Des approches qui tiennent compte à la fois des variables individuelles et contextuelles afin de mieux synthétiser les résultats, d'avoir

une vue d'ensemble, sans négliger ni les particularités de chacun des sous-groupes (ou sous-populations), ni la validité des estimations produites, ni la présentation de résultats éclairants et faciles à opérationnaliser par les décideurs de santé publique. Pour atteindre cet objectif, cette partie mettra en œuvre les modèles marginaux de type GEE (*Generalized Estimating Equations*) et les modèles multiniveaux.

4.3.3. Partie C : Données sanitaires agrégées

Pour résoudre la problématique 3.3, cette partie fait appel à des modèles utilisant des données sanitaires non agrégées (p. ex. décès MCV) et introduisant les variables individuelles et contextuelles telles quelles. Elle est scindée en deux sous-parties C1 et C2 :

✓ Sous-partie C1 : Modèle de Cox multiniveau

La sous-partie C1 rend possible l'inclusion des données individuelles et contextuelles utilisées par des analyses contextuelles. Ces dernières recourent à des facteurs contextuels (p. ex. ID), au-delà des facteurs individuels, et parviennent à une meilleure identification des populations à risque. D'où l'importance de disposer d'outils permettant de saisir cette réalité complexe. Le modèle de Cox multiniveau est privilégié pour étudier les effets du contexte car il distingue les caractéristiques individuelles des caractéristiques contextuelles et tient compte des niveaux d'agrégation emboîtés (individus, AD, RSS).

✓ Sous-partie C2 : Modèle Age-Période-Cohorte (APC), diagramme de Lexis

Tel qu'expliquer dans la section problématique, pour l'étude de la relation météo-décès, il est nécessaire d'appréhender les effets sous-jacents (âge, période et cohorte) aux variables environnementales. Le modèle Age-Période-Cohorte (APC) permet de décortiquer l'effet de la météo/polluants de ces trois effets. Ce modèle interprète le diagramme de Lexis. Ce dernier représente les événements d'intérêt (p. ex., décès par MCV) en fonction de la date calendrier (période), de l'âge ou de la date de naissance (cohorte) et exprime les taux de décès MCV comme un produit des effets de l'âge, et/ou de la période et/ou de la cohorte. Ainsi que exposer l'évolution

des taux de mortalité à travers ces trois axes de temps (voir Figure 1 à titre illustratif, (Carstensen, 2007)). Cette analyse descriptive est nécessaire avant d'effectuer la modélisation.

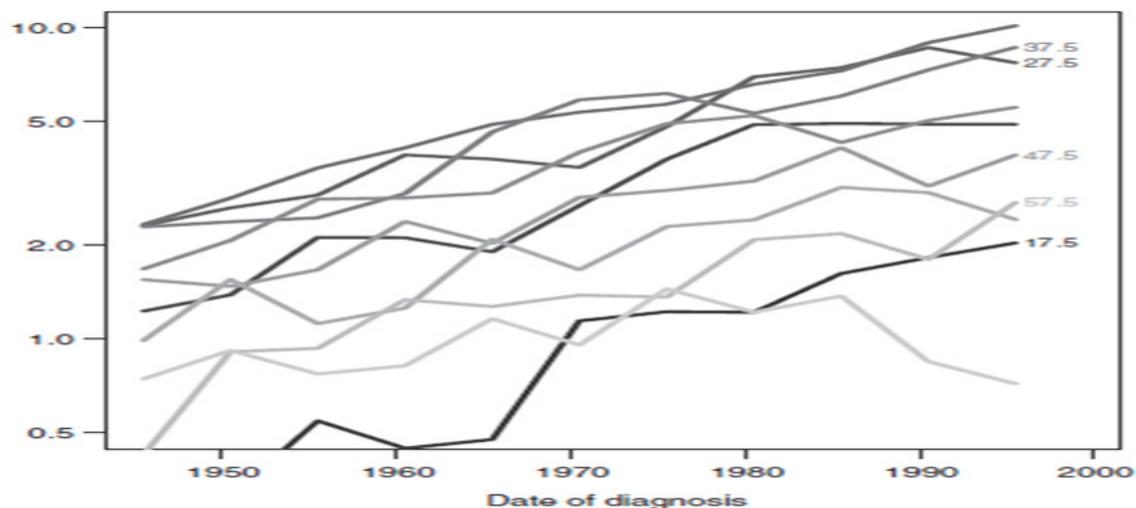


Figure 1 : L'évolution des taux d'incidence du cancer en fonction de la période et l'âge

Les méthodes présentées auparavant utilisent les données des personnes décédées par MCV au Québec de 1981 à 2011, à l'échelle de la province, également les RSS ou communautés métropolitaines (CM) correspondantes de Montréal, de Québec et de Saguenay (selon les distributions observées), en raison des importantes différences régionales soulevées dans le domaine de la santé et du climat (Doyon et al., 2008), notamment pour les MCV (Filate, 2006).

5. Méthodologie

Cette section présente les données et les méthodes utilisées dans ce projet de doctorat.

5.1. Variables étudiées

Quatre catégories de variables sont considérées dans ce projet : les variables sanitaires, les variables individuelles, les variables environnementales et les variables contextuelles par région RSS ou CM selon la méthode. Ces variables sont définies dans les paragraphes suivants, et leur choix y est justifié.

5.1.1. Les variables sanitaires et individuelles

Les variables sanitaires proviennent de l'INSPQ, dans le cadre de l'entente du PMSM. Elles regroupent les décès par MCV (CIM-9 ou CIM-10, transition à partir de 2000), des caractéristiques sociodémographiques, ainsi que des caractéristiques liées aux soins et services.

Dans le présent projet, un décès par MCV se définit comme étant un décès dont la cause principale est l'une des maladies suivantes : une maladie coronarienne (CIM-9 : 410-414; CIM-10 : I20-I25), une insuffisance cardiaque (CIM-9 : 428; CIM-10 : I50), une maladie vasculaire cérébrale ou une ischémie cérébrale transitoire (CIM-9 : 362.3, 430-438; CIM-10 : G45, H34.1, I60-I69). Les causes secondaires de décès ne sont pas considérées, car elles ne sont colligées que depuis 2000.

Parmi les autres variables, sont retenus aux fins du présent projet : la date du décès et le code postal de résidence afin de lier les données sanitaires, les données individuelles (âge au décès et le genre du décédé) aux autres données, dans l'espace et le temps (Tableau 1).

L'âge des personnes décédées sera regroupé sous deux larges strates d'âge pour la partie A. Le seuil à retenir reste à déterminer, car dans la littérature scientifique chaque strate d'âge de 5 ans à partir de 50 ans jusqu'à 85 ans a été étudiée pour la mortalité toutes causes (comme illustré dans les revues systématiques de la littérature scientifique, dont celles de (Bustanza et al., 2010; Tairou et al., 2010). Toutefois, il semble que le taux annuel moyen de décès (par million de populations) selon les groupes d'âge augmenterait nettement chez les 65 ans ou plus, comparativement à leurs cadets, et que la tendance à la hausse s'accroîtrait particulièrement lors du passage de la tranche d'âge des 65-74 ans à celle des 75-84 ans (CDC, 2002; OMS, 2009b). Ces résultats ont été observés chez les Américains de 1979 à 1997, mais leur généralisation à d'autres populations n'avait pas encore été examinée en 2009 (OMS, 2009a). Puisque le projet porte sur la mortalité par MCV, non sur la mortalité toutes causes, et que la population québécoise est nettement moins nombreuse que la population américaine, le choix des groupes d'âge étudiés dans le présent projet dépendra des distributions observées lors de l'analyse préliminaire.

Tableau 1 Périodes couvertes par les variables sanitaires et individuelles

Variables	Périodes
Variable sanitaire dépendante :	
décès MCV (non agrégée)	1981-2011
Nombre de décès (agrégée)	1981-2011
Variables individuelles confondantes :	
• Âge au moment du décès	1981-2011
• Genre du décédé	1981-2011
Variables pour relier les données entre elles :	
• Date du décès (jour-mois-année)	1981-2011
• Code postal de résidence à 6 positions	1981-2011

5.1.2. Variables environnementales

Les variables environnementales se divisent en trois types : les variables météorologiques (données historiques), les polluants atmosphériques (données historiques) et les données climatiques (simulations futures).

✓ Variables météorologiques

Les variables météorologiques ou variables d'exposition sont définies au Tableau 2. Ces données sont issues du portail de téléchargement DAI d'Environnement Canada (*Données Accès et Intégration*, cf. <http://loki.qc.ec.gc.ca/DAI/>); elles seront complétées par celles du Ministère du Développement Durable, de l'Environnement, de la Faune et des Parcs (MDDEFP), si besoin est. Différents indicateurs de la température ont été utilisés, la température minimale (Ebi et al., 2004), maximale (Gosling et al., 2007; Wang et al., 2009). moyenne (Battaieb et al.; Bayentin et al., 2010; Ha & Kim, 2013; Hori. A 2012; Ishigami et al., 2008; Kovats et al., 2004; Lavigne et al., 2014; Liang et al., 2008; Schwartz et al., 2004; Tian et al., 2012; Wang et al., 2014), température diurne (Tam et al., 2009). La température diurne peut être aussi un facteur de risque pour les

maladies cardiovasculaires et respiratoires (Liang et al., 2009). Quelques études ont utilisé des indices biométéorologiques comme la température apparente (Basu & Ostro, 2008; Bell et al., 2008; Madrigano et al., 2013; Michelozzi et al., 2009; O'Neill et al., 2003) et l'humidex (Mastrangelo et al., 2007). Ces indices combinent les mesures de la température et l'humidité, sont considérés les meilleures mesures de l'effet combiné sur la santé (Ye et al., 2012) afin d'éviter la colinéarité. Il n'y a ni critère de sélection du meilleur indicateur (Barnett et al., 2010). Gouveia (2003b) compare la vraisemblance du modèle pour la température minimale, maximale et moyenne, puis suggère d'utiliser la température moyenne car elle était le meilleur prédicteur du froid et du chaud.

Sur la base de cette littérature scientifique, la principale variable météorologique d'intérêt est la température quotidienne. La température moyenne sur 24 heures (un bon indicateur de l'exposition à la chaleur selon Besancenot (1997) permet de comprendre les impacts de la chaleur ou du froid sur le corps humain et fournit des résultats plus faciles à interpréter (Anderson & Bell, 2009). Facilement compréhensible, cet indicateur ne représente pas, toutefois, l'ensemble des efforts imposés à l'organisme et le risque d'accident pathologique qui en résulte (Ledrans, 2003). En effet, une seule variable météorologique (température) ne peut être représentative de l'impact global de la météo sur la santé car d'autres variables météorologiques peuvent l'affecter (Kalkstein, 1991) cité dans (Gosling et al., 2009). Ces variables météorologiques doivent être prises en considération en parallèle, d'où leur sélection afin d'étoffer le contexte lié à la température (Robinson, 2000).

Toutes les variables météorologiques sont des données continues. Dans le présent projet, elles seront mesurées à l'aide de moyennes spatiales des stations (MSDS). Pour la même base de données, selon Giroux et al. (2013), l'utilisation d'autres approches (comme le krigeage) n'améliorerait pas les modèles de régression de façon systématique et suffisante.

Tableau 2 Données météorologiques historiques du portail DAI

Données	Code / Type	Unités
Température maximale	001 / Journalière	Degrés Celsius (°C)
Température minimale	002 / Journalière	Degrés Celsius (°C)
Température moyenne	003 / Journalière	Degrés Celsius (°C)
Précipitations totales	012 / Journalière	Millimètres (mm)
Épaisseur de neige au sol	013 / Journalière	Centimètres (cm)
Pression au niveau de la mer	073 / Horaire	Kilopascals (kPa)
Pression à la station	077 / Horaire	Kilopascals (kPa)
Humidité relative	080 / Horaire	Pour cent (%)

Si l'analyse avec les variables de base (températures, précipitations totales, épaisseur de neige au sol, pression et humidité) ne donne pas de résultats satisfaisants, la construction et l'inclusion d'indices climatiques seront envisagées (Rainham & Smoyer-Tomic, 2003), comme l'indice humidex. Ce dernier est calculé à partir de la température et de l'humidité. La variation spatiale de l'humidex est relativement faible, une moyenne sur la région est considérée représentative de la région entière (Zauli Sajani et al., 2002). Les effets retardés (ou *lags* en anglais) des variables météorologiques seront mesurés jusqu'à 14 jours (p. ex. Braga et al., 2002; Carder et al., 2005; Muggeo, 2008; Schwartz et al., 2004). Enfin, on tiendra compte de la saison sous forme d'une variable binaire, telle que définie par Environnement Canada, soit de mai à octobre pour l'été et de novembre à avril pour l'hiver.

L'ensemble de ces données couvre la période 1981-2011 (Tableau 1).

✓ **Polluants atmosphériques**

Comme rapporté dans le dernier bilan de la qualité de l'air en lien avec la santé, produit par l'INSPQ (Lebel et al., 2012), au Québec, les principaux contaminants atmosphériques réglementés et mesurés à partir de 1975 sont les particules fines de moins de 2,5 microns (PM_{2,5}), les oxydes

d'azote (NO et NO₂), l'ozone (O₃), le dioxyde de soufre (SO₂) et le monoxyde de carbone (CO). Dans le présent projet, les NO_x, le SO₂ et le CO ne sont toutefois pas considérés. Ces polluants accusent une grande déficience en termes de couverture spatiale et de monitoring environnemental, p. ex., les NO_x sont mesurés essentiellement sur l'île de Montréal (Lebel et al., 2012). De même, leurs valeurs dépassent rarement les normes à observer (p. ex., pour les NO₂, une seule journée avec au moins un dépassement de la norme horaire de 220 ppb a été observée à Montréal, de 1995 à 2009. Par ailleurs, les SO₂ et CO sont davantage des indicateurs de sources d'émission de polluants que de surveillance de l'exposition sur les individus. Enfin, les impacts sanitaires associés à tous ces polluants touchent essentiellement le système respiratoire. (Abbey et al., 1999).

Ainsi, seuls les PM_{2,5} et l'O₃ sont étudiés à titre de facteurs aggravants ou de variables potentiellement confondantes. Les PM_{2,5} correspondent à un indicateur majeur de risque sanitaire de la pollution de l'air (Pascal et al., 2012; SFSP, 1996). Quant à l'O₃, même si les mécanismes sous-jacents aux impacts sanitaires qui lui sont associés ne sont pas encore bien connus, un panel d'experts de National Academy of Sciences des États-Unis a malgré tout conclu que l'association entre les changements journaliers des concentrations d'O₃ et les décès au cours des mois d'été avait un lien de cause à effet (National Research Council, 2008). Une situation qui pourrait s'aggraver dans un contexte de CC, puisque les températures élevées pourraient amplifier les effets de l'ozone sur la santé (Künzli, 2010).

Les données d'O₃ et de PM_{2,5} sont mesurées à l'aide d'analyseurs en continu dans 57 des 75 stations que compte le réseau. Elles proviennent de la banque du Réseau National de Surveillance de la Pollution Atmosphérique (RNSPA) d'Environnement Canada, base qui est alimentée par le MDDEFP et la Ville de Montréal. Ces données seront accessibles par l'entremise de l'INSPQ. Les données d'O₃ couvrent toute la période étudiée (Tableau 1). Pour les PM_{2,5}, la période débute plus tard, car ce n'est qu'à partir de 1999 que la moyenne annuelle commence à

avoir une meilleure représentativité pour le Québec, et seulement à partir de 2002 qu'elles sont mesurées au Saguenay-Lac-Saint-Jean (Lebel et al., 2012). Le nombre de stations mesurant les $PM_{2,5}$ varie de 17 à 47 selon l'année pour le Québec, de 7 à 10 pour la RSS de Montréal, de 1 à 6 pour la RSS de la Capitale-Nationale, de 1 à 2 pour la RSS du Saguenay-Lac-Saint-Jean. Pour l' O_3 , ces nombres sont respectivement de 40 à 48, 10 à 11, 6 à 8, et 2. En raison de ces différences, le nombre de stations par jour et par RSS ayant observé un dépassement de la ligne directrice de l'OMS de $10 \mu g/m^3$ de particules fines ou de la norme horaire de 82 ppb du Règlement sur l'assainissement de l'atmosphère (RAA) pour l'ozone sera considéré. Avant de statuer sur la forme finale de la mesure des polluants atmosphériques, quelques scénarios seront explorés (Tableau 3). En particulier pour l'ozone, ce polluant peut voyager loin et les périodes où ses concentrations sont élevées peuvent persister de quelques heures à quelques jours, selon les conditions météorologiques (MDDEP, 2002a). Enfin, les effets retardés pour les polluants atmosphériques seront évalués jusqu'à 3 jours (p. ex. Braga et al., 2001b; Schwartz, 2000b).

Tableau 3 Mesures des polluants atmosphériques

Particules fines ($PM_{2,5}$)	Ozone (O_3)
<ul style="list-style-type: none"> • Variable binaire indiquant le dépassement (ou non) de la ligne directrice de l'OMS; • Concentration moyenne sur 24 h, comme dans Bouchard et Smargiassi (2007) ou dans Pascal et coll. (2009); • Concentration moyenne sur les 3 derniers jours, comme dans Huang et coll. (2005). 	<ul style="list-style-type: none"> • Variable binaire indiquant le dépassement (ou non) de la norme horaire du RAA; • Concentration horaire maximale a plutôt été retenue, comme Bouchard et Smargiassi (2007); • 8h de concentration maximale comme dans Pascal et coll. (2009); • Concentration moyenne sur les 3 derniers jours, comme dans Huang et coll. (2005).

Une attention particulière devra être portée aux données manquantes pour les $PM_{2,5}$ (colligées depuis 1999 ou 2002, selon la région). Si les $PM_{2,5}$ contribuent à la relation entre les MCV et les données météorologiques, alors la période étudiée sera tronquée à 2002-2011.

✓ **Variables météorologiques simulées pour le climat futur**

Les variables météorologiques simulées pour le climat futur de la partie A du projet seront fournies par Ouranos à partir de plusieurs modèles régionaux de climat (MRC). Les MRC sont des modèles climatiques permettant de simuler la réponse du système planétaire à la variation de la concentration des gaz à effet de serre (ou GES) dans le futur. Les MRC ont une résolution spatiale d'environ 50 km, mais ne fournissent des résultats que pour un sous-domaine de la planète. En comparaison des modèles de circulation générale (MCG; ces modèles couvrent l'ensemble du globe, mais à une résolution plus faible : 150 à 350 km), les MRC permettent une meilleure représentation de la surface et de certains phénomènes physiques rares ou extrêmes sur une région spécifique, tout en restant cohérent avec le reste du climat planétaire (Lepage et al., 2011). Ces données simulées de MRC seront incorporées aux modèles retenues à cette partie, afin d'établir les projections futures.

5.1.3. Variables Contextuelles

La densité de la population est un facteur influençant la relation température mortalité (Johnson 2009) souvent omis. Elle est élevée dans les zones urbaines par rapport aux zones rurales. Les zones urbaines contiennent souvent des immeubles de grande hauteur, ce qui limite les espaces verts et la circulation de l'air. Elle sera définie pour chaque AD par le nombre d'habitants au km^2 . L'ID aussi un facteur de confusion rarement utilisé. Ce dernier sera défini pour chaque AD. Quant aux données manquantes d'ID (disponible depuis 1989), l'utilisation d'une variable indicatrice permettra de conserver l'entièreté de la période 1981-2011. Le nombre de catégories

sera vraisemblablement réduit, par exemple, en considérant les classes défavorisées et très défavorisées comme groupe à risque, comparativement aux autres classes (voir Figure 2).

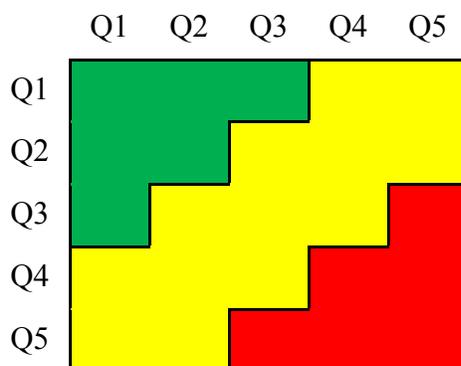


Figure 2 : Combinaison de ID matériel et social (Vert = groupe de personnes favorisées ou très favorisées; jaune = groupe de personnes plus ou moins favorisées (ou défavorisées); rouge = groupe de personnes défavorisées ou très défavorisées)

5.2. Constitution des banques de données

Les données sanitaires, les données individuelles, les données environnementales et les données contextuelles seront jumelées pour créer de nouvelles bases de données qui serviront aux analyses statistiques envisagées. Selon le cas, il pourra s'agir de données sanitaires agrégées (parties A et B) ou de données sanitaires non agrégées (partie C).

Tableau 4 Association entre les différents types de variables et les méthodes

Parties du projet	Zone	Variables			
		Individuelles	Sanitaires	Environnementales	Contextuelles
A : Procédure méthodologique	CM	Âge, genre pour stratifier	nombre de décès MCV par jour	Météo + Polluants	ID pour stratifier
B : GEE-Modèle multiniveau	RSS, AD	Âge, genre pour stratifier	nombre de décès MCV par jour et par strates	Météo + Polluants	ID pour stratifier
C1 : Cox Multiniveau	RSS, AD	Âge, genre introduits dans le modèle	Évènement sanitaire : Décès MCV	Météo + Polluants	ID + densité de population introduits dans le modèle
C2 : Modèle APC, diagramme de Lexis	RSS	Âge introduit dans le modèle	Évènement sanitaire : Décès MCV	Météo + Polluants	N/A

5.3. Méthodologies statistiques

Cette sous-section présente les méthodologies statistiques appliquées dans ce projet qui s’articule autour de trois parties principales A, B et C (qui est découpée en C1 et C2). La présentation des méthodologies est brève et l’accent est mis sur leurs applications. Pour leurs détails techniques, se référer à l’annexe.

5.3.1. Partie A : Procédure méthodologique

La procédure méthodologique qui sera conçue à cette partie, sera appliquée pour modéliser la relation entre les décès (données agrégées) pour cause de MCV et les variables météorologiques, de 1981 à 2011. Les modèles obtenus seront projetés sur les horizons 2030, 2050 ou 2080, grâce aux données de simulations futures fournies par le consortium Ouranos.

Une première série d’analyses, basée sur la régression de Poisson et les DLNM, nous permettra d’examiner l’association entre le nombre de décès par unité de temps (jour) et les variables météorologiques, les polluants atmosphériques, tout en contrôlant pour les tendances à long terme et les tendances saisonnières.

Les DLNM effectuent une analyse complexe, leur fonctionnalité est basée sur le calcul des effets non linéaires d'un facteur à travers les valeurs de ce facteur et à un ordre déterminé de ses effets retardés (p. ex. température et ses effets retardés jusqu'aux 14 jours). Cette fonctionnalité s'effectue en trois étapes :

1. Un ensemble de fonctions de base (p. ex. Spline) peut être défini pour la relation entre l'exposition et la variable d'intérêt. L'exposition est transformée par l'ensemble de ces fonctions de base pour créer une nouvelle matrice des bases de l'exposition. Ces transformations peuvent être plus ou moins complexes, selon les fonctions choisies. A ce moment, la matrice des bases de l'exposition décrit simplement l'effet de l'exposition;
2. En appliquant cette fois-ci un ensemble de fonctions de base (p. ex. Spline) décrivant les effets de retard de l'exposition, une nouvelle matrice des bases des effets retardés est construite;
3. En combinant les deux jeux de fonctions de base on obtient un ensemble de fonctions de base croisée, qui peut être utilisé pour décrire le comportement de l'exposition dans les trois dimensions: l'exposition, l'effet (risque relatif) et à l'échelle du décalage.

Dans les modèles de Poisson, on utilisera les DLNM que pour la température car elle est la variable d'exposition qui a des longs effets retardés. L'effet des autres variables météorologiques sera introduit dans le modèle, selon la forme de la relation avec les décès MCV qui va être déterminée à l'étape de l'analyse exploratoire. Comme mentionné par plusieurs auteurs (Ishigami et al., 2008; Madrigano et al., 2013; Michelozzi et al., 2009; Vutcovici et al., 2013), les polluants ont un effet linéaire monotone sur la mortalité (p.ex. un taux élevé de pollution est associé à un taux élevé de mortalité). De plus, puisque les polluants peuvent agir comme modificateurs d'effet des variables météorologiques sur les décès MCV, des termes d'interaction d'ordre deux entre une variable météorologique et un polluant seront testés, comme dans Ren et coll. (2011). Aucun

terme d'interaction entre deux variables potentiellement confondantes ne sera toutefois considéré, car seule la mesure de la variabilité liée à ces variables est à documenter (Kleinbaum et al., 2007).

La température a des valeurs optimales correspondant aux points minima de la courbe de mortalité (Curriero et al., 2002). Le seuil thermique est le point où la mortalité atteint son minimum, et la mortalité augmente si la température dépasse ou diminue par rapport à ce point (Kalkstein & Davis, 1989).

Parfois, le seuil thermique n'est pas unique. C'est le cas de la forme U de la courbe d'association entre la température et la mortalité, où la température peut être divisée en trois intervalles avec un seuil pour le chaud et un seuil pour le froid et la section entre les deux a une pente égale à zéro. La forme « V » de l'association est un cas particulier quand le seuil chaud et le seuil froid sont égaux (Armstrong, 2006). L'effet froid (inférieur au seuil thermique) et chaud (supérieur au seuil thermique) sont souvent étudiés séparément à cause de la relation non linéaire température – mortalité. Le seuil thermique varie selon la population, la zone et la cause du décès. Il simplifie l'estimation des effets de la température sur les décès MCV. Ainsi les décès attribuables au chaud et au froid seront estimés en utilisant la méthode d'Armstrong (2011) :

$$AF = \frac{\sum_i p_i (RR_i - 1)}{\sum_i p_i (RR_i - 1) + 1}$$

où

AF est la fraction attribuable dans la population indiquant le nombre (ou proportion) des cas qui n'auraient pas eu lieu dans une population si le facteur a été éliminé (p. ex. combien de vies seraient sauvées si les gens ne fumaient pas ?);

i est l'indice pour un niveau de température ;

p_i est la proportion des jours à chaque niveau de température ;

RR_i est le risque relatif pour le niveau i . Le nombre journalier attribué au décès MCV est calculé en multipliant AF par le nombre total de décès.

Enfin, une fois les modèles déterminés, une analyse de sensibilité de ces modèles (avant et après l'ajout des termes d'interaction, le cas échéant) sera effectuée. Cette analyse permettra de déterminer les variables d'entrée du modèle qui contribuent le plus à l'explication des décès MCV, celles qui n'ont pas d'influence, celles qui interagissent au sein du modèle, puis de hiérarchiser les variables d'entrée les plus influentes (Faivre, 2013,; Fassò & Perri, 2006). Le AIC (Akaike Information Criterion) sera utilisé pour comparer les différents modèles de régression. Le modèle avec le plus petit AIC sera le modèle retenu pour l'interprétation et les projections futures. Si le test d'égalité entre la moyenne et la variance est significatif, les modèles de Poisson seront remplacés par les modèles Quasi-Poisson. Dans ce cas, le QAIC qui sera utilisé pour la comparaison, le modèle retenu est celui qui minimise le QAIC donné par (Gasparrini et al., 2010) :

$$QAIC = -2\ell(\hat{\theta}) + 2\hat{\phi}k$$

où,

ℓ est la log-vraisemblance du modèle avec les paramètres estimés par $\hat{\theta}$;

$\hat{\phi}$ est l'estimation du paramètre de dispersion;

k est le nombre de paramètres estimés.

5.3.2. Partie B : GEE-Modèle multiniveau

Tel que décrit dans les sections précédentes, la partie A du projet se limite à modéliser la relation entre les décès MCV et les variables météorologiques, tout en évaluant l'influence des polluants atmosphériques. Les modèles qui y sont proposés ne tiennent pas compte des facteurs contextuels (p. ex. ID) et individuels (p. ex. âge) qui peuvent influencer cette association. Comme complément

aux modèles précédents, nous proposons d'explorer les structures hiérarchiques des données et d'utiliser les méthodes statistiques appropriées, dont les approches GEE (Liang & Zeger, 1986) et les modélisations de type multiniveau (Goldstein, 2003; Rabe-Hesketh & Skrondal, 2012).

Rappelons que nous disposons de quatre catégories de variables : les variables sanitaires, les variables individuelles, les variables environnementales et les variables démographiques par RSS. Ces variables possèdent diverses structures hiérarchiques qui seront mises à profit selon la question de recherche posée.

Les modèles multivariés présentés dans le cadre de la partie A du projet seront améliorés en leur adjoignant des facteurs individuels et contextuels. En ce qui concerne l'association entre les taux de mortalité par MCV et la météorologie, les caractéristiques des individus (p. ex., âge) ne pourront pas être utilisées telles quelles. À l'instar de l'étude de Demers et Karouz (2011), nous construirons des groupes (strates) composés des combinaisons de genre (2 classes), de classes d'âge (2 à 3, selon les données observées) et de l'ID (4 classes). Ces strates, qui sont fixes dans le temps, seront jumelées aux variables météorologiques et aux polluants atmosphériques qui varient géographiquement et temporellement. Nous aurons, ainsi une structure hiérarchique dans laquelle ces groupes seront nichés dans les AD, les AD nichés dans les RSS, pour lesquelles nous disposons de variables qui changent dans le temps : mortalité par MCV, météorologie et polluants atmosphériques.

Une première amélioration des modèles discutés dans le cadre de la partie A du projet, consisterait à construire des modèles marginaux de type GEE dans lesquels on tient compte des corrélations entre les observations de strates similaires. Plus précisément, on pourrait supposer que les observations d'une année à l'autre sont indépendantes, alors que les observations d'une même saison sont corrélées. Similairement, les individus habitant une même AD/RSS auront, vraisemblablement, des données météorologiques et des caractéristiques socioéconomiques corrélées.

Aussi, ces mêmes modèles seront reformulés sous forme de modèles mixtes/multiniveau (Rabe-Hesketh & Skrondal, 2008). Par exemple, l'association entre la mortalité par MCV et les variables météorologiques sera réexaminée via une régression de Poisson multiniveau qui incorpore la structure hiérarchique décrite ci-haut (voir Figure 3) ainsi que les variables associées (Tarkiainen et al., 2010).

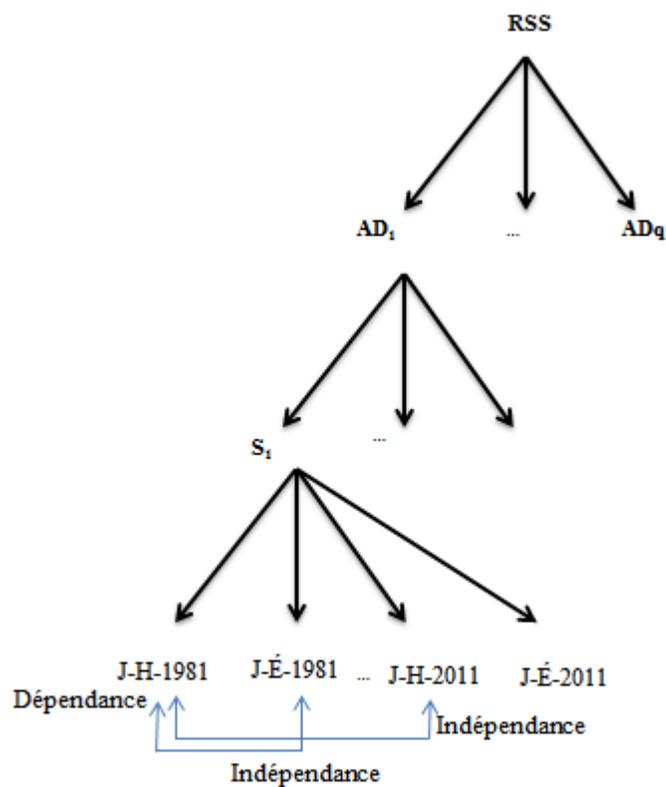


Figure 3 Structures des données (1981 à 2011) en trois niveaux, S : strate construite, J-H : les jours d'hiver et J-É : les jours d'été.

5.3.3. Partie C : Données sanitaires agrégées

Cette partie est scindée en deux, la sous-partie C1 et la sous-partie C2. C1 porte sur la modélisation par la régression de Cox multiniveau de la relation décès/météo. C2 concerne la mise en œuvre du modèle APC et l'exploration du diagramme de Lexis.

✓ **Sous-partie C1 : Cox multiniveau**

Finalement, les données individuelles (p. ex., âge) et contextuelles pourraient être incorporées dans des modèles multiniveau où l'évènement d'intérêt est le décès par MCV (données sanitaires non agrégées). Nous chercherons à modéliser le délai de survenue du décès par MCV chez les individus dont les suivis varient de 1981 à 2011. Les suivis qui se terminent par un décès non MCV ou par la fin de l'étude seront considérés comme des données censurées. Ce type de variable est usuellement modélisé par une régression de Cox (Cox, 1972; Goldstein, 2003; Rabe-Hesketh & Skrondal, 2012). Plusieurs études ont utilisé ce modèle statistique pour expliquer la relation entre une condition sanitaire et la pollution (Bates et al., 2006; Miller et al., 2007; Pinheiro et al., 2008). L'une des raisons évoquées pour expliquer ce constat serait le fait que les polluants atmosphériques ont un effet linéaire monotone sur la mortalité, tandis que la température lui est souvent associée de façon non linéaire (Zanobetti et al., 2012).

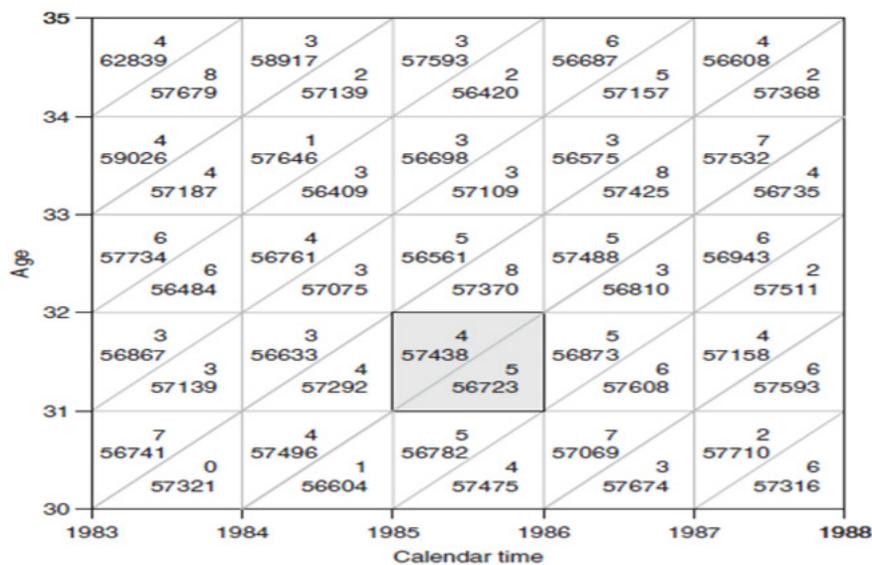
Nous sommes donc, ici aussi, en présence d'une structure hiérarchique à trois niveaux : sujets nichés dans les AD, les AD nichées dans les RSS. Ainsi, nous disposons de variables individuelles, de variables environnementales et contextuelles. Nous supposons que les données sur les décès au Québec entre 1981 et 2011 sont complètes et qu'elles reflètent bien l'évolution de la mortalité toutes causes confondues. Toute personne ne figurant pas dans le registre de mortalité est implicitement considérée vivante. Si besoin est, les valeurs des variables météorologiques dans les zones sans station de mesure pourraient être estimées à l'aide d'une interpolation spatiale, comme suggéré par Firestone (2012).

En somme, nous construirons des modèles à risques proportionnels de Cox adaptés à la structure hiérarchique des données et pour lesquels l'évènement d'intérêt est le décès par MCV. Nous y tiendrons compte de la variabilité temporelle des variables explicatives dépendantes du temps (p. ex. Liang & Zeger, 1986; Rabe-Hesketh & Skrondal, 2008) et de facteurs individuels (âge et genre), contextuels (ID et densité de population de l'AD de résidence), ainsi que des variables qui

touchent une plus vaste région (les variables environnementales). Enfin, l'effet mesuré est un risque relatif; le seuil de rejet statistique retenu (α) est de 5 %.

✓ Sous-partie C2 : Modèle APC, Diagrammes de Lexis

Le diagramme de Lexis (Carstensen, 2005; Carstensen, 2007; Clayton & Schiffers, 1987; Efron, 2002; Holford, 1983; Holford et al., 2006) est initialement créé pour un usage démographique (représentation synthétique des données de mortalité), mais rapidement introduit en épidémiologie. Il a été construit à l'origine pour représenter les taux de mortalité par âge (Carstensen, 2005). Actuellement, il est utilisé, notamment en cancérologie, pour déterminer les taux d'incidence par âge et cohorte, ou âge et période, et étudier leur évolution (Belot et al., 2008; Holford et al., 2006; Remontet et al., 2008). L'exploitation du diagramme passe par une organisation des données et le calcul des différents effectifs associés au regroupement des données (Voir Figure 4 à titre illustratif, (Carstensen, 2007)).



Danish testis cancer cases and male person-years in ages 30–34 and period 1983–1987, subdivided by age, period and cohort in 1-year classes (subdivision of the bold face entry in Table I). For example, in the highlighted square, there were 9 cases diagnosed in age 31 in the year 1985; 4 born in 1953 (upper triangle, persons who cross the 32 age line in 1985) and 5 born in 1954 (lower triangle, persons who cross the 31 age line in 1985).

Figure 4 Nombre de personnes-années et le nombre de cas.

Le modèle « âge-période-cohorte » (APC) a été développé comme outil d'interprétation du diagramme de Lexis. Il permet ainsi d'étudier l'évolution au cours du temps (tendance) d'un

phénomène ou d'une mesure, à travers les effets temporels de la cohorte (effet de génération) et de la période d'observation (temps calendaire), ajustés sur l'âge des individus au moment de l'évènement d'intérêt (Carstensen, 2007). L'objectif de l'analyse APC est de déterminer les contributions respectives des effets de l'âge, de la période et de la cohorte à l'évolution du phénomène étudié (Holford, 1983) (Voir Figure 5 à titre illustratif, (Carstensen, 2007))

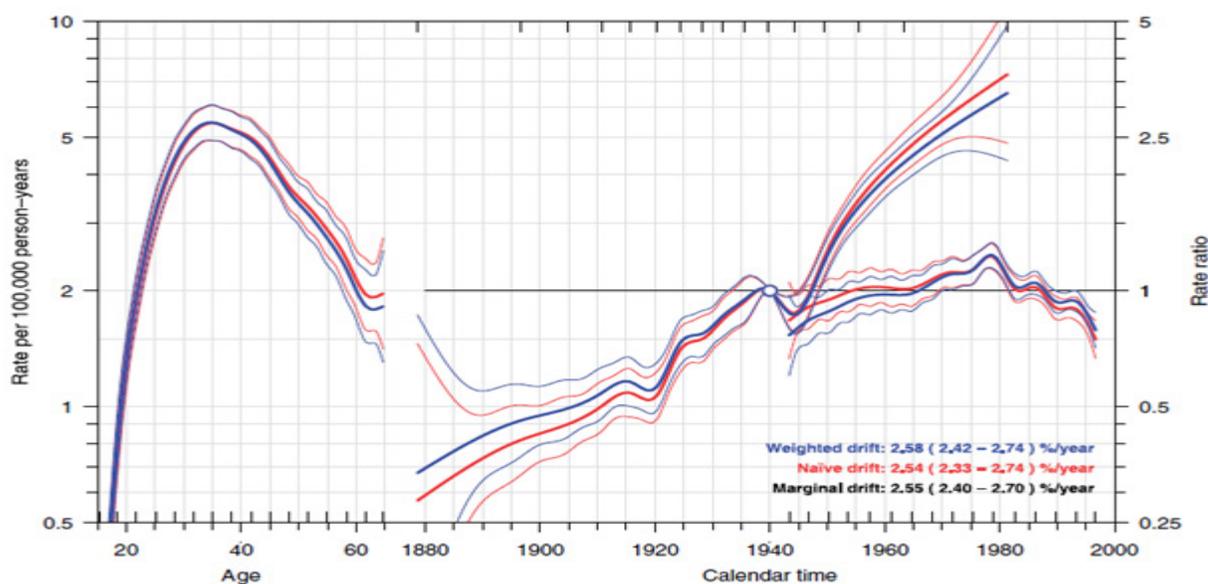


Figure 5 Illustration du modèle APC pour les données de cancer au Danemark

Dans un premier temps, nous utiliserons le diagramme de Lexis pour représenter les évènements d'intérêt (p. ex., décès par MCV) par des trajectoires (lignes obliques de 45 degrés) en fonction de la date calendaire (période), de l'âge ou de la date de naissance (cohorte), etc. (voir Figure 6 à titre illustratif, (Carstensen, 2007)). Puis nous exploiterons les diverses tables qui lui sont associées pour construire un modèle APC qui exprimera les taux de décès MCV comme un produit des effets de l'âge, et/ou de la période et/ou de la cohorte (Carstensen, 2007; Clayton & Schifflers, 1987). Nous aurons ainsi un aperçu sur les amplitudes des taux, sur la variation par âge et sur la tendance temporelle de ces taux. Cette approche repose sur un découpage assez fin des échelles de temps, et traite le temps et les variables qui en dépendent comme de simples covariables, sans

aucun statut particulier. Les analyses stratifiées deviennent des interactions entre le temps et les variables catégorielles, alors que les coefficients qui dépendent du temps deviennent des interactions entre le temps et les variables continues. Les modèles construits dans le cadre de cette approche se résument à de simples régressions de Poisson (Carstensen, 2006).

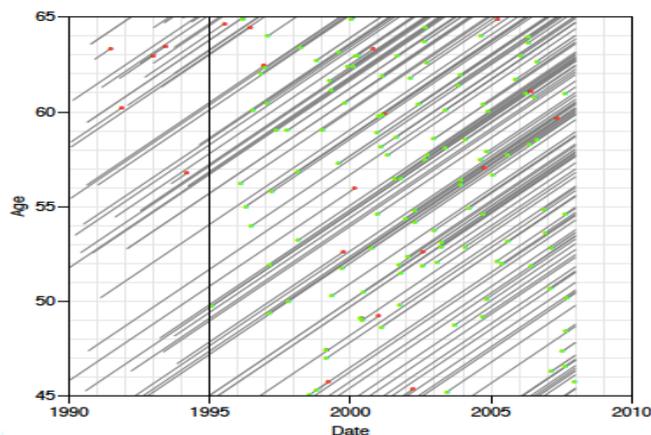


Figure 6 Durées de vie avec deux échelles de temps : date calendrier et âge; les points rouges=décès ou autre évènement, les points verts=diagnostic d'une maladie ou autre évènement.

6. Apport à la recherche

Le but ultime du projet est de soutenir la surveillance en santé publique, en améliorant les connaissances épidémiologiques et méthodologiques sur la mortalité pour cause de MCV dans un contexte de CC.

6.1. Volet épidémiologique

À notre connaissance, seules trois publications ont porté sur la relation entre la météorologie et les MCV au Québec. La première concernait la morbidité par cardiopathies ischémiques de 1989 à 2006 (Bayentin et al., 2010). La seconde avait trait à la mortalité saisonnière et aux tendances d'hospitalisations pour cause d'insuffisance cardiaque congestive au Québec de 1990 à 1998

(Feldman et al., 2004). La troisième portait aussi sur la mortalité cause d'insuffisance cardiaque congestive à Montréal de 1984 à 1993 (Kolb et al., 2007). Ces maladies sont incluses dans le regroupement de MCV étudiées dans ce projet. Bien qu'intéressantes, ces études n'explorent pas le lien entre la mortalité par MCV et les données météorologiques historiques, tout en introduisant les effets confondants ou aggravants plausibles des polluants atmosphériques ou en évaluant d'autres effets potentiellement confondants (comme l'ID). Aucune étude n'a projeté la relation entre les MCV et la météorologie dans le futur, et aucune n'a proposé un spectre aussi large de variables environnementales à mettre en relation avec les MCV.

Les résultats finaux de ces travaux serviront :

- 1- À soutenir les efforts des pouvoirs publics pour mettre en place des stratégies et plans d'actions pouvant contribuer à atténuer les effets du climat, notamment sur le plan clinique auprès des personnes atteintes de problèmes cardiovasculaires.
- 2- Adopter des mesures préventives ou/et en instaurer des systèmes d'alerte et de surveillance. Ainsi, un nombre important de ces décès pourraient être évités avec un avertissement adéquat et une réponse appropriée.

De façon générale, ils seront à même d'éclairer davantage les effets de la météorologie/climat sur la santé humaine et déterminer l'impact réel de la météo sur les décès pour causes de MCV.

6.2. Volet méthodologique

Les outils méthodologiques développés dans le cadre du présent projet seront appliqués sur des données des décès MCV. Les résultats de cette application serviront à des fins de prévention pour les MCV (volet épidémiologique). Ces outils méthodologiques sont généraux et ainsi pourront être utilisés pour d'autres maladies dans le cadre des CC. L'utilisation de ces outils s'avère donc des plus innovateurs pour répondre aux problématiques de la section 3 :

1. Les études antérieures nous présentent une panoplie de choix au niveau des variables confondantes, des lags et des méthodes de modélisation. Ce qui rend difficile le choix

pertinent de chacun de ces aspects. Prenons par exemple le cas de la température, ses effets retardés (froid ou chaud) dans une zone géographique ne peuvent pas être appliqués à une autre zone (Anderson et Bell 2009) car la sensibilité de la population à ces différents effets varie selon le lieu géographique (p. ex. Kovats 2004, Robinson 2001). Certaines études ont opté pour des effets retardés d'une façon subjective en se référant à des études précédentes. Voilà pourquoi, ce projet vise à concevoir une procédure méthodologique claire et globale orientant le choix des différents aspects.

2. Notre stratégie basée sur la construction des strates en combinant les modalités des variables individuelles (âge, genre) et contextuelles (ID) puis faire leurs suivi dans le temps est une approche novatrice dans le cadre santé/climat. Cette stratégie fait appel aux modèles marginaux de type GEE et multiniveau permettant d'avoir une vue d'ensemble, sans négliger les particularités de chacun des sous-groupes de la population (strate).
3. Les données sanitaires agrégées étaient utilisées dans toutes les études de la relation mortalité ou morbidité climat. L'exploitation de ces données ne permet pas de mettre en évidence la disparité entre les personnes au niveau individuel et contextuel, sans compter l'absence de mesure d'effet permettant d'évaluer la contribution des différences observées pour chacune de ces variables. La reconstitution des bases de données d'une façon rétrospective et le retraçage de l'historique des sujets à l'étude en allant chercher de l'information disponible dans les fichiers médicaux et les jumeler aux bases de données environnementales (météo, polluants), nous permettra d'aboutir à des bases de données avec de l'information individuelle avec des données non agrégées. Le traitement de ces bases en tenir compte des niveaux d'agrégation emboîtés, met en œuvre les modèles de survie multiniveau (régression de Cox). Cette manœuvre est aussi novatrice dans le cadre des relations santé climat.

4. Faire une analyse exploratoire des taux de mortalité MCV via Le digramme de Lexis, nous permet de voir leur évolution à travers les trois axes du temps, âge, période et cohorte. Tel que expliqué dans la partie problématique, ces trois axes de temps influencent l'évolution des phénomènes étudiés. Il est très important de contrôler au maximum la variabilité qui ne serait pas attribuable aux effets de l'environnement. Faire une analyse sans tenir compte de l'effet de chacun de ces axes ne garantit pas la fiabilité des résultats. Le modèle APC est utilisé pour quantifier l'effet de l'âge, cohorte et période sur le phénomène étudié (p.ex. taux d'incidence du cancer). Notre stratégie est d'intégrer les variables environnementales dans ce type de modèle déjà adapté à modéliser ces trois effets. Ainsi que vérifier l'impact de l'intégration de ces variables sur la modélisation APC. Cette intégration est la majeure innovation dans l'ensemble du présent projet.

D'où la pertinence de l'ensemble des approches présentées dans ce projet en termes méthodologies et en termes d'application aux MCV au Québec – une première au Canada, à notre connaissance.

7. Implémentation

La majorité de l'implémentation de l'analyse statistique sera réalisée avec le logiciel R (R Development Core Team, 2012). De plus, de nombreux packages fournissent les programmes nécessaires à la réalisation du projet.

Partie A : Procédure méthodologique

Le logiciel R sera considéré avec les packages MASS (Ripley et al., 2014), DLNM (Gasparrini, 2011) et sensitivity (Pujol et al., 2014) pour l'analyse de sensibilité du modèle.

Partie B : GEE-Modèle multiniveau

Le logiciel R sera considéré avec les packages lme4 (Bates et al., 2006), nlme (Pinheiro et al., 2008) ou le logiciel MLwin.

La section 8 du présent document a été retirée pour des raisons de droits de propriété intellectuelle sur les données qu'elle modélisait.

Bibliographie

- Abbey D, Nishino N, William F, McDonnell R, Burchette R, Knutsen S, Beeson W & Yang J (1999) Long-Term inhalable particles and other air pollutants related to mortality in nonsmokers. *Am J Respir Crit Care Med* 159:373 - 382.
- Anderson BG & Bell ML (2009) Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology (Cambridge, Mass.)* 20(2):205.
- Armstrong B (2006) Models for the relationship between ambient temperature and daily mortality. *Epidemiology* 17(6):624-631.
- Armstrong BG, Chalabi Z, Fenn B, Hajat S, Kovats S, Milojevic A & Wilkinson P (2011) Association of mortality with high temperatures in a temperate climate: England and Wales. *Journal of epidemiology and community health* 65(4):340-345.
- Ballester F, Corella D, Pérez-Hoyos S, Sáez M & Hervás A (1997) Mortality as a function of temperature. A study in Valencia, Spain, 1991-1993. *International journal of epidemiology* 26(3):551-561.
- Barnett A, Hajat S, Gasparrini A & Rocklöv J (2012) Cold and heat waves in the United States. *Environmental research* 112:218-224.
- Barnett AG, Tong S & Clements A (2010) What measure of temperature is the best predictor of mortality? *Environmental research* 110(6):604-611.
- Basu R (2009) High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environ Health* 8(1):40.
- Basu R & Ostro BD (2008) A multicounty analysis identifying the populations vulnerable to mortality associated with high ambient temperature in California. *American journal of epidemiology* 168(6):632-637.
- Basu R & Samet J (2002) Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiol Rev* 24:190 - 202.
- Bates D, Maechler M & Bolker B (2006) *Package 'lme4'*. <http://cran.r-project.org/web/packages/lme4/lme4.pdf>
- Battaieb J, Toumi A, Leffondre K, Ben Alaya N, Boulthir A, Chlif S, Hajem S & Ben Salah E (2010) Relation temperature-mortalite pour la ville de Tunis: 2005-2007.
- Bayentin L, El Adlouni S, Ouarda TBMJ, Gosselin P, Doyon B & Chebana F (2010) Spatial variability of climate effects on ischemic heart disease hospitalization rates for the period 1989-2006 in Quebec, Canada. *International Journal of Health Geographics* 9.
- Bélanger D, Gosselin P, Valois P & Abdous B (2006) Vagues de froid au Québec méridional : adaptations actuelles et futures. .).
- Bell ML, O'Neill MS, Ranjit N, Borja-Aburto VH, Cifuentes LA & Gouveia NC (2008) Vulnerability to heat-related mortality in Latin America: a case-crossover study in São Paulo, Brazil, Santiago, Chile and Mexico City, Mexico. *International Journal of Epidemiology* 37(4):796-804.
- Belot A, Grosclaude P, Bossard N, Jouglu E, Benhamou E, Delafosse P, Guizard A-V, Molinié F, Danzon A & Bara S (2008) Cancer incidence and mortality in France over the period 1980–2005. *Revue d'épidémiologie et de santé publique* 56(3):159-175.
- Besancenot J-P (2002) Vagues de chaleur et mortalité dans les grandes agglomérations urbaines. *Environnement, risques & santé* 1(4):229-240.
- Besancenot JP & Besancenot J-P (1997) Les grands paroxysmes climatiques et leurs répercussions sur la santé. *Bioclimatologie*. 134(4):237-246.

- Blais C, Hamel D & Rinfret S (2012) Impact of socioeconomic deprivation and area of residence on access to coronary revascularization and mortality after a first acute myocardial infarction in Québec. *Canadian Journal of Cardiology* 28(2):169-177.
- Blum L, Bresolin L & Williams M (1998) AMA Council on Scientific Affairs: Heat-related illness during extreme weather emergencies. *Journal of the American Medical Association*, vol. 279, no 19, :p. 1514.
- Bouchard M & Smargiassi A (2009) *Estimation des impacts sanitaires de la pollution atmosphérique au Québec: essai d'utilisation du Air Quality Benefits Assessment Tool (AQBAT)*. Direction des risques biologiques, environnementaux et occupationnels, Institut national de santé publique du Québec,
- Braga A, Zanobetti A & Schwartz J (2001a) The time course of weather-related deaths. *Epidemiology* 12:662 - 667.
- Braga ALF, Saldiva PHN, Pereira LAA, Menezes JJC, Conceição G, Lin CA, Zanobetti A, Schwartz J & Dockery DW (2001b) Health effects of air pollution exposure on children and adolescents in São Paulo, Brazil*†. *Pediatric Pulmonology* 31(2):106-113.
- Braga ALF, Zanobetti A & Schwartz J (2002) The effect of weather on respiratory and cardiovascular deaths in 12 US cities. *Environmental Health Perspectives* 110(9):859.
- Breitner S, Stölzel M, Cyrus J, Pitz M, Wölke G, Kreyling W, Küchenhoff H, Heinrich J, Wichmann HE & Peters A (2009) Short-term mortality rates during a decade of improved air quality in Erfurt, Germany. *Environmental Health Perspectives* 117(3):448.
- Bustinza R, Tairou FO, Gosselin P & Bélanger D (2010) Proposition d'indicateurs aux fins de vigie et de surveillance des troubles de la santé liés au froid. .)
- Carder M, McNamee R, Beverland I, Elton R, Cohen G, Boyd J & Agius R (2005) The lagged effect of cold temperature and wind chill on cardiorespiratory mortality in Scotland. *Occupational and environmental medicine* 62(10):702-710.
- Carstensen B (2005) Demography and epidemiology: Practical use of the Lexis diagram in the computer age. *Annual meeting of Finnish Statistical Society*. p 24.
- Carstensen B (2007) Age–period–cohort models for the Lexis diagram. *Statistics in medicine* 26(15):3018-3045.
- Carstensen B, Plummer M, Hills M & Laara E (2013) *Package 'Epi'*. <http://cran.r-project.org/web/packages/Epi/Epi.pdf>
- Chaix B & Chauvin P (2002) L'apport des modèles multiniveau dans l'analyse contextuelle en épidémiologie sociale: une revue de la littérature. *Revue d'épidémiologie et de santé publique* 50(5):489-499.
- Chen R, Pan G, Zhang Y, Xu Q, Zeng G, Xu X, Chen B & Kan H (2011) Ambient carbon monoxide and daily mortality in three Chinese cities: the China Air Pollution and Health Effects Study (CAPES). *Science of The Total Environment* 409(23):4923-4928.
- Clayton D & Schifflers E (1987) Models for temporal variation in cancer rates. II: age–period–cohort models. *Statistics in medicine* 6(4):469-481.
- Conceicao G, Miraglia S, Kishi HS, Saldiva P & Singer JM (2001) Air pollution and child mortality: a time-series study in São Paulo, Brazil. *Environmental Health Perspectives* 109(Suppl 3):347.
- Conti S, Meli P, Minelli G, Solimini R, Toccaceli V, Vichi M, Beltrano C & Perini L (2005) Epidemiologic study of mortality during the Summer 2003 heat wave in Italy. *Environ Res* 98:390 - 399.
- Cox DR (1972) Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 34(2):187-&.

- Crabbe H (2012) Risk of respiratory and cardiovascular hospitalisation with exposure to bushfire particulates: new evidence from Darwin, Australia. *Environmental geochemistry and health* 34(6):697-709.
- Curriero F, Heiner K, Samet J, Zeger S, Strug L & Patz J (2002) Temperature and mortality in 11 cities of the eastern United States. *Am J Epidemiol* 155:80 - 87.
- Dominici F, McDermott A, Zeger S & Samet J (2002) On the Use of Generalized Additive Models in Time-Series Studies of Air Pollution and Health. *Am J Epidemiol* 156:193 - 203.
- Doyon B, Belanger D & Gosselin P (2008) The potential impact of climate change on annual and seasonal mortality for three cities in Quebec, Canada. *International Journal of Health Geographics* 7(1):23.
- Doyon B, Bélanger D & Gosselin P (2006) Effets du climat sur la mortalité au Québec méridional de 1981 à 1999 et simulations pour des scénarios climatiques futurs. Québec, Québec, Canada : Institut national de santé publique du Québec. .).
- Ebi KL, Teisberg TJ, Kalkstein LS, Robinson L & Weiher RF (2004) Heat Watch/Warning Systems Save Lives: Estimated Costs and Benefits for Philadelphia 1995–98. *Bulletin of the American Meteorological Society* 85(8).
- Efron B (2002) The two-way proportional hazards model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64(4):899-909.
- Faivre R, Looss, B., Makowski, D., Monod, H., (2013,) *Analyse de sensibilité et exploration de modèles, Application aux sciences de la nature et de l'environnement.*, 324 p
- Fassò A & Perri PF (2006) Sensitivity Analysis. *Encyclopedia of Environmetrics*, John Wiley & Sons, Ltd, 10.1002/9780470057339.vas015.
- Feldman DE, Platt R, Déry V, Kapetanakis C, Lamontagne D, Ducharme A, Giannetti N, Frenette M & Beck E (2004) Seasonal congestive heart failure mortality and hospitalisation trends, Quebec 1990–1998. *Journal of epidemiology and community health* 58(2):129-130.
- Filate WA, Johansen, H. L., Courtney, C .K., & Tu, J. V. (2006) Regional variations in cardiovascular mortality in Canada. J. V. Tu, W. A. Ghali, L. Pilote, & S. Brien, S. (Eds.), CCORT Canadian Cardiovascular Atlas. Toronto (Canada). *Pulsus Group Inc and Institute for Clinical Evaluative Sciences* 41-48.
- Filion P & Lauzier JJ (2002) L'énergie au Québec, ministère des Ressources naturelles.), p 131 p.
- Firestone SM, Cogger N, Ward MP, Toribio J-AL, Moloney BJ & Dhand NK (2012) The influence of meteorology on the spread of influenza: Survival analysis of an equine influenza (A/H3N8) outbreak. *PloS one* 7(4):e35284.
- Gasparrini A (2011) Distributed lag linear and non-linear models in R: the package dlnm. *Journal of Statistical Software* 43(8):1.
- Gasparrini A, Armstrong B & Kenward M (2010) Distributed lag non-linear models. *Statistics in medicine* 29(21):2224-2234.
- Gauthier J, Morais J & Mallet L (2005) Impact des vagues de chaleur en gériatrie et risques associés aux médicaments, Pharmactuel, .).
- Gauvin L & Dassa C (2004) L'analyse multiniveaux: avancées récentes et retombées anticipées pour l'étude des inégalités sociales et de santé. *Santé, société et solidarité* 3(2):187-195.
- GIEC (1990) Groupe d'experts intergouvernemental sur l'évolution du climat, Le premier Rapport d'évaluation.).
- GIEC (2001) Groupe d'experts intergouvernemental sur l'évolution du climat, Changements climatiques : Rapport de synthèse. http://www.grida.no/publications/other/ipcc_tar/
- GIEC (2013) Changements Climatiques 2013 Les éléments scientifiques, Rapport d'évaluation du Groupe d'experts intergouvernemental sur l'évolution du climat.).

- Giroux J-X (2011a) *"Effet du climat sur la morbidité extrême au Québec."* Mémoire de maîtrise. (INRS-ETE. 285p., Québec, Canada).
- Giroux J-X, Chebana F, Bélanger D, Gloaguen E, Ouarda T & St-Hilaire A (2013) Projet M1 : Comparaison de l'utilisation des moyennes spatiales à celle du krigeage appliquée à la relation mortalité par MCV - météorologie, au Québec, de 1996 à 2007, Rapport final - version préliminaire, INRS-ETE, Québec, Canada.
- Giroux JX (2011b) Exploration de l'effet de la météorologie sur la survenue de journées avec valeurs extrêmes de consultations aux urgences pour problèmes cardiovasculaires ou d'hospitalisations pour traumatismes dans la région de Montréal (Canada) de 1988 à 2008. Mémoire de maîtrise, Sciences de l'eau, INRS-ETE, Québec, Canada.
- Goldberg MS, Burnett RT, Bailar Iii JC, Brook J, Bonvalot Y, Tamblyn R, Singh R & Valois M-F (2001) The Association between Daily Mortality and Ambient Air Particle Pollution in Montreal, Quebec: 1. Nonaccidental Mortality. *Environmental Research* 86(1):12-25.
- Goldberg MS, Gasparrini A, Armstrong B & Valois M-F (2011) The short-term influence of temperature on daily mortality in the temperate climate of Montreal, Canada. *Environmental research* 111(6):853-860.
- Goldstein H (2003) *Multilevel statistical models*. Chichester: Wiley.,
- Gosling SN, Lowe JA, McGregor GR, Pelling M & Malamud BD (2009) Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. *Climatic Change* 92(3-4):299-341.
- Gosling SN, McGregor GR & Páldy A (2007) Climate change and heat-related mortality in six cities Part 1: model construction and validation. *International journal of biometeorology* 51(6):525-540.
- Gouveia N, Hajat S & Armstrong B (2003a) Socioeconomic differentials in the temperature-mortality relationship in Sao Paulo, Brazil. *Int J Epidemiol* 32:390 - 397.
- Gouveia N, Hajat S & Armstrong B (2003b) Socioeconomic differentials in the temperature-mortality relationship in São Paulo, Brazil. *International Journal of Epidemiology* 32(3):390-397.
- Grahame TJ & Schlesinger RB (2010) Cardiovascular health and particulate vehicular emissions: a critical evaluation of the evidence. *Air Quality, Atmosphere & Health* 3(1):3-27.
- Ha J & Kim H (2013) Changes in the association between summer temperature and mortality in Seoul, South Korea. *International journal of biometeorology* 57(4):535-544.
- Hajat S, Kovats RS, Atkinson RW & Haines A (2002) Impact of hot temperatures on death in London: a time series approach. *Journal of epidemiology and community health* 56(5):367-372.
- Hastie T & Tibshirani R (1986) Generalized Additive Models. *Statistical Science* 1(3):297-310.
- Hastie T & Tibshirani RJ (1990) Generalized Additive Models—Monographs on Statistics and Applied Probability 43. *Generalised additive models: Monographs on Statistics and Applied Probability* 43.
- Herjean P (2006) L'approche multiniveau de la santé. *Cahiers de géographie du Québec* 50(141):347-355.
- Hoek G, Brunekreef B, Goldbohm S, Fischer P & Brandt P (2002) Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *Lancet* 360(9341):1203 - 1209.
- Holford TR (1983) The estimation of age, period and cohort effects for vital rates. *Biometrics*:311-324.

- Holford TR, Cronin KA, Mariotto AB & Feuer EJ (2006) Changing patterns in breast cancer incidence trends. *JNCI Monographs* 2006(36):19-25.
- Holguín F, Téllez-Rojo MM, Hernández M, Cortez M, Chow JC, Watson JG, Mannino D & Romieu I (2003) Air pollution and heart rate variability among the elderly in Mexico City. *Epidemiology* 14(5):521-527.
- Hori. A HM, Tsuda. Y, Tsukahara. T, Nomiya. T, (2012) Effects of weather variability and air pollutants on emergency admissions for cardiovascular and cerebrovascular diseases. *International Journal of Environmental Health Research*.
- Huang C, Barnett AG, Wang X & Tong S (2012) Effects of extreme temperatures on years of life lost for cardiovascular deaths: a time series study in brisbane, australia. *Circulation: Cardiovascular Quality and Outcomes* 5(5):609-614.
- Huang C, Vaneckova P, Wang X, FitzGerald G, Guo Y & Tong S (2011) Constraints and Barriers to Public Health Adaptation to Climate Change: A Review of the Literature. *American Journal of Preventive Medicine* 40(2):183-190.
- Ibald-Mulli A, Timonen KL, Peters A, Heinrich J, Wölke G, Lanki T, Buzorius G, Kreyling WG, De Hartog J & Hoek G (2004) Effects of particulate air pollution on blood pressure and heart rate in subjects with cardiovascular disease: a multicenter approach. *Environmental Health Perspectives* 112(3):369.
- INSPQ (2009) Guide méthodologique «L'indice de défavorisation sociale: en bref».)
- IPCC (ca. 2011,) Les maladies cardiovasculaires. .)
- Ishigami A, Hajat S, Kovats RS, Bisanti L, Rognoni M, Russo A & Paldy A (2008) An ecological time-series study of heat-related mortality in three European cities. *Environ Health* 7(5):1-7.
- ISQ (2000) Perspectives démographiques du Québec 1996-2041, régions administratives, régions métropolitaines et municipalités régionales de comté.)
- ISQ (2012) *Population et structure par âge et sexe.*, http://www.stat.gouv.qc.ca/donstat/societe/demographie/dons_regnl/regional/213.htm
- Jbilou J & El Adlouni S (2012) Generalized Additive Models in Environmental Health : A Literature Review.
- Kalkstein L (1991) A new approach to evaluate the impact of climate on human mortality. *Environ Health Perspect* 96:145 - 150.
- Kalkstein LS & Davis RE (1989) Weather and human mortality: an evaluation of demographic and interregional responses in the United States. *Annals of the Association of American Geographers* 79(1):44-64.
- Kenney WL & Hodgson JL (1987) Heat tolerance, thermoregulation and ageing. *Sports Medicine* 4(6):446-456.
- Kleinbaum DG, Kupper LL & Muller KE (2007) *Applied regression analysis and other multivariable methods*. Duxbury Pr,
- Kolb S, Radon K, Valois M-F, Héguay L & Goldberg MS (2007) The short-term influence of weather on daily mortality in congestive heart failure. *Archives of environmental & occupational health* 62(4):169-176.
- Kovats R, Hajat S & Wilkinson P (2004) Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in Greater London, UK. *Occup Environ Med* 61:893 - 898.
- Kovats R, Hajat, S. (2008) Heat stress and public health: a critical review,. *Annual Review of Public Health*, vol. 29, :p. 41-55.

- Künzli N, Perez, L., Rapp, R. (2010) Qualité de l'air et santé, European Respiratory Society, Swiss Tropical and Public Health Institute, Switzerland, 68 p. Cité dans: Lebel, G., D. Busque, M. Therrien, P. Walsh, J. Paradis, M.-P. Brault et M. Canuel (2012). Bilan de la qualité de l'air au Québec en lien avec la santé, 1975-2009. .).
- Lavigne E, Gasparrini A, Wang X, Chen H, Yagouti A, Fleury MD & Cakmak S (2014) Extreme ambient temperatures and cardiorespiratory emergency room visits: assessing risk by comorbid health conditions in a time series study. *Environmental Health* 13(1):5.
- Lebel G BD, Therrien M, Walsh P, Paradis J, Brault M-P, Canuel M, (2012) Bilan de la qualité de l'air au Québec en lien avec la santé, 1975-2009. .).
- Lebel G, Busque D, Therrien M, Walsh P, Paradis J, Brault M-P & Canuel M (2012) Bilan de la qualité de l'air au Québec en lien avec la santé, 1975-2009.).
- Ledrans M, Isnard, H (2003) *Impact sanitaire de la vague de chaleur d'août 2003 en France : bilan et perspectives, octobre 2003. Institut de veille sanitaire, Saint-Maurice, France, 120 p.*,
- Lepage MP, Bourdages L & Bourgeois G (2011) Interprétation des scénarios de changements climatiques.
- Liang KY & Zeger SL (1986) Longitudinal Data-Analysis Using Generalized Linear-Models. *Biometrika* 73(1):13-22.
- Liang W-M, Liu W-P, Chou S-Y & Kuo H-W (2008) Ambient temperature and emergency room admissions for acute coronary syndrome in Taiwan. *International journal of biometeorology* 52(3):223-229.
- Liang W-M, Liu W-P & Kuo H-W (2009) Diurnal temperature range and emergency room admissions for chronic obstructive pulmonary disease in Taiwan. *International journal of biometeorology* 53(1):17-23.
- Liu L, Breitner S, Pan X, Franck U, Leitte AM, Wiedensohler A, von Klot S, Wichmann H-E, Peters A & Schneider A (2011) Associations between air temperature and cardio-respiratory mortality in the urban area of Beijing, China: a time-series analysis. *Environ Health* 10(1):51.
- Luber G, McGeehin, M. (2008) Climate change and extreme heat events,. *American Journal of Preventive Medicine*, vol. 35, no 5, (p. 429-435).
- Lundgren L & Jonsson A (2012) Assessment of social vulnerability. A literature review of vulnerability related to climate change and natural hazards.).
- Madrigano J, Mittleman MA, Baccarelli A, Goldberg R, Melly S, Von Klot S & Schwartz J (2013) Temperature, myocardial infarction, and mortality: Effect modification by individual- and area-level characteristics. *Epidemiology* 24(3):439-446.
- Mastrangelo G, Fedeli U, Visentin C, Milan G, Fadda E & Spolaore P (2007) Pattern and determinants of hospitalization during heat waves: an ecologic study. *BMC Public Health* 7(1):200.
- McMichael AJ, Woodruff RE & Hales S (2006) Climate change and human health: present and future risks. *The Lancet* 367(9513):859-869.
- MDDEP (2002a) *Indice de qualité de l'air (IQA)*. Accessible au: <http://www.iqa.mddefp.gouv.qc.ca/contenu/polluants.htm>. Consulté le 2 octobre 2013.
- MDDEP (2002b) *Info-smog*. . Accessible au: <http://www.mddefp.gouv.qc.ca/air/info-smog/smog.htm>. Consulté le 2 octobre 2013.
- Medina-Ramón M & Schwartz J (2007) Temperature, temperature extremes, and mortality: a study of acclimatisation and effect modification in 50 US cities. *Occupational and environmental medicine* 64(12):827-833.

- Medina-Ramón M, Zanobetti A, Cavanagh DP & Schwartz J (2006) Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. *Environmental health perspectives* 114(9):1331.
- Michelozzi P, Accetta G, De Sario M, D'Ippoliti D, Marino C, Baccini M, Biggeri A, Anderson HR, Katsouyanni K & Ballester F (2009) High temperature and hospitalizations for cardiovascular and respiratory causes in 12 European cities. *American journal of respiratory and critical care medicine* 179(5):383-389.
- Middleton N, Yiallourous P, Kleanthous S, Kolokotroni O, Schwartz J, Dockery DW, Demokritou P & Koutrakis P (2008) A 10-year time-series analysis of respiratory and cardiovascular morbidity in Nicosia, Cyprus: the effect of short-term changes in air pollution and dust storms. *Environ Health* 7(39):10.1186.
- Miller K, Siscovick D, Sheppard L, Shepherd K, Sullivan J, Anderson G & Kaufman J (2007) Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med* 356(5):447 - 458.
- Muggeo VM & Hajat S (2009) Modelling the non-linear multiple-lag effects of ambient temperature on mortality in Santiago and Palermo: a constrained segmented distributed lag approach. *Occupational and Environmental Medicine* 66(9):584-591.
- Muggeo VMR (2008) Modeling temperature effects on mortality: multiple segmented relationships with common break points. *Biostatistics* 9(4):613-620.
- National Research Council (2008) *Estimating Mortality Risk Reduction and Economic Benefits from Controlling Ozone Air Pollution*. The National Academies Press. http://www.nap.edu/openbook.php?record_id=12198
- Nelder JA & Wedderburn RWM (1972) Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)* 135(3):370-384.
- O'Neill MS, Zanobetti A & Schwartz J (2003) Modifiers of the temperature and mortality association in seven US cities. *American Journal of Epidemiology* 157(12):1074-1082.
- OMS (2009a) Health and Global Environmental Change.).
- OMS (2009b) *Health and Global Environmental Change, Series, no 2*. . Accessible au :http://www.euro.who.int/data/assets/pdf_file/0008/96965/E82629.pdf, Consulté le 5 mai 2014.
- Ouranos (2010) Savoir s'adapter aux changements climatiques. Montréal, Québec, Canada.).
- Park AK, Hong YC & Kim H (2011) Effect of changes in season and temperature on mortality associated with air pollution in Seoul, Korea. *Journal of epidemiology and community health* 65(4):368-375.
- Pascal M, Falq G, Wagner V, Chatignoux E, Corso M & Declercq C (2012) Influence de la saison et des épisodes de fortes chaleurs sur les liens entre ozone, particules et mortalité dans neuf villes françaises. Saint-Maurice: Institut de veille sanitaire.).
- Pattenden S, Nikiforov B & Armstrong B (2003) Mortality and temperature in Sofia and London. *J Epidemiol Community Health* 57:628 - 633.
- Pelucchi C, Negri E, Gallus S, Boffetta P, Tramacere I & La Vecchia C (2009) Long-term particulate matter exposure and mortality: a review of European epidemiological studies. *BMC Public Health* 9(1):453.
- Peng RD & Dominici F (2008) *Statistical Methods for Environmental Epidemiology with R*.
- Peng RD, Dominici F & Louis TA (2006) *Model choice in time series studies of air pollution and mortality*. Blackwell Publishing Ltd. 179-203 p. <http://dx.doi.org/10.1111/j.1467-985X.2006.00410.x>

- Pinheiro J, Bates D, DebRoy S & Deepayan S (2008), <http://cran.r-project.org/web/packages/nlme/nlme.pdf>
- Pope C, Burnett R, Thurston G, Thun M, Calle E, Krewski D & Godleski J (2004) Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation* 109(1):71 - 77.
- Pope III CA, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D & Godleski JJ (2004) Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution Epidemiological Evidence of General Pathophysiological Pathways of Disease. *Circulation* 109(1):71-77.
- Pope III CA, Ezzati M & Dockery DW (2009) Fine-particulate air pollution and life expectancy in the United States. *New England Journal of Medicine* 360(4):376-386.
- Portier CJ, Thigpen Tart, K., Carter, S.R., Dilworth, C.H., Grambsch, A.E., Gohlke, J., (2010) A Human Health Perspective On Climate Change: A Report Outlining the Research Needs on the Human Health Effects of Climate Change. Research Triangle Park, NC: Environmental Health Perspectives. National Institute of Environmental Health Sciences).
- Pujol G, Iooss B, Janon A, Gilquin L, Le Gratiet L & Lemaitre P (2014) *Package 'sensitivity'*. <http://cran.r-project.org/web/packages/sensitivity/sensitivity.pdf>
- R Development Core Team (2012) R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.*
- Rabe-Hesketh S & Skrondal A (2008) *Multilevel and longitudinal modelling using Stata*. STATA press,
- Rabe-Hesketh S & Skrondal A (2012) *Multilevel and longitudinal modelling using Stata, Third Edition* STATA press,
- Rainham D & Smoyer-Tomic K (2003) The role of air pollution in the relationship between a heat stress index and human mortality in Toronto. *Environ Res* 93:9 - 19.
- Reid CE, O'Neill MS, Gronlund CJ, Brines SJ, Brown DG, Diez-Roux AV & Schwartz J (2009) Mapping community determinants of heat vulnerability. *Environmental Health Perspectives* 117(11):1730.
- Remontet L, Esteve J, Bouvier A-M, Grosclaude P, Launoy G, Menegoz F, Exbrayat C, Tretare B, Carli P-M & Guizard A-V (2008) Cancer incidence and mortality in France over the period 1978-2000.
- Ren C, O'Neill MS, Park SK, Sparrow D, Vokonas P & Schwartz J (2011) Ambient Temperature, Air Pollution, and Heart Rate Variability in an Aging Population. *American journal of epidemiology* 173(9):1013-1021.
- Ripley B, Venables B, Bates D, Hornik K, Gebhardt A & Firth D (2014) *Package 'MASS'*. <http://cran.r-project.org/web/packages/MASS/MASS.pdf>
- Robinson PJ (2000) Temporal trends in United States dew point temperatures. *International Journal of Climatology* 20(9):985-1002.
- Sala C-A (2009) *Contribution du modèle Age-Période-Cohorte à l'étude de l'épizootie d'Encéphalopathie Spongiforme Bovine en France et en Europe*. (Université Claude Bernard-Lyon I).
- Santé Canada (2011) *Lignes directrices à l'intention des travailleurs de la santé pendant les périodes de chaleur accablante : Un guide technique*. Bureau de l'eau, de l'air et des changements climatiques, Direction générale de la santé environnementale et de la sécurité des consommateurs, Ottawa, Ontario. 165 p
- Schwartz J (1994) Generalized additive models in epidemiology. *International Biometric Society, Invited Papers. 17th International Biometric Conference*. p 12.

- Schwartz J (2000a) Assessing confounding, effect modification, and thresholds in the association between ambient particles and daily deaths *Environ Health Perspect* 108: 563–568. *Find this article online.*
- Schwartz J (2000b) The distributed lag between air pollution and daily deaths. *Epidemiology* 11(3):320-326.
- Schwartz J, Laden F & Zanobetti A (2002) The concentration-response relation between PM (2.5) and daily deaths. *Environmental Health Perspectives* 110(10):1025.
- Schwartz J, Samet JM & Patz JA (2004) Hospital admissions for heart disease: the effects of temperature and humidity. *Epidemiology* 15(6):755-761.
- SFSP (1996) *La pollution atmosphérique d'origine automobile et la santé publique, collection santé et société série N° 4.*, <http://fulltext.bdsp.ehesp.fr/Sfsp/Rapports/pollution1996.pdf>
- Tairou FO, Bélanger D & Gosselin P (2010) Proposition d'indicateurs aux fins de vigie et de surveillance des troubles de la santé liés à la chaleur.).
- Tam WW, Wong TW, Chair S & Wong AH (2009) Diurnal temperature range and daily cardiovascular mortalities among the elderly in Hong Kong. *Archives of environmental & occupational health* 64(3):202-206.
- Tarkiainen L, Martikainen P, Laaksonen M & Leyland AH (2010) Comparing the effects of neighbourhood characteristics on all-cause mortality using two hierarchical areal units in the capital region of Helsinki. *Health & Place* 16(2):409-412.
- Therneau T (2014) *Package 'survival'*. <http://cran.r-project.org/web/packages/survival/survival.pdf>
- Tian Z, Li S, Zhang J, Jaakkola JJ & Guo Y (2012) Ambient temperature and coronary heart disease mortality in Beijing, China: a time series study. *Environ Health* 11(56):10.1186.
- Vallée J (2009) Les disparités spatiales de santé en ville: l'exemple de Vientiane (Laos). *Cybergeo: European Journal of Geography*.
- Vutcovici M, Goldberg M & Valois M-F (2013) Effects of diurnal variations in temperature on non-accidental mortality among the elderly population of Montreal, Québec, 1984–2007. *International Journal of Biometeorology* 10.1007/s00484-013-0664-9:1-10.
- Wang L, Tong S, Toloo G & Yu W (2014) Submicrometer particles and their effects on the association between air temperature and mortality in Brisbane, Australia. *Environmental Research* 128(0):70-77.
- Wang XY, Barnett AG, Hu W & Tong S (2009) Temperature variation and emergency hospital admissions for stroke in Brisbane, Australia, 1996–2005. *International journal of biometeorology* 53(6):535-541.
- WHO (2009) Protecting health from climate change : global research priorities. *World Health Organisation*.
- Wood S (2006) *Generalized additive models: an introduction with R*.
- Wood SN (2001) mgcv: GAMs and generalized ridge regression for R. *R news* 1(2):20-25.
- Ye X, Wolff R, Yu W, Vaneckova P, Pan X & Tong S (2012) Ambient temperature and morbidity: a review of epidemiological evidence. *Environmental health perspectives* 120(1):19.
- Zanobetti A, O'Neill MS, Gronlund CJ & Schwartz JD (2012) Summer temperature variability and long-term survival among elderly people with chronic disease. *Proceedings of the National Academy of Sciences* 109(17):6608-6613.
- Zauli Sajani S, Garaffoni G, Goldoni CA, Ranzi A, Tibaldi S & Lauriola P (2002) Mortality and bioclimatic discomfort in Emilia-Romagna, Italy. *Journal of Epidemiology and Community Health* 56(7):536-537.

Zhang Y, Guo Y, Li G, Zhou J, Jin X, Wang W & Pan X (2012) The spatial characteristics of ambient particulate matter and daily mortality in the urban area of Beijing, China. *Science of The Total Environment* 435–436(0):14-20.

Annexe

1. DLNM

La formulation de DLNM est un peu complexe, impliquant des matrices à trois dimensions.

Soit C , une matrice des fonctions de base pour les lags de dimension $(L+1, vl)$ avec L est le lag maximal et Z une base pour la matrice X des prédicteurs de dimension (n, vx) . On ajoute les dimensions de lags pour chaque prédicteur dans X ce qui produit une matrice R de dimension $(n, vx, L+1)$, le DLNM peut être exprimé comme suit :

$$s(x_t, \eta) = \sum_{j=1}^{vx} \sum_{k=1}^{vl} r_{tj}^T c_{.k} \eta_{jk} = w_t^T \eta$$

où,

r_{tj} : est le vecteur des effets retardés pour le jour t transformée par la fonction de base j ;

η : un vecteur de paramètres inconnus ;

w_t : est le vecteur obtenu par l'application de la fonction des bases croisées vx et vl à la matrice X des prédicteurs.

Une fois ces bases sont construites, elles sont incluses dans un modèle GLM ou GAM.

2. Modèles GLM, GAM

Si l'on se restreint l'analyse au modèle linéaire standard, étant donné une variable réponse Y , supposée suivre une loi normale, et p prédicteurs X_1, \dots, X_p observés sur n individus, on a

l'expression suivante :

$$E[Y | X] = \alpha_0 + \alpha_1 X_1 + \dots + \alpha_p X_p.$$

Ce modèle est très utilisé en statistique, puisque les chercheurs désirent souvent connaître l'effet d'un groupe de variables explicatives (X_1, \dots, X_p) sur une variable dépendante « Y ». Quoiqu'il soit très intéressant par sa simplicité de calcul et d'interprétation, son utilisation est limitée quand la condition de normalité n'est pas remplie. Le modèle linéaire généralisé (GLM) est son extension. Ici l'hypothèse de normalité de Y est remplacée par une hypothèse plus large, celle d'appartenance de la distribution de Y à la famille exponentielle.

L'expression du modèle GLM est :

$$g(\mathbf{E}[Y | X]) = \alpha_0 + \alpha_1 X_1 + \dots + \alpha_p X_p.$$

Le modèle GAM étend à son tour le modèle GLM en utilisant des fonctions non linéaires pour représenter les variables explicatives. La forme générale du modèle GAM est donnée par l'équation suivante :

$$g(\mathbf{E}[Y | X]) = g[\mu] = \eta = \beta_0 + f_1(X_1) + \dots + f_p(X_p)$$

où y suit une distribution de la famille exponentielle, g est la fonction de lien qui dépend de la distribution de Y , les X_j sont les variables explicatives et les f_j sont des fonctions de lissage. Ces dernières sont représentées par une somme de fonctions de base $b_i(X_j)$ nommées bases des *splines* telles que :

$$f_j(X_j) = \sum_{i=1}^k b_i(X_j) \beta_i$$

où les β_i sont des paramètres inconnus à estimer. Ainsi, chaque variable explicative n'est plus nécessairement reliée de façon linéaire à la variable réponse. L'interprétation ne se fait plus à travers les coefficients α_i comme dans le GLM, mais selon les f_j il devient ainsi possible

d'analyser l'influence de chaque variable sur Y plus en détail au travers de ces fonctions. Ces dernières sont non linéaires et sont lissées à l'aide de splines, moyennes mobiles ou noyaux.

L'estimation du modèle se fait par une méthode itérative nommée P-IRLS². Pour estimer chaque fonction f_j , on utilise le maximum de vraisemblance pénalisé L_p où la pénalité contrôle le lissage de la courbe de f_j (Wood, 2006).

$$L_p = L(\beta) + \frac{1}{2} \sum_k \lambda_k \beta^T S_k \beta$$

Paramètres de la modélisation

✓ *La fonction de lissage*

Un lisseur est un outil qui permet de résumer la tendance de la variable réponse Y en une fonction d'une ou plusieurs variables prédictives. Cela produit une estimation plus stable de cette tendance. Une propriété importante du lisseur est sa nature non paramétrique : il ne suppose pas une forme rigide et prédéfinie pour la dépendance de Y en fonction des variables explicatives. Dans cette étude, nous nous concentrons uniquement sur la procédure de lissage nommée, en anglais, *cubic smoothing spline*.

✓ *Le choix du paramètre de lissage*

Chaque fonction de lissage n'a qu'un seul paramètre de lissage λ_k . Chaque paramètre gouverne le niveau de lissage ainsi que le compromis entre le biais et la variance. Quand les paramètres de lissage tendent vers 0, il n'y a pas de pénalité de lissage et le modèle fournit un ajustement parfait : les valeurs ajustées sont les données elles-mêmes. Ce qui fait diminuer le biais et implique que la variance augmente. À l'inverse, quand ils sont grands (tendent vers l'infini), les valeurs

² *Penalized iteratively re-weighted least squares.*

ajustées tombent le long d'une ligne droite forçant ainsi la liaison à être linéaire. Il en résulte une faible variance, mais un biais important. Les paramètres de lissage doivent être compris entre ces deux cas extrêmes pour produire le niveau désiré de lissage qui permet d'avoir un bon ajustement du modèle avec un équilibre biais - variance. Le choix de ces paramètres de lissage peut se faire à l'aide de la méthode par GCV.

3. Modèle de Cox

Le modèle de Cox (1972) en temps continu, connu aussi sous le nom de modèle semi-paramétrique à risques proportionnels, est un modèle d'analyse biographique (ou survie) du type régression. Il exprime le risque instantané (*hazard rate*, en anglais) de connaître l'événement étudié après une durée d'exposition donnée en fonction d'une combinaison linéaire de facteurs explicatifs.

Ce modèle permet de quantifier et de tester les effets propres de caractéristiques individuelles telles que genre, classe sociale, etc., sur le risque de transition. Il prend la formule suivante :

$$h(t/ Z) = h_0(t) \exp(Z' \beta)$$

où $h(t)$ est le risque instantané de survenue de l'événement;

$h_0(t)$ est le risque de base, non spécifié de façon paramétrique;

Z est la matrice des variables explicatives pouvant dépendre du temps;

β est le vecteur des coefficients de régression des variables explicatives.

Estimation des paramètres

Étant donné un échantillon $(T_j, \delta_j, Z_j)_{1 \leq j \leq n}$ où $T_j = \inf(X_j, C_j)$ avec

- C_j : la censure droite et X_j la durée de vie du $j^{\text{ème}}$ individu;
- δ_j : l'indicatrice de l'évènement;

- Z_j : le vecteur des covariables pour le $j^{\text{ème}}$ individu.

Soit

- t_1, \dots, t_n : l'échantillon observé dans lequel on a D événements et n-D censures;
- $0 < X_1^* < X_2^* < \dots < X_D^*$: l'échantillon ordonné des D instants d'évènements, supposés distincts;
- Z_j^* : le profil ou vecteur de réalisation des variables, associé à l'individu qui décède en X_j^* ;
- $\mathfrak{R}(X_j^*)$: l'ensemble des individus à risque à X_j^* i.e., $\mathfrak{R}(X_j^*) = \{i : X_i \geq X_j^*\}$;
- $\varepsilon(X_j^*)$: l'évènement pour l'individu X_j^* appartenant à $\mathfrak{R}(X_j^*)$;
- $\varepsilon(X_i^*)$: l'évènement pour l'individu à profil Z_j^* .

Soit $L_j(\beta)$ la contribution multiplicative à la vraisemblance, de l'observation (X_j^*, Z_j^*) c.-à-d.,

$$L_j(\beta) = P(\varepsilon(Z_j^*) / \varepsilon(X_j^*))$$

S'agissant d'un modèle de Cox, on obtient

$$L_j(\beta) = \frac{\exp(\beta' Z_j^*)}{\sum_{i \in \mathfrak{R}(X_j^*)} \exp(\beta' Z_i^*)}$$

Par la suite, la vraisemblance partielle de Cox est donnée par

$$L(\beta) = \prod_{j=1}^D L_j(\beta) = \prod_{j=1}^D \frac{\exp(\beta' Z_j^*)}{\sum_{i \in \mathfrak{R}(X_j^*)} \exp(\beta' Z_i^*)}$$

La recherche de l'estimateur de β est faite à partir du log-vraisemblance partielle, en résolvant un système d'équations non linéaires par des procédures itératives comme l'algorithme de Newton-Raphson.