

#### **4. ANALYSE BAYESIENNE DE LA PART THERMOSENSIBLE DE LA CONSOMMATION ELECTRIQUE REGIONALE**

Le but de ce chapitre est d'évaluer de l'intérêt de la méthode bayésienne pour combiner les approches top-down et bottom-up pour l'analyse de la part thermosensible de la demande d'électricité. Cette section présente les outils élémentaires nécessaires au développement et à la mise en œuvre de l'inférence bayésienne. Dans un premier temps, on présente le principe de l'inférence statistique bayésienne. Ensuite, on aborde les différents aspects techniques tels que les modèles statistiques utilisés et les algorithmes de simulations permettant l'estimation des paramètres. Finalement, on présente les analyses bayésiennes adaptées à l'échelle régionale.

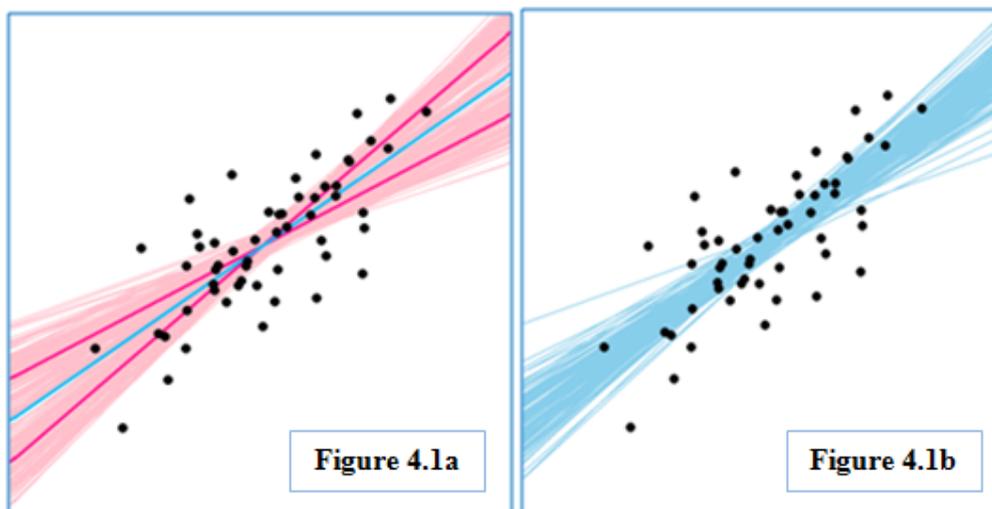
##### **4.1 Pourquoi adopter une approche statistique Bayésienne ?**

L'un des éléments permettant de distinguer les statistiques fréquentiste et bayésienne est la direction de l'inférence. L'approche fréquentiste est déductive. Elle part de l'hypothèse et la confronte aux données, tandis que l'approche bayésienne est inductive. Elle part des données pour estimer la distribution du paramètre inconnu. [136].

Rivot (2003) [137] a mentionné que le cadre statistique bayésien se distingue du cadre fréquentiste par deux différences conceptuelles fondamentales. La première différence réside dans l'interprétation de la notion de probabilité : dans le cadre bayésien, la probabilité exprime un degré de crédibilité, tandis que dans le cadre fréquentiste, la probabilité est une fréquence limite d'une infinité de réalisations. La seconde différence est une conséquence de la première : le cadre bayésien considère les paramètres comme des variables aléatoires et leur assigne une distribution de probabilité qui reflète le degré de crédibilité des différentes valeurs. De son côté, le cadre fréquentiste considère les paramètres comme des grandeurs fixes pour lesquelles la notion même de distribution de probabilité n'a pas de sens. Lecoutre (2006) [138] a expliqué que dans l'inférence fréquentiste toutes les probabilités sont

conditionnelles aux paramètres qui sont supposés connus. Cela conduit en particulier aux tests de signification dans lesquels la valeur d'au moins un paramètre est fixée par hypothèse et aux intervalles de confiance. Mais les paramètres ne peuvent pas être probabilisés (des fréquences empiriques ne sont pas disponibles). Dans l'inférence bayésienne, au contraire, on probabilise explicitement les valeurs possibles du paramètre. A partir d'un état de connaissance initial formalisé par une distribution *a priori*, et des données, nous obtenons par la formule de Bayes (4.2.1 ci-après) une distribution *a posteriori* qui exprime directement notre incertitude sur le paramètre, conditionnellement à l'échantillon observé. La distribution *a posteriori* combine l'information initiale avec l'information apportée par les données.

Prenons comme exemple la régression linéaire. Dans la Figure 4.1, les données sont représentées par des points noirs. Dans l'estimation fréquentiste (Figure 4.1.a), la ligne bleue foncée représente le meilleur ajustement. Les deux lignes roses foncées indiquent les limites de l'intervalle de confiance. Les lignes roses claires montrent les distributions d'échantillonnage autour de chacune des limites de l'intervalle de confiance. Il faut noter que la ligne du meilleur ajustement tombe à l'extrémité de chaque distribution d'échantillonnage. Il n'y a pas de distribution de probabilités sur les possibilités, il y a seulement les valeurs de trois points: le meilleur ajustement et les deux limites de l'intervalle de confiance [139].



Source: John K. Kruschke (2014) [139]

**Figure 4.1 :** Estimation fréquentiste et estimation bayésienne.

La Figure 4.1.b montre l'estimation bayésienne. L'élément essentiel de l'analyse bayésienne est d'inférer l'incertitude (crédibilité relative) de paramètres dans un espace modèle, compte tenu des données. Les lignes bleues claires indiquent la distribution *a posteriori* des estimations possibles. Ce graphique représente l'incertitude par différentes pentes et interceptes, avec une plus grande densité au milieu des lignes. On voit que la régression bayésienne révèle plusieurs possibilités crédibles au lieu d'une seule « meilleure » ligne [139].

Les résultats d'une analyse bayésienne sont des degrés de crédibilité associés à chaque hypothèse. Elle rend plus facile de représenter et d'expliquer les incertitudes liées aux modèles et aux valeurs de paramètres. Elle donne la possibilité de rassembler et de fusionner des connaissances de diverses natures dans un même modèle : retour d'expérience (données historiques ou empiriques), expertise (exprimée sous forme de règles logiques, d'équations, de statistiques ou de probabilités subjectives), observations. Cependant, la critique la plus souvent faite à l'approche bayésienne est sa subjectivité [140-141]. L'explicitation des opinions *a priori* d'experts du domaine peut être utile dans certaines études, mais il faut insister sur le fait que cela nécessite des techniques appropriées [142].

Rouanet et al. (1986) [143] ont proposé une certaine approche bayésienne qui permet de dépasser la subjectivité de l'approche bayésienne et d'apporter une réponse satisfaisante. Ils ont mentionné que cette approche bayésienne illustrée notamment par Jeffreys [144] depuis les années trente consiste à utiliser des distributions initiales visant à représenter un état d'ignorance sur les paramètres. Ces distributions, souvent qualifiées de « non-informatives », sont des distributions très diffuses (4.2.3 ci-après). Du point de vue technique, les distributions *a posteriori* correspondantes peuvent être dérivées par des algorithmes fondés exclusivement sur le modèle d'échantillonnage et les données. Elles peuvent donc être interprétées comme l'apport propre aux données et fournissent ainsi une référence objective pour l'interprétation et la communication des résultats. Cette approche de l'inférence bayésienne est maintenant reconnue comme un standard. Elle permet des propositions concrètes pour répondre parfois aux insuffisances des procédures fréquentistes.

Efron (1986) [145] remarque que « les méthodes bayésiennes ont rarement été appliquées car les probabilités critiques des tests fréquentistes sont beaucoup plus

faciles à calculer que les distributions de probabilité *a posteriori* ». Le développement récent des méthodes numériques a permis de contourner ces difficultés. Cela a permis le développement des analyses bayésiennes dans de nombreux domaines de la statistique appliquée. Cette méthode est devenue un outil important pour l'analyse de données et la présentation de résultats scientifiques. Les avantages de l'approche bayésienne pour les analyses statistiques dans le domaine énergétique ont été mis en évidence par certains auteurs [146-150]. A notre connaissance, il n'existe pas de méthode bayésienne pour estimer la part thermosensible de la demande d'électricité et la part des usages thermiques non-linéaires.

La quantification exacte de la thermosensibilité nécessite des données détaillées sur le parc des bâtiments, sur les équipements installés et sur le comportement des usagers. C'est la raison pour laquelle elle est difficile à quantifier. La possibilité de se baser sur des campagnes de mesures est limitée pour des raisons économiques. Les paramètres qui influencent la partie thermosensible de la consommation sont variables en fonction du secteur (résidentiel ou tertiaire), du type de bâtiments, du dimensionnement et de la performance des équipements de chauffage. Aujourd'hui, même si on a des données assez détaillées sur le parc de bâtiments, on ne dispose pas de données sur le dimensionnement réel des équipements dans différents secteurs. L'incertitude traduit essentiellement les limites des connaissances. Pour prendre en compte la variabilité des paramètres du modèle et l'incertitude associée, l'approche bayésienne semble être une des meilleures pistes. Elle propose une modélisation de l'incertitude des paramètres du modèle qui seront bien prises en compte au niveau des prévisions et fournissent des distributions prédictives au lieu d'une seule quantité ou d'un intervalle de confiance [151]. En partant des données connues et en adoptant le point de vue bayésien, on peut utiliser en plus des probabilités d'échantillonnage, des probabilités qui expriment notre incertitude sur toutes les valeurs possibles des paramètres. Pour ces raisons, on a choisi d'adopter la perspective bayésienne pour répondre de façon adéquate à la question posée.

## 4.2 Approche Bayésienne en Modélisation

Dans cette partie les concepts de base de la statistique bayésienne seront expliqués. La partie suivante est principalement basée sur le livre de Gelman et al. (2004) [152].

### 4.2.1 Théorème de Bayes

Un modèle statistique bayésien est basé sur le Théorème de Bayes pour caractériser des situations réelles à partir de données observées. L'inférence bayésienne sur un paramètre  $\theta$  se fait en terme de probabilités conditionnelles aux valeurs observées  $y$  notées  $p(\theta|y)$ .

Afin de formuler des probabilités de  $\theta$  en fonction de  $y$ , on doit spécifier un modèle avec une distribution de probabilité pour  $\theta$  et  $y$ . La fonction de densité de probabilité peut être écrite comme le produit de deux distributions souvent appelées la distribution *a priori*  $p(\theta)$  et la distribution d'échantillonnage, ou la distribution de données  $p(y|\theta)$  [152]:

$$p(\theta, y) = p(\theta).p(y | \theta) \quad (4.1)$$

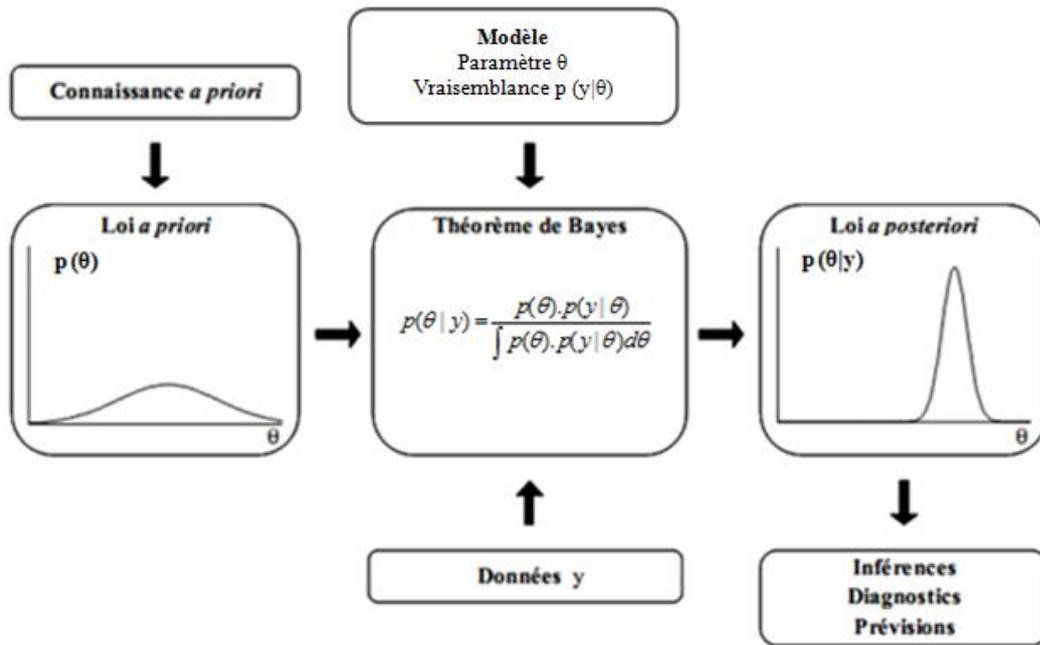
Le théorème de Bayes amène au calcul d'une distribution *a posteriori*  $p(\theta|y)$  par le biais de probabilités conditionnelles à des valeurs observées des données  $y$  :

$$p(\theta | y) = \frac{p(\theta).p(y | \theta)}{p(y)} \quad (4.2)$$

$p(y|\theta)$  est la distribution d'échantillonnage appelée aussi la fonction de vraisemblance à travers laquelle les données affectent la distribution *a posteriori*,  $p(\theta)$  est la distribution *a priori* et  $p(y)$  est la vraisemblance marginale.  $p(y)$  est la somme qui couvre toute les valeurs possibles de  $\theta$ . Puisque cette dernière ne dépend pas de  $\theta$ , elle est constante et le théorème se réduit à :

$$p(\theta | y) \propto p(\theta).p(y | \theta) \quad (4.3)$$

Dans la figure ci-dessous, les étapes fondamentales de l'analyse bayésienne sont représentées.



Source : Rivot (2003) [137]

**Figure 4.2 :** Les étapes fondamentales de l'analyse bayésienne.

Dans les cas les plus simples, tels que les modèles linéaires, l'expression analytique de la loi *a posteriori*  $p(\theta|y)$  peut être déterminée. Cependant, ces cas sont relativement rares en pratique, et il est souvent difficile d'obtenir une expression analytique pour  $p(\theta|y)$  en raison des termes d'intégration multidimensionnelle contenus dans les expressions de ces lois. Les méthodes Monte Carlo par Chaînes de Markov (MCMC) ont permis de lever un certain nombre de ces difficultés [153]. Ces méthodes sont basées sur l'approche de Monte Carlo (les paramètres suivent des distributions) ; les données observées (expérimentales) sont traitées par des chaînes de Markov pour caractériser une distribution optimale des paramètres. Cette approche a l'avantage de générer des distributions représentatives des données, à condition que celles-ci soient assez riches en information (voir 4.2.4 ci-après).

#### 4.2.2 Fonction de vraisemblance

La vraisemblance est souvent considérée comme une fonction des paramètres conditionnellement aux données. Mais formellement, elle exprime le degré de crédibilité des données conditionnellement aux paramètres. La combinaison de la loi *a priori* et de la vraisemblance permet de mettre à jour la distribution *a priori* pour obtenir la distribution *a posteriori*. Le calcul des probabilités garantit la cohérence

mathématique de la combinaison des sources d'information par la formule de Bayes [137].

La forme de la vraisemblance dépend d'une part du modèle déterministe mais aussi du modèle d'erreur sur les données. En général, le modèle est calibré à l'aide de données entachées d'erreur qu'il est nécessaire de prendre en compte. Pour définir la variance du terme d'erreur, deux approches sont possibles: soit la variance est fixée (à l'aide d'expériences précédentes par exemple), soit elle est estimée lors de la calibration du modèle [153].

#### 4.2.3 Distribution *a priori*

Le choix de la loi *a priori* représente une étape cruciale dans l'analyse statistique bayésienne, puisqu'elle influence directement le reste de l'inférence. On entend par information *a priori* sur le paramètre  $\theta$  toute information disponible sur  $\theta$  en dehors de celle apportée par les observations [154]. Son indépendance par rapport aux données est obligatoire afin d'éviter d'utiliser deux fois la même information.

Suivant la disponibilité de ces informations, on distingue deux types de distributions *a priori* : les distributions informatives et non-informatives [153]. La première repose sur l'information disponible sur le paramètre  $\theta$  obtenue dans la littérature ou par l'expertise. Des distributions informatives peuvent être utilisées pour les paramètres dont les valeurs possibles sont bien connues [153]. Le choix de la loi *a priori* comporte une certaine part d'arbitraire. Il convient par conséquent d'examiner dans quelle mesure le choix de la famille de probabilités retenue pour la loi *a priori* influence l'estimation bayésienne. Dans la pratique, on se contente d'examiner, dans quelle mesure, une petite perturbation des paramètres modifie l'estimation bayésienne. C'est ce qu'on appelle faire l'étude de sensibilité de la réponse bayésienne à la loi *a priori* [154].

Les distributions non-informatives (ou vagues) sont conçues dans le but de faire de l'analyse bayésienne en absence d'information *a priori* sur le paramètre  $\theta$  ou dans le cas où il est difficile de traduire en terme de loi *a priori* l'information disponible sur ce paramètre par une loi de probabilité [155]. Les distributions *a priori* non-informatives visent à représenter un "état d'ignorance" sur le paramètre  $\theta$ ; la distribution *a posteriori* correspondante pourra alors être interprétée comme résultant

de l'apport propre des données [156]. Lorsque la distribution *a priori* est non-informative, elle donne des probabilités identiques pour les différentes valeurs de  $\theta$ , comme par exemple une distribution de  $p(\theta)$  uniforme [136].

#### 4.2.4 Estimation des paramètres

##### 4.2.4.1 Méthodes de Monte Carlo par Chaînes de Markov

Dans cette partie, les principes de base de méthodes de Monte Carlo par Chaînes de Markov (MCMC) seront brièvement expliqués. Cette partie est principalement basée sur les documents de Rivot (2003) [137] et Peyret (2008) [157].

Une difficulté centrale de l'analyse bayésienne est le calcul des distributions *a posteriori*. Ces calculs nécessitent des intégrations sur de grandes dimensions [137] et il n'est pratiquement jamais possible de calculer de façon analytique. Les méthodes de simulation de Monte Carlo par Chaînes de Markov (MCMC) permettent de réaliser des inférences bayésiennes sur des modèles complexes et de générer des échantillons de la distribution *a posteriori* des paramètres du modèle.

Les méthodes de Monte Carlo sont des méthodes d'intégration numérique qui utilisent l'aléatoire. Rivot (2003) [137] explique que les méthodes de Monte Carlo consistent à générer par des simulations un échantillon  $(\theta^1, \dots, \theta^n)$  par tirage aléatoire dans la distribution de probabilité du vecteur des inconnues du modèle ( $\theta$ ). Si l'échantillon est suffisamment grand, l'histogramme des tirages obtenu est lisse et fournit alors une bonne approximation de la distribution de probabilité de  $\theta$ . Il souligne que l'avantage de ces méthodes est qu'il n'est pas nécessaire de connaître l'expression analytique de la distribution de  $\theta$  pour réaliser les simulations de Monte Carlo. On ne détaillera pas ici les méthodes de Monte-Carlo, les lecteurs intéressés peuvent se référer à Robert et Casella (2004) [158] ou à Marin et Robert (2007) [159].

Les méthodes MCMC sont un cas particulier des méthodes Monte Carlo et sont employées pour simuler la distribution *a posteriori*. La simulation MCMC est une méthode pour dresser des valeurs de  $\theta$  à partir de distributions approximatives et pour corriger ses valeurs afin de mieux faire l'approximation des distributions *a posteriori*  $p(\theta|y)$ . Des échantillons sont tirés séquentiellement des distributions fixées *a priori* à partir de la dernière valeur tirée, les suites de valeurs tirées forment une

chaîne de Markov. Ainsi une chaîne de Markov est une séquence de variables aléatoires  $\theta^1, \theta^2, \dots$ , pour laquelle, pour tout temps  $t$ , la distribution de  $\theta^t$  de tous les  $\theta$  précédents dépend seulement de la plus récente valeur ( $\theta^{t-1}$ ).

A chaque étape de la simulation, la distribution approximative est améliorée en convergeant vers la distribution cible. La simulation MCMC est utilisée quand on ne connaît pas la distribution  $p(\theta|y)$ . Elle utilise des distributions transitoires qui dépendent uniquement du tirage précédent ( $j-1$ ). Sa finalité est d'arriver à une distribution stationnaire se rapprochant le plus de  $p(\theta|y)$  [152].

Plusieurs algorithmes existent pour réaliser des simulations MCMC. L'échantillonneur de Gibbs et l'algorithme de Metropolis-Hastings sont deux algorithmes de simulation MCMC fréquemment utilisés. L'échantillonneur de Gibbs sera brièvement expliqué car dans toutes les applications présentées dans cette thèse, il a été mis en œuvre grâce au logiciel WinBUGS<sup>6</sup>.

#### 4.2.4.2 Échantillonneur de Gibbs

La méthode de l'échantillonneur de Gibbs repose sur la simulation successive suivant les diverses lois conditionnelles associées à la loi *a posteriori* cible. Cette décomposition en lois simples permet de faire face à des problèmes de simulation dans des espaces dont les dimensions sont arbitrairement grandes en les ramenant à de nombreux problèmes de petite dimension [160]. Les explications ci-dessous sont tirées du livre de Dodge et Melfi (2008) [161]. Pour plus de détails, le lecteur intéressé peut se référer aussi aux études de Casella et George (1992) [162] et Smith et Roberts (1993) [163].

L'échantillonneur de Gibbs est une manière de générer des distributions de deux (ou plusieurs) variables à partir d'un modèle qui définit les distributions de probabilité conditionnelles. Dans le cas d'un modèle à deux variables, la méthode consiste à prendre un élément de départ ( $x_0, y_0$ ) et à générer à l'aide de nombres aléatoires les éléments d'un échantillon fictif ( $x_n, y_n$ ) par itération en choisissant dans l'ordre  $x_n$

---

<sup>6</sup> WinBUGS est un logiciel statistique adapté à l'analyse bayésienne  
<http://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/>

d'une variable aléatoire de densité  $f(x|y)=y_{n-1}$  et  $y_n$  d'une variable aléatoire de densité  $g(y|x)=x_n$  où  $f(x|y)=y_{n-1}$  et  $g(y|x)=x_n$  sont les distributions de probabilité conditionnelles supposées connues ou modélisées. Pour le cas d'un modèle à plusieurs variables où les distributions conditionnelles seraient connues, l'élément au départ serait  $(x_1^0, x_2^0, \dots, x_k^0)$  l'échantillon est construit en simulant  $(x_1^n, x_2^n, \dots, x_k^n)$  en choisissant dans l'ordre :

$x_1^n$  d'une variable aléatoire de densité  $f(x_1 | x_2 = x_2^{n-2}, \dots, x_k = x_k^{n-k})$ ;

$x_2^n$  d'une variable aléatoire de densité  $f(x_2 | x_1 = x_1^n, x_3 = x_3^{n-1}, \dots, x_k = x_k^{n-k})$ ;

$x_3^n$  d'une variable aléatoire de densité  $f(x_3 | x_1 = x_1^n, x_2 = x_2^n, \dots, x_k = x_k^{n-k})$ ;

et ainsi de suite jusqu'à

$x_k^n$  d'une variable aléatoire de densité  $f(x_k | x_1 = x_1^n, \dots, x_{k-1} = x_{k-1}^n)$

L'échantillon ainsi généré sera une simulation conforme aux probabilités conditionnelles imposées et la distribution se conformera d'autant plus aux probabilités conditionnelles que le nombre d'itérations sera élevé. Un point est donc généré à partir du point précédent, sans utiliser l'information sur les autres points générés : il s'agit donc d'une chaîne de Markov [161].

L'utilisation intensive des méthodes de simulation MCMC est due à leur adaptabilité pour une vaste classe de problèmes et de modèles mais également à l'amélioration des technologies informatiques. L'une des limitations potentielles de l'utilisation des MCMC est le problème de la convergence de ces chaînes. En effet, il faut s'assurer qu'un nombre suffisant de générations de MCMC a été réalisé afin d'avoir une estimation la plus exacte possible de la distribution des probabilités postérieures [164].

#### 4.2.5 Contrôle de la convergence

Quand l'approche bayésienne est utilisée pour faire de l'inférence, il est important de s'assurer que les distributions postérieures obtenues correspondent bien à la distribution stationnaire cible. En d'autres termes la convergence, vers la distribution cible, doit être atteinte pour ne pas obtenir des résultats menant à de fausses interprétations [157].

En pratique, deux approches sont utilisées lors de la calibration d'un modèle : soit une seule chaîne MCMC est simulée pour un nombre d'itérations pour lequel la

convergence est supposée atteinte, soit plusieurs chaînes MCMC sont simulées à partir de points initiaux différents [153]. Peyret (2008) [157] note que selon Gelman et Rubin (1992) [165] le fait d'utiliser différentes chaînes suffisamment dispersées permet d'exploiter la variabilité de la distribution initiale et ensuite d'obtenir facilement une estimation de la variabilité d'échantillonnage sans avoir à faire de l'inférence sur la structure des séries temporelles de la simulation.

Le diagnostic de convergence le plus utilisé est celui de Gelman et Rubin (1992) [165] corrigé par Brooks et Gelman (1998) [166]. Ce diagnostic est un critère de mélange basé sur la comparaison entre plusieurs chaînes MCMC indépendantes, obtenues en lançant le processus d'échantillonnage à partir d'états initiaux différents, dispersés dans l'espace des paramètres. Ce diagnostic consiste à vérifier que différentes chaînes, initiées à différents endroits, vont effectivement "perdre la mémoire" de leur état initial et explorer de la même façon le domaine des paramètres. [167].

Brochet (2006) [153] explique que cette méthode est basée sur la comparaison de la variance inter-chaînes (entre les valeurs des différentes chaînes) et de la variance intra-chaîne (entre les valeurs d'une même chaîne), par l'intermédiaire d'un critère appelé facteur de réduction potentiel de variance  $\hat{R}$ . La comparaison s'effectue pour chacun des paramètres indépendamment. Les chaînes sont arrivées à convergence lorsque  $\hat{R}$  est égal à 1. Cependant, la valeur 1 n'est atteinte qu'asymptotiquement. En pratique, on considère que la convergence est probablement atteinte lorsque  $\hat{R}$  est inférieur à 1,2.

Au cours de ce travail, la convergence des algorithmes MCMC a été vérifiée par un examen visuel des chaînes MCMC complété par un calcul de la statistique de convergence de Gelman-Rubin.

### **4.3 Application de l'Approche Bayésienne**

Il s'agit de construire un cadre de modélisation cohérent et transposable à de nombreux cas d'étude, qui puisse permettre d'améliorer des modèles de la demande d'électricité avec des chroniques d'observations, afin d'assimiler l'information contenue dans les données et de quantifier l'incertitude.

On a utilisé l'approche bayésienne pour tester l'identification des composantes de la part thermosensible de la consommation d'électricité avec un modèle linéaire et un modèle non-linéaire. Un intérêt est qu'avec l'approche bayésienne, on va pouvoir limiter les risques d'obtenir des valeurs de paramètres peu réalistes lorsqu'on cherche à estimer des paramètres qui ont une signification physique.

Les paramètres du modèle sont estimés à pas horaire par rapport à la température extérieure. Une approche similaire à celle de Rouanet et al. [143] est adoptée pour identifier un modèle de la variation de la consommation d'électricité par rapport à la température en utilisant des données de la partie thermosensibles de la consommation (Qts\*) décrit précédemment dans le Chapitre 3.

On a identifié les signaux de la partie thermosensible de la consommation par l'approche bayésienne. On a analysé deux cas différents :

- Analyse de thermosensibilité actuelle sans chauffage non-linéaire
- Analyse de la thermosensibilité avec détermination de la pénétration actuelle des pompes à chaleur

Le logiciel WinBUGS est utilisé pour mener ces analyses bayésiennes. Ce logiciel de programmation a été développé à l'Université de Cambridge. Il utilise le programme BUGS : Bayesian Analysis using the Gibbs Sampler, qui prend en entrée un modèle composé de loi(s) *a priori* et de valeurs initiales, et retourne les sorties de l'algorithme de Gibbs pour cette loi *a posteriori* [168]. WinBUGS est très utilisé actuellement pour l'inférence bayésienne dans de nombreux domaines des statistiques appliquées.

#### **4.3.1 Analyse de thermosensibilité actuelle sans chauffage non-linéaire**

La première analyse est basée sur l'hypothèse que la part des usages thermosensibles dont la consommation est non-linéaire en termes de température est négligeable actuellement. Dans un premier temps, il faut décrire le système d'équations à résoudre, ainsi que les relations entre les différentes variables du modèle. Ensuite, il est nécessaire de spécifier la vraisemblance des données et les distributions *a priori* des paramètres.

### 4.3.1.1 Modèles

On a supposé que la partie thermosensible du modèle représente la demande de chauffage. On s'est basé sur les équations de la norme européenne EN 12831 (2004) [124] afin de décrire le besoin de chauffage. Selon cette norme, la charge thermique  $\Phi_{HL,i}$  pour un espace chauffé (i), se calcule comme suit :

$$\Phi_{HL,i} = \sum \Phi_{T,i} + \sum \Phi_{V,i} + \sum \Phi_{RH,i} \quad (4.4)$$

- $\Phi_{T,i}$  sont les déperditions par transmission de l'espace chauffé (i) [W]
- $\Phi_{V,i}$  sont les déperditions par renouvellement d'air de l'espace chauffé (i) [W]
- $\Phi_{RH,i}$  est la surpuissance de relance nécessaire dans l'espace chauffé (i) pour compenser les effets de l'intermittence du chauffage [W]

Les déperditions par transmission peuvent être décrites d'une manière simplifiée comme suit:

$$\Phi_{T,i} = \sum U_j \cdot A_j \cdot (T_{int,i} - T_{ext}) \quad [W] \quad (4.5)$$

- $U_j$  est le coefficient de transmission thermique de chaque paroi (j) ( $W/m^2K$ )
- $A_j$  est l'aire de chaque paroi (j) ( $m^2$ )
- $(T_{int,i} - T_{ext})$  est la différence entre la température intérieure et la température extérieure

Les déperditions par renouvellement d'air,  $\Phi_{V,i}$  pour un espace chauffé (i), se calculent comme suit :

$$\Phi_{V,i} = H_{V,i} \cdot (T_{int,i} - T_{ext}) \quad [W] \quad (4.6)$$

- $H_{V,i}$  sont les déperditions par renouvellement d'air [ $W/K$ ]
- $(T_{int,i} - T_{ext})$  est la différence entre la température intérieure et la température extérieure

Le but de cette thèse n'étant pas de traiter tous les aspects de la thermique du bâtiment, l'effet des ponts thermiques, les différents facteurs de réduction et la surpuissance de relance nécessaire  $\Phi_{RH,i}$  comme décrit dans la norme EN 12831 (2004) [124] n'ont pas été discutés.

Pour le calcul des déperditions de base d'un espace chauffé, on prend en compte les déperditions de base par transmission et les déperditions de base par renouvellement d'air.

$$\Phi_{HL} = U_{moyenne} \cdot A_{total} \cdot (T_{int} - T_{ext}) + H_V \cdot (T_{int} - T_{ext}) \quad (4.7)$$

$$\Phi_{HL} = (U_{moyenne} \cdot A_{total} + H_V) \cdot (T_{int} - T_{ext}) \quad (4.8)$$

Les déperditions sont compensées en partie par les apports gratuits ou non (internes et externes) AG qui participent directement au chauffage. Le besoin de chauffage est donc :

$$B = (U_{moyenne} \cdot A_{total} + H_V) \cdot (T_{int} - T_{ext}) - AG \quad (4.9)$$

Le problème est de connaître les différents apports gratuits. Les apports internes (occupants, éclairages, appareils ménagers et autres) sont fonction du nombre d'occupants, de leurs comportements et de leur mode de vie. Les apports externes sont principalement les apports solaires. Ces apports sont difficiles à quantifier en absence de données comme par exemple les données climatiques et les données comportementales. De plus, la température intérieure est liée aux apports gratuits. Dans un premier temps, on a essayé d'identifier AG mais les modèles décrits ci-dessous mais il n'a pas été possible de les identifier (Voir l'Annexe B). C'est pour cette raison que l'on a exclu AG de ces modèles.

On a construit trois modèles :

<b>Modèle 1a</b>	$Q_{ts} = K \cdot (T_i - T_{ext}) + \varepsilon$	Estimation par la température brute mesurée d'un coefficient global du besoin thermique
<b>Modèle 1b</b>	$Q_{ts} = K \cdot (T_i - T_{lis}) + \varepsilon$	Estimation par la température lissée d'un coefficient global de besoin thermique
<b>Modèle 1c</b>	$Q_{ts} = K \cdot (T_i - T_{lis}) + H \cdot (T_i - T_{ext}) + \varepsilon$	Prise en compte de la dynamique lente et de la dynamique rapide des bâtiments

Les Modèles 1a et 1b sont construits avec trois paramètres à identifier : un coefficient de perte de chaleur global K qui inclut les caractéristiques du parc de

bâtiments, la température intérieure ( $T_i$ ) et l'erreur ( $\varepsilon$ ). Si on suppose que cette erreur suit une loi Normale  $N(0,1)$ , le modèle s'écrit sous la forme suivante :

$$\begin{aligned} Q_{ts}^* &\sim N(Q_{ts}, \sigma^2) \\ \text{avec } Q_{ts} &= K(T_i - T_{ext}) \end{aligned} \quad (4.10)$$

où  $\sigma^2$  est la variance de l'erreur  $\varepsilon$ .

Nous supposons ainsi que la partie thermosensible de la consommation électrique ( $Q_{ts}^*$ ) est normalement distribuée autour de la valeur de  $Q_{ts}$  avec une variance  $\sigma^2$  inconnue. Dans WinBUGS, ce paramètre est estimé via la précision ( $\tau$ ), définie comme l'inverse de la variance ( $\sigma^2$ ).

Les variables observées du *Modèle 1a* sont la partie thermosensible de la consommation d'électricité ( $Q_{ts}$ ) et la température extérieure ( $T_{ext}$ ). La partie thermosensible du modèle est supposée représenter la demande de chauffage par effet Joule.

Le modèle a été calibré à l'aide des données de consommation ( $Q_{ts}$ ) issues du modèle Metehore et des données de température extérieure ( $T_{ext}$ ) pour le département des Alpes Maritimes (zone essentiellement résidentielle, pénétration des PAC allant jusqu'à 10%, signature des PAC visible en hiver). Ce sont les données horaires de l'hiver 2008-2009 (mi-Novembre, Décembre, Janvier, Février, mi-Mars) qui sont utilisées. Ces données correspondent aux jours ouvrables (lundi au vendredi hors jours fériés), en excluant les EJP.

Pour le *Modèle 1b*, au lieu d'utiliser les températures brutes, on a utilisé les températures lissées pour tenir compte de l'inertie des bâtiments. Les températures lissées sont calculées en utilisant le lissage exponentiel simple :

$$T_{lis(h,t)} = \alpha_h T_t + \alpha_h (1 - \alpha_h) T_{t-1} + \alpha_h (1 - \alpha_h)^2 T_{t-2} + \alpha_h (1 - \alpha_h)^3 T_{t-3} \dots \quad (4.11)$$

Le coefficient  $\alpha$ , compris entre 0 et 1, s'applique à la dernière réalisation. On a déterminé une constante de lissage pour chaque heure ( $\alpha_h$ ). Afin d'ajuster le modèle de lissage sur une série temporelle horaire continue, on a considéré l'ensemble des données horaires de la période hiver 2008-2009 (tous jours confondus, y compris les jours fériés, EJP et week-end). On a ensuite sélectionné pour chacune des 24 heures de la journée la valeur de  $\alpha$  qui minimisait la somme des carrés des erreurs. On a

ainsi déterminé que pour les heures allant de 19h jusqu'à 11h du matin, le meilleur coefficient était  $\alpha_h=0.05$  ; entre 12h et 15h,  $\alpha_h=0.35$  ; pour 16h,  $\alpha_h=0.2$  ; et pour 17h et 18h,  $\alpha_h=0.1$ .

Le *Modèle 1c* est construit avec quatre paramètres à identifier : le paramètre K qui représente la dynamique lente, le paramètre H qui représente la dynamique rapide, la température intérieure ( $T_i$ ) et l'erreur ( $\epsilon$ ).

$$Q_{ts}^* \sim Normal(Q_{ts}, \sigma^2)$$
$$Q_{ts} = K \cdot (T_i - T_{lis}) + H \cdot (T_i - T_{ext}) \quad (4.12)$$

La dynamique lente représente les déperditions à travers les parois des bâtiments et la dynamique rapide correspond plutôt aux déperditions par renouvellement d'air.

#### 4.3.1.2 Choix des distributions *a priori*

Comme expliqué dans les parties précédentes, le choix des distributions *a priori* est un aspect controversé des méthodes bayésiennes. Nous allons ici utiliser des lois *a priori* les plus générales et les moins informatives possibles, en l'occurrence des lois gamma et uniformes [169]. Les inconnues dans nos *Modèles 1a* et *1b* sont i) le coefficient de perte de chaleur (K) ii) la température intérieure ( $T_i$ ), iii) le terme d'erreur du modèle (la précision  $\tau$ ). Les lois *a priori* spécifiées dans le modèle sont diffuses (peu informatives). La loi *a priori* sur le paramètre de K (MW/°C) est choisie uniforme sur l'intervalle [0,100]. On a donné peu d'information sur le paramètre  $T_i$  avec une distribution normale centrée autour de 15°C et de variance 10. La précision  $\tau$  ( $\tau = 1/\sigma^2$ ) du modèle est estimée en mettant une loi uniforme sur l'écart type du modèle ( $\sigma$ ). C'est également l'approche adoptée par Gelman et Hill (2006) [170] et Jackman (2009) [171]. Une autre approche est d'utiliser une loi gamma non-informative. Dans les modèles BUGS<sup>7</sup>, les distributions *a priori* gamma pour des précisions ont tendance à causer plus de problèmes numériques que les distributions *a priori* uniformes pour les écarts-types. Une loi gamma qui est plate

---

<sup>7</sup> BUGS : Bayesian Analysis using the Gibbs Sampler

pour la plupart de sa distribution, devient très informative autour de zéro et cause des problèmes si la vraie valeur du paramètre de précision est proche de zéro [172]. Par conséquent, on a choisi d'utiliser une loi uniforme pour la précision sur l'intervalle [0,1000].

Pour le *Modèle 1c*, on a une inconnue supplémentaire  $H$  pour laquelle on a choisi une loi *a priori* uniforme sur l'intervalle [0,50]. Les lois *a priori* sur les autres paramètres restent identiques sauf pour le paramètre  $K$ . La distribution *a posteriori* de  $K$  déterminée lors des calages précédents est utilisée comme nouvelle distribution *a priori*. On a choisi une loi normale centrée autour de 35 MW/°C de variance 10.

Dans le schéma ci-dessous, on résume le *Modèle 1c* et ses différentes dépendances.

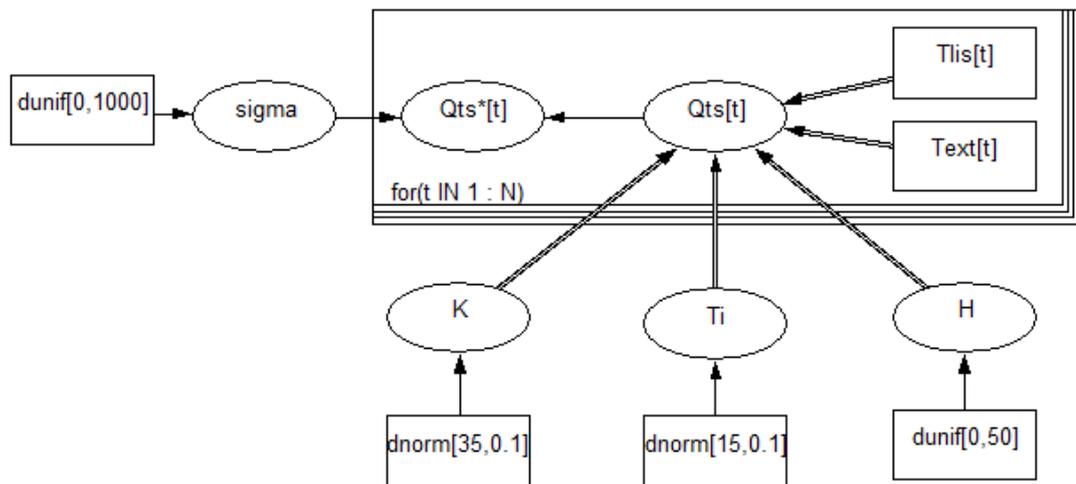


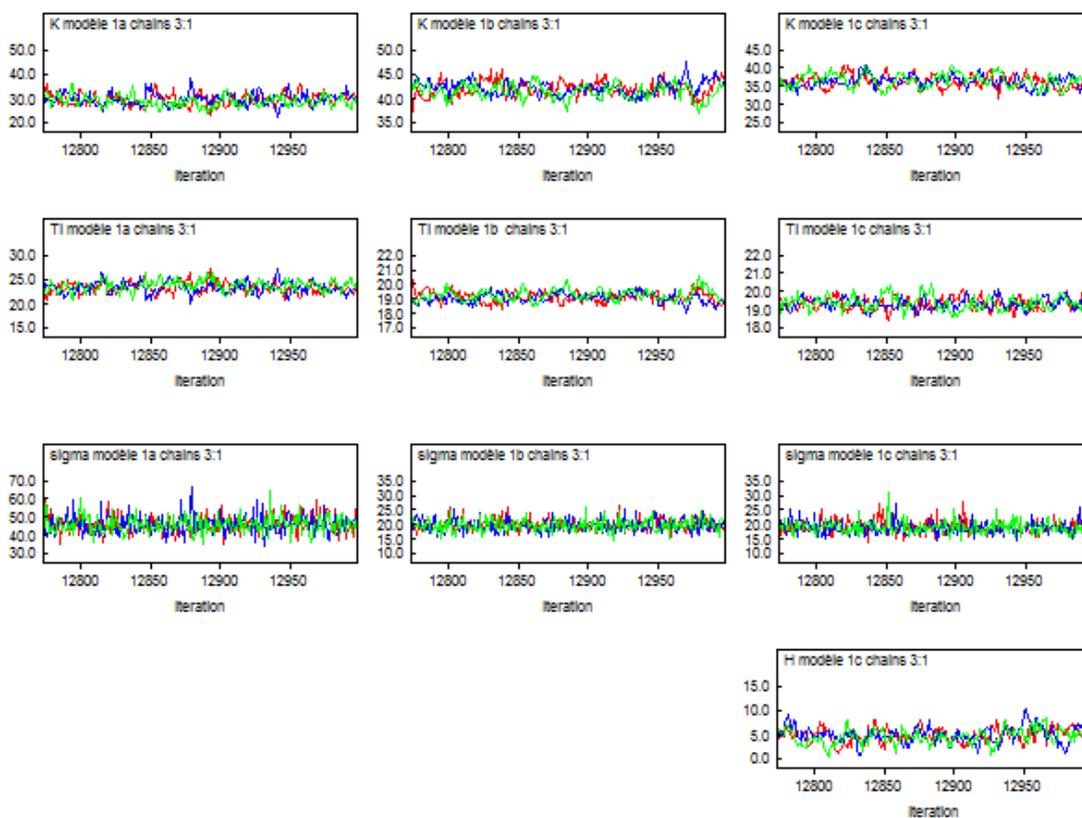
Figure 4.3 : Représentation du *Modèle 1c*.

#### 4.3.1.3 Analyse des résultats : distributions *a posteriori*

Pour chacun des paramètres, une distribution *a posteriori* a été obtenue. Tout d'abord, les paramètres de chaque modèle sont estimés pour chaque heure. Les simulations MCMC ont été conduites en lançant trois chaînes en parallèle, initialisées en trois points différents. 13 000 tirages ont été effectués pour chaque chaîne. Il a été considéré que les trois chaînes avaient perdu la mémoire de leur état initial après 3000 itérations de l'échantillonneur de Gibbs. Les 3000 premiers tirages de chaque chaîne n'ont pas été considérés pour les inférences. Les inférences ont été réalisées sur la base d'un échantillon de taille 30 000, formé par trois chaînes de taille 10 000. Ces trois chaînes indépendantes permettent de visualiser la

convergence du modèle et de vérifier la sensibilité des modèles aux valeurs initiales. Les autocorrélations et les convergences vers une distribution stationnaire ont ensuite été examinées. Les valeurs de la statistique de convergence de Gelman-Rubin ont ensuite été calculées dans WinBUGS.

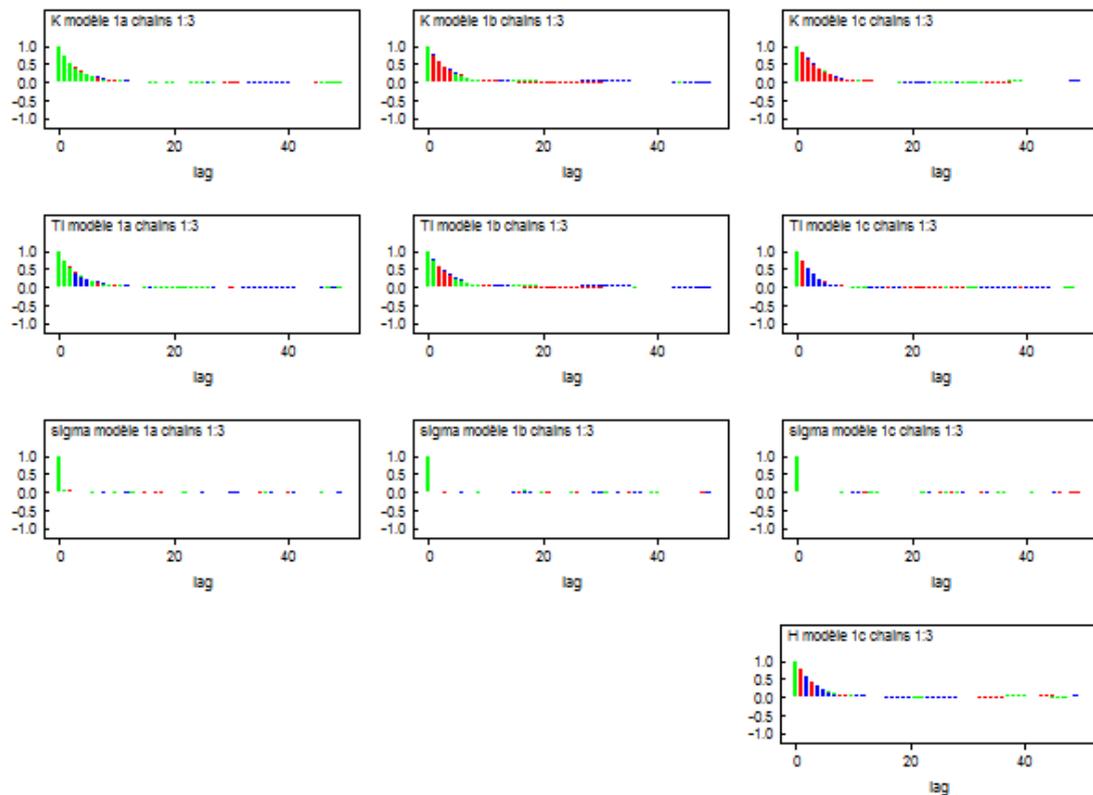
Sur la Figure 4.4, les résultats de l'analyse de 19h de différents modèles en utilisant le jeu de données du modèle RTE (Metehore) sont donnés à titre indicatif. On peut constater graphiquement que les chaînes sont bien mélangées et qu'il y a une certaine stationnarité des valeurs *a posteriori* tout au long des 10 000 itérations de la chaîne. Ceci indique que l'échantillon obtenu provient bien de la distribution souhaitée.



**Figure 4.4 :** 3 chaînes MCMC générées par l'échantillonneur de Gibbs pour les paramètres de chaque modèle pour 19h.

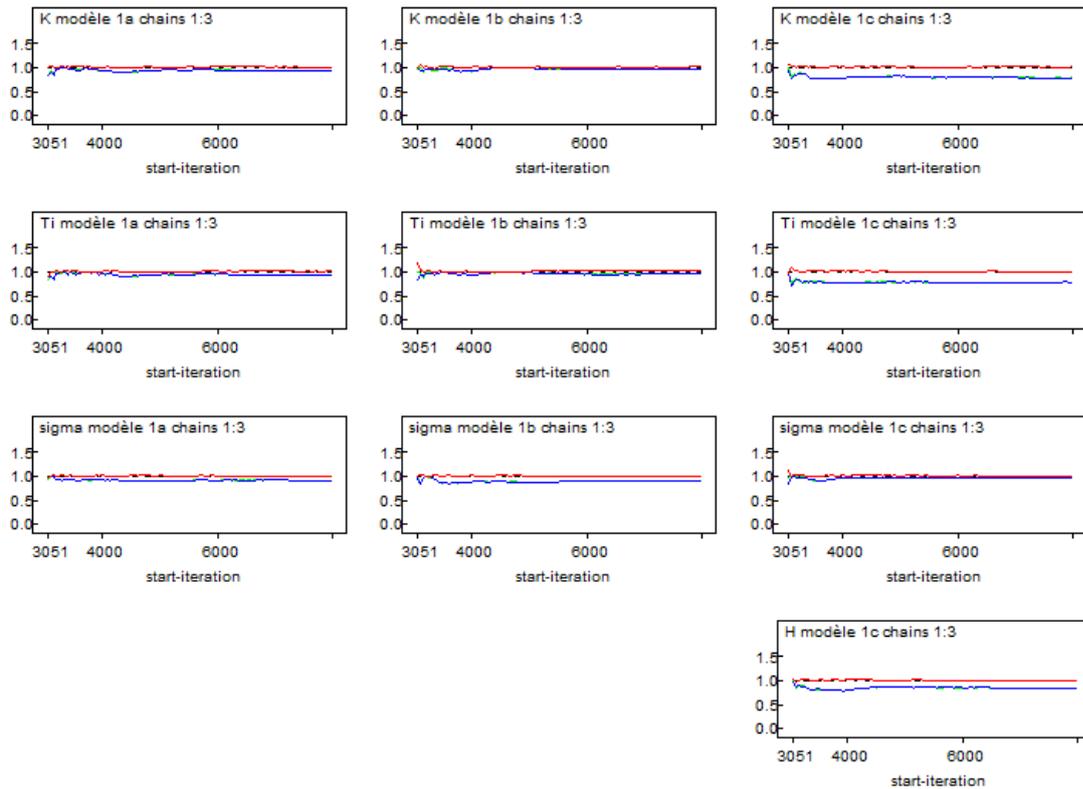
L'autocorrélation se réfère à un modèle de corrélation en série dans la chaîne, où les tirages séquentiels d'un paramètre à partir de la distribution conditionnelle sont corrélés. La cause de l'autocorrélation est que les paramètres du modèle peuvent être fortement corrélés, donc l'échantillonneur de Gibbs est lent à estimer la distribution *a posteriori*. WinBUGS permet de visualiser graphiquement le niveau d'autocorrélation jusqu'au rang  $n = 50$ . Le rang est appelé «lag» dans WinBUGS. En

règle générale, le niveau d'autocorrélation diminue avec un nombre croissant de rang dans la chaîne. Lorsque les autocorrélations sont élevées, cela indique que la chaîne de Markov est faiblement mélangeante. Une chaîne faiblement mélangeante peut impliquer une convergence lente de la chaîne. La Figure 4.5 montre que l'autocorrélation est relativement faible pour chacun des paramètres.



**Figure 4.5 :** Autocorrélation des paramètres  $K$ ,  $T_i$  et  $\sigma$  du modèle 19h.

Ensuite, il faut vérifier si les chaînes de Markov ont en effet atteint une distribution d'équilibre stable, c'est à dire, ont convergé. Le diagnostic de convergence Brooks-Gelman-Rubin (BGR) est utilisé. Les valeurs autour de 1 indiquent la convergence, la limite acceptable est considérée comme 1,1 par Gelman et Hill (2006) [170]. La Figure 4.6 montre l'évolution de la statistique de Brooks-Gelman-Rubin (BGR) au cours du temps. Selon ce diagnostic, les chaînes convergent vers une distribution stationnaire presque instantanément. On remarque ici qu'elle converge vers 1 pour les paramètres des 3 chaînes de Markov simulées. Ainsi donc, la convergence des chaînes est vérifiée.

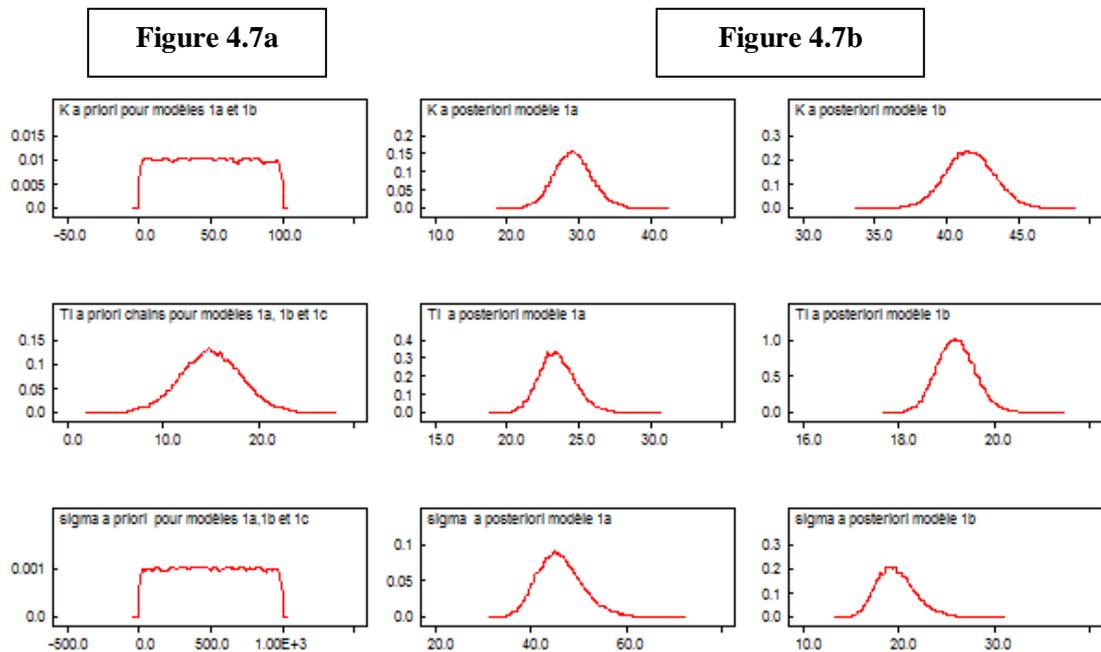


**Figure 4.6 :** Le statistique Brooks-Gelman-Rubin (BGR) des distributions *a posteriori* des trois modèles pour 19h.

La comparaison des distributions *a posteriori* et *a priori* permet d'estimer l'effet d'apprentissage du processus inférentiel. Les distributions *a posteriori* des paramètres peuvent être plus resserrées, cela veut dire que le processus inférentiel a permis d'améliorer la connaissance des paramètres du modèle. Quand on utilise les distributions *a priori* non-informatives et que les formes des deux distributions sont proches ça montre que les observations utilisées pour le calage ne contiennent pas ou peu d'informations sur ce paramètre. Les distributions peuvent présenter des changements de position et/ou de forme significatifs. Cette situation met en évidence une divergence entre ce que révèlent les observations et l'information *a priori* [173].

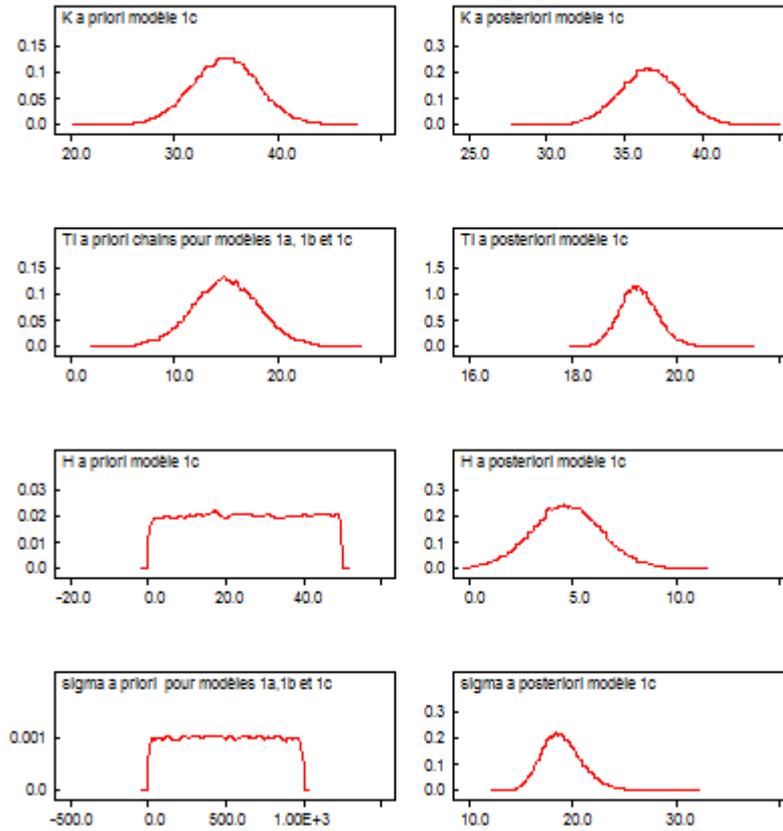
Dans la Figure 4.7, les histogrammes lissés de la distribution *a priori* et *a posteriori* des paramètres sont représentés. Chaque distribution *a posteriori* est moins diffuse que celle définie *a priori*, indiquant que les données ont pu informer sur la valeur du paramètre. Sur la Figure 4.7a, les distributions *a priori* des paramètres K et  $\sigma$  sont uniformes, alors que sur la Figure 4.7b, les distributions *a posteriori* ne sont plus uniformément dispersées, mais concentrées autour d'une valeur moyenne. L'incertitude sur ces paramètres a donc diminué. La distribution *a posteriori* du

paramètre  $T_i$  est moins diffuse que celle définie *a priori*, indiquant que les données ont pu informer sur la valeur du paramètre.



**Figure 4.7 :** Distribution *a priori* et distribution *a posteriori* des paramètres des Modèles 1a et 1b 19h.

Le paramètre  $K$  dans le *Modèle 1c* a une distribution *a priori* plutôt informative. Si la connaissance *a priori* sur la valeur de ce paramètre est informative, la calibration devrait donner une distribution *a posteriori* proche de celle *a priori*. On voit dans la Figure 4.8 que c'est bien le cas. L'incertitude sur le paramètre  $H$  et  $\sigma$  a diminué ce qui montre que les données ont pu informer sur la valeur de ce paramètre. La distribution *a posteriori* du paramètre  $T_i$  est aussi moins diffuse que celle définie *a priori*.



**Figure 4.8 :** Distribution *a priori* et distribution *a posteriori* des paramètres du *Modèle 1c* pour 19h.

Une fois les distributions *a posteriori* des paramètres obtenues, les statistiques descriptives (moyenne, écart type, quantiles, intervalles de crédibilité sous forme d'intervalles inter-quantiles) de chaque paramètre du modèle peuvent être extraites (Tableau 4.1).

**Tableau 4.1 :** Statistiques descriptives des paramètres des modèles pour 19h.

Modèle 1a	Moyenne	Ecart-type	Erreur MC	2,50%	Médiane	97,50%
Ti	23,52	1,254	0,01868	21,23	23,45	26,17
K	29,25	2,614	0,03856	24,38	29,17	34,68
sigma	46,14	4,716	0,03064	38,06	45,76	56,45
Modèle 1b	Moyenne	Ecart-type	Erreur MC	2,50%	Médiane	97,50%
Ti	19,21	0,3984	0,006181	18,47	19,19	20,03
K	41,63	1,684	0,02617	38,35	41,62	44,93
sigma	19,74	2,01	0,01308	16,33	19,56	24,2
Modèle 1c	Moyenne	Ecart-type	Erreur MC	2,5%	Médiane	97,5%
Ti	19,28	0,3603	0,004403	18,62	19,26	20,04
K	36,65	1,886	0,03175	32,98	36,64	40,34
H	4,728	1,654	0,02533	1,484	4,711	8,015
sigma	18,99	1,932	0,01143	15,66	18,83	23,17

On constate que les écart-types des estimations sont faibles. Le fait d'avoir un écart-type faible pour chaque estimation engendre un intervalle plus restreint pour chacun des paramètres, ce qui signifie que l'estimation est plus précise. L'estimation de l'erreur de Monte Carlo permet aussi de vérifier la convergence des algorithmes de MCMC. En pratique, lorsque cette erreur est inférieure à 5% de l'écart type estimé du paramètre, il est considéré que l'algorithme converge. Dans le Tableau 4.2, on montre que les erreurs MC sont inférieures à la limite considérée.

**Tableau 4.2 :** Erreur MC et convergence.

Modèle 1a	Ecart-type	5% de l'écart type	Erreur MC
<b>Ti</b>	1,254	0,0627	0,01868
<b>K</b>	2,614	0,1307	0,03856
<b>sigma</b>	4,716	0,2358	0,03064
Modèle 1b	Ecart-type	5% de l'écart type	Erreur MC
<b>Ti</b>	0,3984	0,01992	0,006181
<b>K</b>	1,684	0,0842	0,02617
<b>sigma</b>	2,01	0,1005	0,01308
Modèle 1c	Ecart-type	5% de l'écart type	Erreur MC
<b>Ti</b>	0,3603	0,018015	0,004403
<b>K</b>	1,886	0,0943	0,03175
<b>H</b>	1,654	0,0827	0,02533
<b>sigma</b>	1,932	0,0966	0,01143

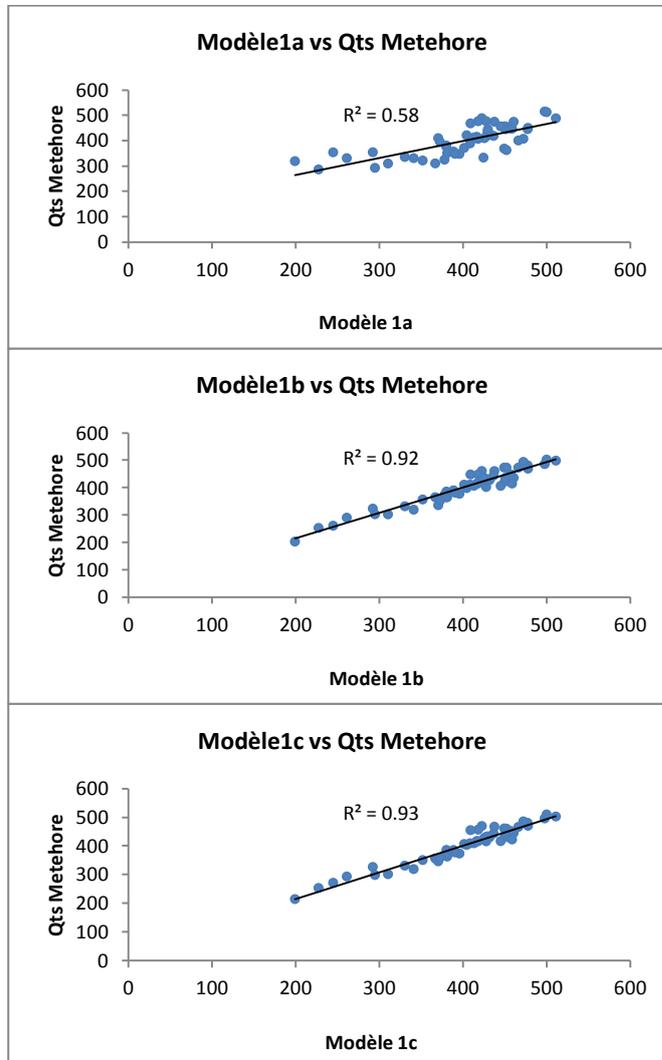
L'intervalle crédible à 95% est utilisé comme indicateur de dispersion. Cet intervalle est défini comme l'intervalle entre le percentile 2.5% et le percentile 97.5% de la distribution *a posteriori*, qui contient donc 95% des valeurs *a posteriori*.

La convergence des estimations des trois modèles pour 24 heures est vérifiée. On ne peut pas ici rapporter les résultats correspondants pour chacun des 3 modèles et chacune des 24 heures. On fait donc le choix de présenter les résultats pour le Modèle 1c, qui est le plus complexe des trois. Les résultats correspondant à ce modèle pour 01h, 07h, 13h sont donnés dans l'Annexe C.

#### 4.3.1.4 Vérification de l'ajustement des modèles

Après la calibration du modèle, il est nécessaire de vérifier son ajustement aux données. Plusieurs représentations graphiques permettent de visualiser les différences entre les prédictions du modèle et les données. La plus répandue superpose les données à la courbe d'ajustement du modèle. La Figure 4.9 présente les données  $Q_{ts}$

ainsi que l'estimation du modèle. On voit que le *Modèle 1a* qui utilise les températures brutes n'ajuste pas bien les données. Les *modèles 1b* et *1c* offrent un meilleur ajustement.



**Figure 4.9 :** Ajustement des modèles aux données.

Afin de valider l'analyse bayésienne et de vérifier si les informations *a priori* sont en adéquation avec les données, nous avons besoin d'un certain critère. En statistique bayésienne, l'ajustement d'un modèle peut être testé en générant des nouvelles données du modèle et en les comparant aux données observées. Si le modèle est bon, alors les données simulées et les données observées seront similaires. Si le modèle est mauvais alors il y aura des différences systématiques entre les données observées et simulées. La « Bayesian p-value » tient compte de la vraisemblance des résultats et des valeurs échantillonnées à partir de la distribution *a posteriori* du modèle. Elle correspond à la différence entre l'écart des valeurs observées (mesure du degré de

différence entre les valeurs observées et les valeurs estimées par le modèle) et l'écart des valeurs échantillonnées (mesure du degré de différence entre les valeurs échantillonnées et les valeurs estimées par le modèle). Un bon modèle aura une « Bayesian p-value », de préférence, proche de 0,5 [174-175].

$$p - value = 2 \sum_i r_i \ln \frac{r_i}{r_i'} - 2 \sum_i r_i^* \ln \frac{r_i^*}{r_i'} \quad (4.13)$$

avec

- $r_i$  : valeurs observées
- $r_i'$  : valeurs calculées partir de la distribution a posteriori du modèle
- $r_i^*$  : valeurs échantillonnées partir de la distribution a posteriori du modèle chaque itération

Le deuxième paramètre permettant de vérifier l'adéquation du modèle est le critère d'information de déviance (DIC pour Deviance Information Criterion). Le DIC est un critère de sélection de modèle en vertu duquel le rendement d'un modèle est évalué comme la somme d'une mesure d'ajustement (la moyenne *a posteriori* de la déviance D) et d'une mesure de complexité correspondant à la différence entre D et la déviance évaluée à la moyenne *a posteriori* du paramètre. Le modèle présentant la meilleure adéquation entre les données et les informations *a priori* sera celui qui possède le DIC le plus faible et positif. En d'autres termes, la vraisemblance doit être la plus élevée possible [174-175].

**Tableau 4.3** : Bayesian p-value et DIC des *Modèles 1a, 1b et 1c*.

	Bayesian p-value	DIC
<b>Modèle1a</b>	0,525	568,115
<b>Modèle1b</b>	0,5235	476,646
<b>Modèle1c</b>	0,5272	473,075

Dans le Tableau 4.3, les valeurs de Bayesian p-value et de DIC correspondant aux trois modèles sont données. Tous les modèles ont une valeur de p autour de 0,5 qui montre que les modèles sont bons. Le *Modèle 1c* ayant le DIC le plus faible présente la meilleure adéquation.

#### 4.3.1.5 Discussion des résultats des modèles sans chauffage non-linéaire

Dans le graphe ci-dessous les températures intérieures estimées pour chaque heure avec trois modèles sont représentées. Ces valeurs sont les valeurs de la mode de chaque distribution *a posteriori*.

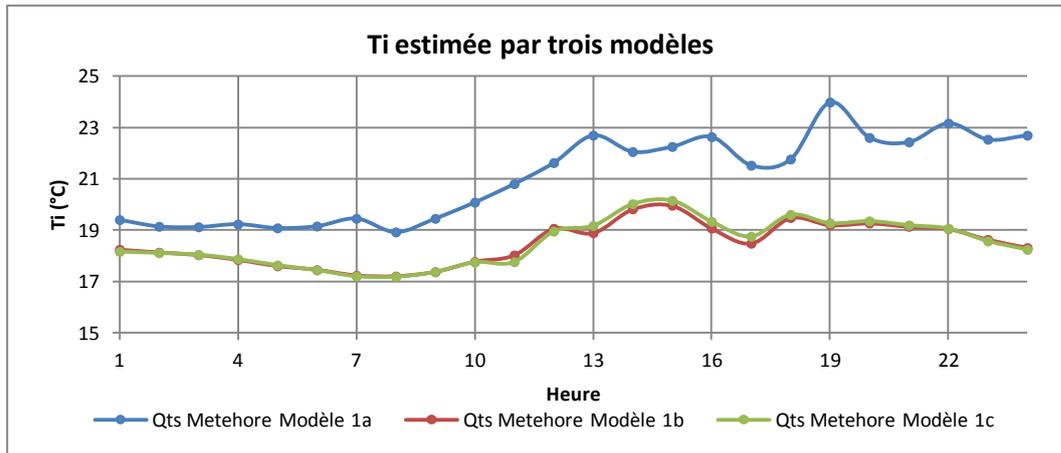


Figure 4.10 : Température intérieure estimée par trois modèles.

On observe que la température intérieure diminue pendant la nuit et augmente pendant la journée. Les estimations du *Modèle 1a* sont plus élevées que les *modèles 1b* et *1c*. Ceci est probablement dû à l'utilisation des températures brutes, les effets liés à l'inertie du bâtiment sont ignorés par ce modèle. La fluctuation de la température est beaucoup plus prononcée. Pour les *modèles 1b* et *1c*, la température oscille entre 17°C et 20°C ce qui est en cohérence avec l'étude de l'ADEME [176] et les résultats de l'enquête Sofres [121].

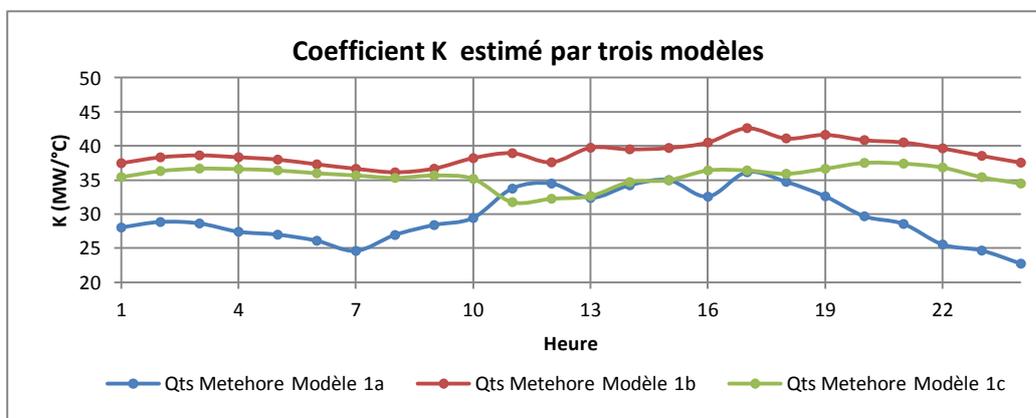
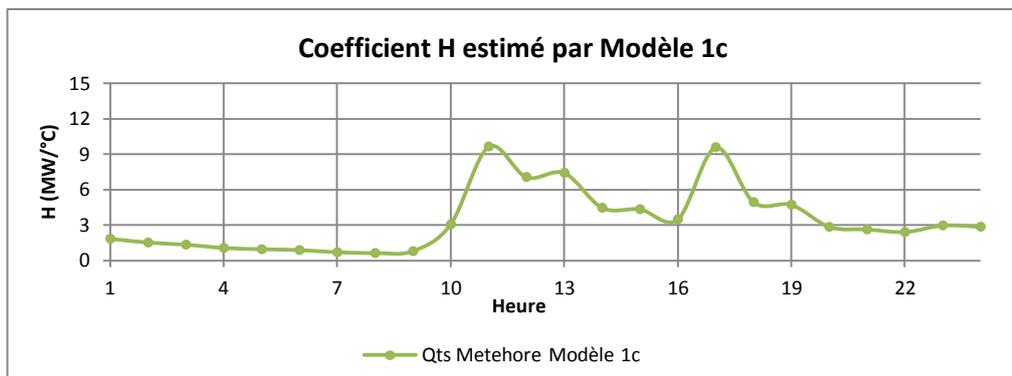


Figure 4.11 : Coefficient K estimée par trois modèles.

Dans la Figure 4.11, les valeurs estimées du paramètre K sont représentées. Les estimations réalisées en utilisant les températures brutes causent une fluctuation assez importante tout au long de la journée. On pourrait attendre une petite fluctuation due au fait que même si les caractéristiques physiques du parc de bâtiments n'évoluent pas la surface chauffée change pendant la journée. On pense que l'utilisation de la température brute et les apports gratuits qui ne sont pas explicitement modélisés perturbent les résultats.

Le *Modèle 1b* et *1c* donnent des résultats assez similaires pour les heures de la nuit, on voit que K a une variation plus petite dans un intervalle de [36-43] pour le *Modèle 1b* et [32-38] pour le *Modèle 1c*. On pense que l'inertie des bâtiments est mieux captée grâce à l'utilisation de la température lissée.

Les estimations horaires du paramètre H du *Modèle 1c* sont données dans la Figure 4.12. Le paramètre H qui représente l'inertie rapide de la consommation (e.g. la ventilation, l'activité humaine).



**Figure 4.12 :** Coefficient H estimée par *Modèle 1c*.

On observe que H est quasi-stable pendant la nuit avec une légère diminution. On pourrait penser qu'il n'y aura pas de changement comportemental pendant la nuit. Il semble logique que la valeur de H augmente à partir de 9h du matin en raison de la ventilation des bâtiments (plutôt tertiaires) et aussi du comportement humain. On constate une fluctuation entre 0.6 et 9.6. Le fait de combiner les observations et les a priori a permis d'utiliser toutes les informations disponibles et d'éviter d'avoir des valeurs irréalistes des paramètres.

Afin de comparer les erreurs des modèles, l'erreur moyenne absolue en pourcentage (MAPE) de trois modèles est représentée dans le graphique ci-dessous.

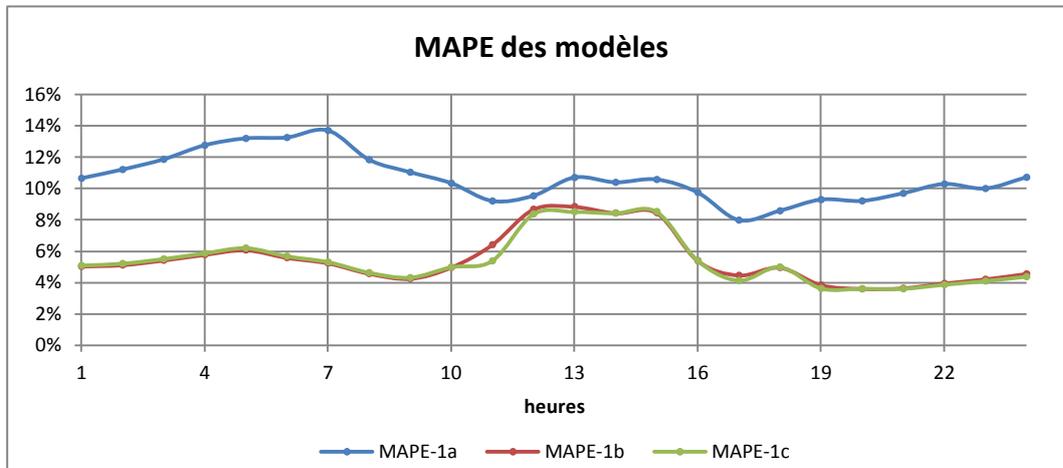


Figure 4.13 : Erreur moyenne absolue en pourcentage (MAPE) des Modèles 1a, 1b et 1c.

L'erreur du Modèle 1a est plus élevée que celles des autres modèles. Les erreurs des Modèles 1b et 1c sont très proches avec une légère différence entre 10h et 13h. La moyenne de la MAPE est de 10,6%, 5,6% et 5,5% pour le Modèle 1a, 1b et 1c respectivement.

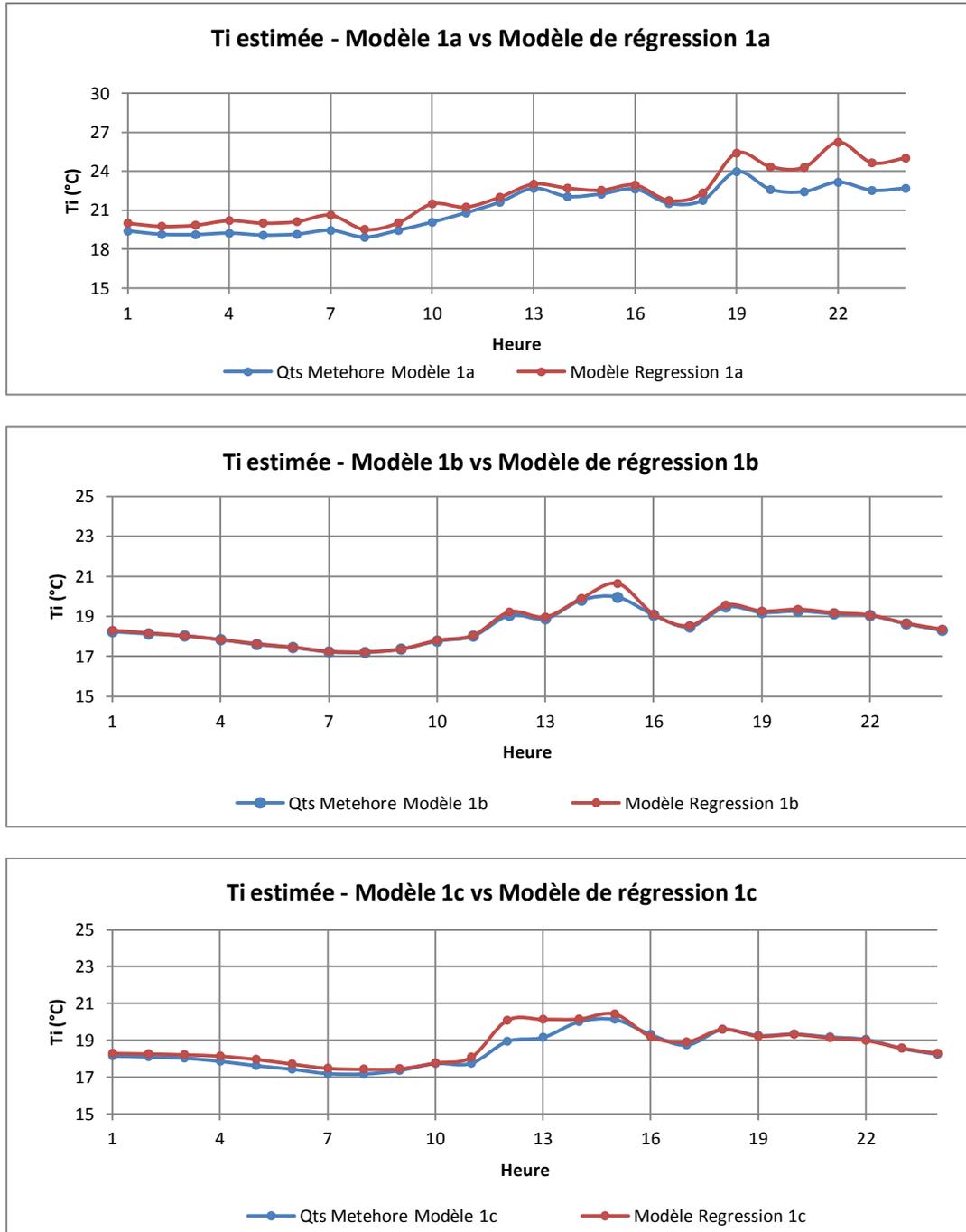
#### 4.3.1.6 Comparaison des résultats avec les modèles de régression

Par la suite on a comparé les résultats de l'estimation bayésienne avec les modèles de régression linéaire. Dans le tableau ci-dessous, les erreurs moyennes absolues en pourcentage (MAPE) des estimations sont présentées. On a trouvé des erreurs très proches pour les deux approches.

**Tableau 4.4 :** Comparaison des MAPE de l'estimation bayésienne et de la régression.

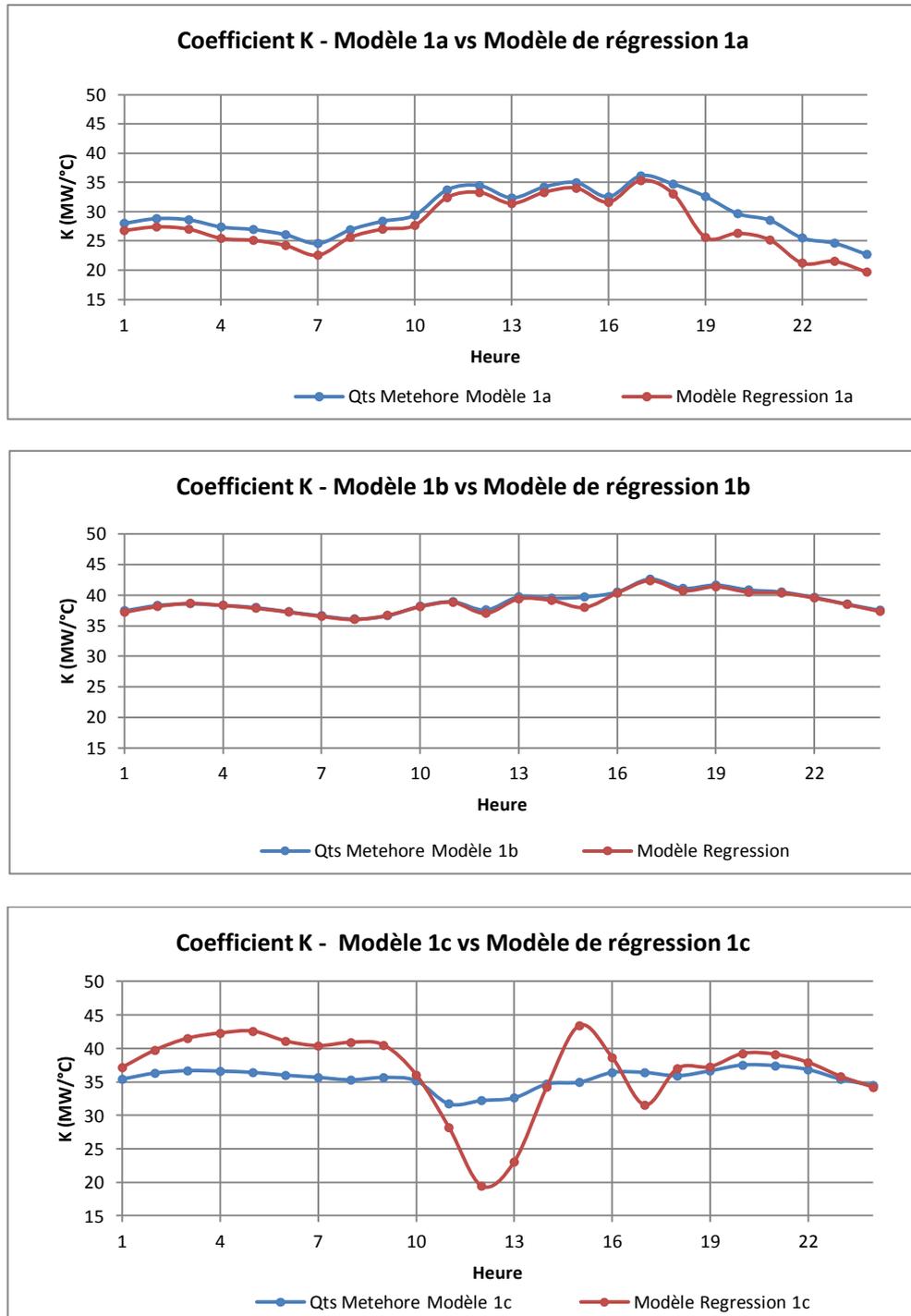
Heure	MAPE Bayes-1a	MAPE Reg-1a	MAPE Bayes-1b	MAPE Reg-1b	MAPE Bayes-1c	MAPE Reg-1c
1	10,6%	10,8%	5,0%	5,0%	5,1%	5,0%
2	11,2%	11,3%	5,1%	5,2%	5,2%	5,2%
3	11,9%	12,0%	5,4%	5,4%	5,5%	5,3%
4	12,8%	13,0%	5,8%	5,7%	5,9%	5,6%
5	13,2%	13,4%	6,1%	6,0%	6,2%	5,9%
6	13,3%	13,4%	5,6%	5,5%	5,7%	5,4%
7	13,7%	13,8%	5,2%	5,2%	5,3%	5,1%
8	11,8%	11,9%	4,6%	4,5%	4,6%	4,4%
9	11,0%	11,1%	4,2%	4,3%	4,3%	4,2%
10	10,3%	10,5%	5,0%	5,0%	5,0%	4,9%
11	9,2%	9,3%	6,4%	6,4%	5,4%	5,4%
12	9,5%	9,6%	8,7%	8,8%	8,4%	8,5%
13	10,7%	10,7%	8,8%	8,8%	8,5%	8,7%
14	10,4%	10,4%	8,4%	8,4%	8,4%	8,4%
15	10,6%	10,6%	8,5%	8,9%	8,5%	8,8%
16	9,7%	9,7%	5,4%	5,4%	5,4%	5,4%
17	8,0%	8,0%	4,4%	4,5%	4,1%	4,2%
18	8,6%	8,5%	4,9%	5,0%	5,0%	5,0%
19	9,3%	9,2%	3,8%	3,9%	3,6%	3,7%
20	9,2%	9,1%	3,6%	3,6%	3,6%	3,5%
21	9,7%	9,5%	3,6%	3,7%	3,6%	3,6%
22	10,3%	10,3%	3,9%	3,9%	3,8%	3,8%
23	10,0%	10,0%	4,2%	4,2%	4,1%	4,1%
24	10,7%	10,8%	4,5%	4,6%	4,4%	4,4%
<b>Moyenne</b>	10,6%	10,7%	5,5%	5,5%	5,4%	5,4%

Ensuite, on a comparé les paramètres estimés par la méthode bayésienne et par la régression. La figure ci-après montre les températures intérieures estimées par les deux approches. Pour le *Modèle 1a*, la régression surestime la température intérieure par rapport à l'estimation bayésienne. On a trouvé des résultats assez proches pour les deux autres modèles.



**Figure 4.14 :** Température intérieures estimées par l’approche bayésienne et l’approche classique avec trois différents modèles.

Les estimations du paramètre K par l’approche bayésienne et par l’approche classique donnent des valeurs très proches pour le *Modèle 1a* et *1b* tandis qu’on observe une grande différence pour le *Modèle 1c* (Figure 4.15).



**Figure 4.15 :** Paramètre  $K$  estimé par l'approche bayésienne et l'approche classique avec trois différents modèles.

L'estimation du paramètre  $H$  par la régression présente des valeurs négatives tandis que ce paramètre représente l'inertie rapide de la demande thermosensible et il devrait avoir des valeurs positives (Figure 4.16).

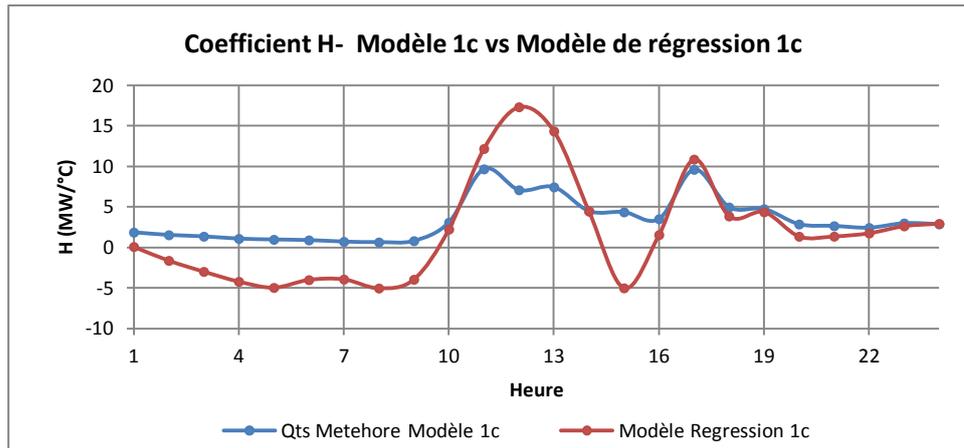


Figure 4.16 : Paramètre H estimé par l'approche bayésienne et l'approche classique.

On pourrait conclure que l'estimation par régression peut nous emmener à des résultats qui n'ont pas de sens physique. L'approche bayésienne permet d'éviter les coefficients négatifs ou déraisonnables grâce à l'information *a priori* sur les paramètres et l'information apportée par les données. On a limité les risques d'obtenir des valeurs de paramètres peu réalistes lorsqu'on cherche à estimer des paramètres qui ont une signification physique.

#### 4.3.2 Analyse de la thermosensibilité avec détermination de la part thermosensible non-linéaire

Cette analyse est basée sur la situation actuelle en supposant que la partie thermosensible de la consommation ( $Q_{ts}^*$ ) comprend déjà un certain pourcentage de pompes à chaleur.

Dans le premier cas, on a choisi des distributions *a priori* dites non informatives, c'est-à-dire, des distributions qui expriment peu d'informations sur une variable aléatoire. Ce modèle possède des paramètres de variances plus faibles et définit donc plutôt des distributions *a priori* informatives, i.e. donnant une information sur la variable aléatoire. Si la connaissance *a priori* sur la valeur de certains paramètres est informative, la calibration devrait donner une distribution *a posteriori* proche de celle *a priori*. Dans la mesure où ces distributions seraient notablement différentes, il conviendrait de remettre en question le modèle, ou encore les données [18].

#### 4.3.2.1 Modèle

Le modèle comporte cinq paramètres ( $K$ ,  $H$ ,  $T_i$ ,  $\tau$  et la part du chauffage linéaire  $\beta$  donc la part des pompes à chaleur est égale à  $(1-\beta)$ ). Si on suppose que l'erreur suit une loi Normale  $N(0,1)$ , le modèle s'écrit sous la forme suivante :

$$Q_{ts}^* \sim Normal(Q_{ts}, \sigma^2)$$

$$Q_{ts} = (K \cdot (T_i - T_{lis}) + H \cdot (T_i - T_{ext})) \cdot (\beta + (1-\beta) \frac{1}{COP}) \quad (4.14)$$

$$COP = 0.098 * T_{ext} + 2.23 \quad (4.15)$$

Cette équation du COP est obtenue par la modélisation bottom-up des PAC air-air et air-eau (Voir Chapitre 3.8).

#### 4.3.2.2 Choix des distributions *a priori*

Les distributions *a priori* des paramètres estimés  $K$ ,  $T_i$  et  $\sigma$  sont identiques au *Modèle 1c*. Contrairement au *Modèle 1c*, on a précisé une distribution informative pour le coefficient  $H$ . En se basant sur les calages du *Modèle 1c*, on a choisi une loi normale centrée autour de 5 de variance 2,5. On a choisi une loi beta<sup>8</sup> [1,1] pour le paramètre  $\beta$  qui est défini entre 0 et 1. Dans le schéma ci-dessous, on résume le *Modèle 2* et ses dépendances.

---

<sup>8</sup> La loi bêta est une famille de lois de probabilités continues, définies sur [0,1], paramétrée par deux paramètres de forme.

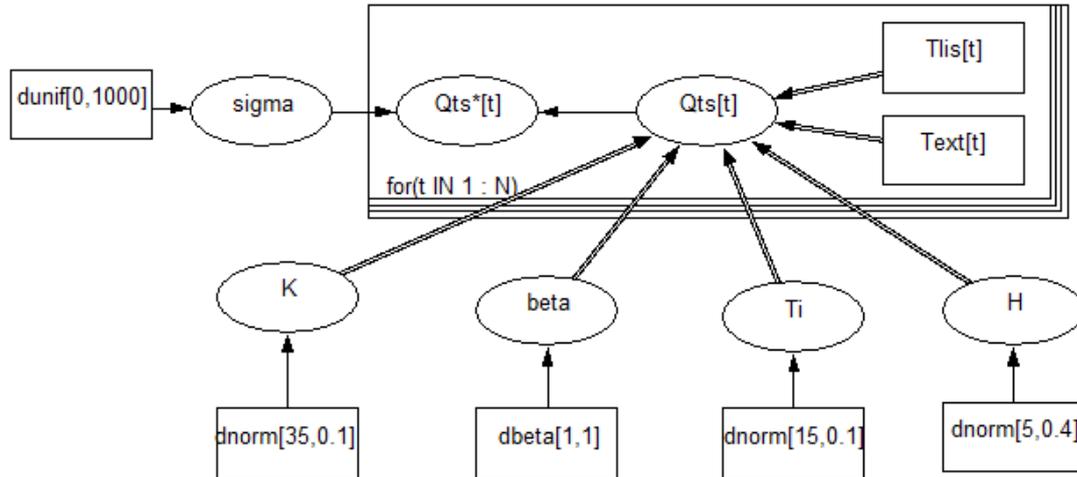
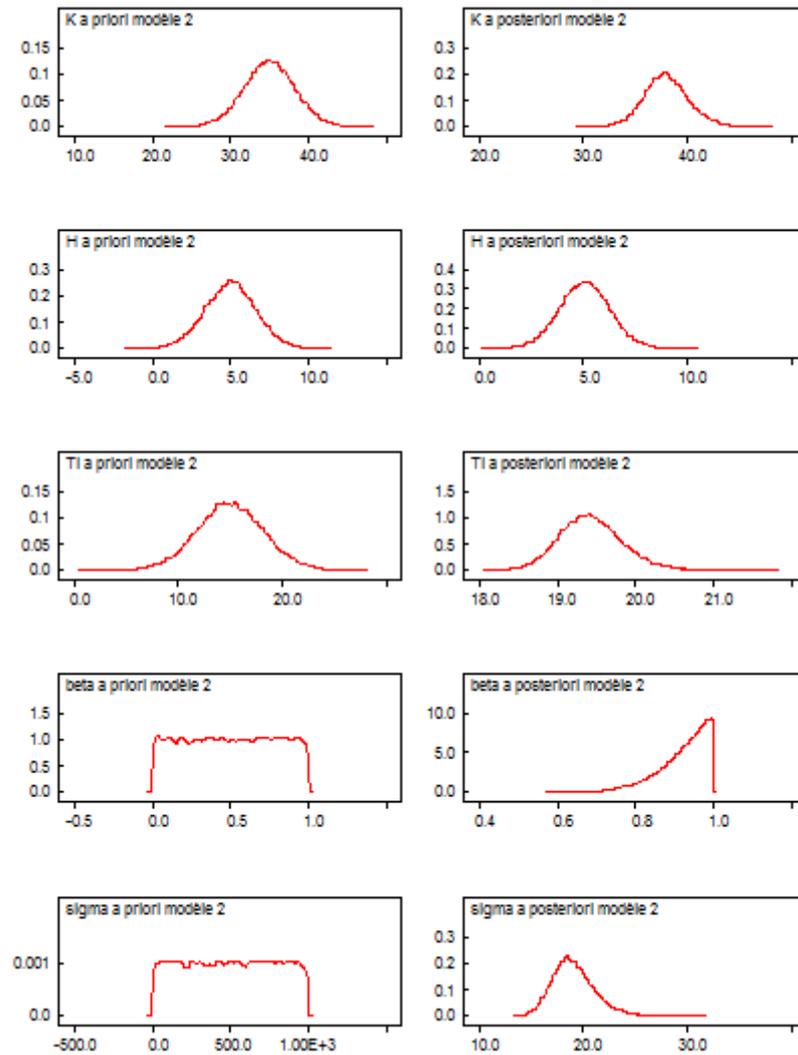


Figure 4.17 : Représentation du *Modèle 2*.

#### 4.3.2.3 Analyse des résultats : distributions a posteriori

Afin de vérifier si les chaînes de Markov ont convergé, on a analysé les autocorrélations et les valeurs de la statistique de convergence de Gelman-Rubin. Ainsi on a validé qu'on a un échantillon valable de la distribution *a posteriori* des paramètres du modèle.

Dans les graphiques ci-dessous, les histogrammes lissés des distributions *a priori* et *a posteriori* des paramètres du modèle pour 19h sont représentés. Les graphiques correspondants à la statistique de Gelman-Rubin et aux distributions a posteriori des paramètres à 01h, 07h et 13h du *Modèle 2* sont donnés dans l'Annexe C.



**Figure 4.18 :** Distribution *a priori* et *a posteriori* des paramètres du *Modèle 2*.

Les distributions *a posteriori* des paramètres  $T_i$ ,  $\beta$ ,  $\sigma$  présentent des changements de position et/ou de forme significatifs qui montrent que le processus inférentiel a permis d'améliorer la connaissance des paramètres du modèle. On a choisi une loi beta [1,1] qui est comme une loi uniforme pour le paramètre  $\beta$  et qui représente le chauffage linéaire en température. On a obtenu une distribution *a posteriori* proche de 1 qui est en cohérence avec la situation réelle.

Les distributions *a posteriori* des paramètres seront proche de leurs distributions *a priori* si les distributions *a priori* sont informatives [177]. Si on ne trouve pas des distributions similaires, cela montrera que les distributions *a priori* ne sont pas bien choisies. Les paramètres pour lesquels on a choisi des distributions *a priori* plus

informatives (avec une plus petite variance), on a obtenue des distributions *a posteriori* proches de celles *a priori*.

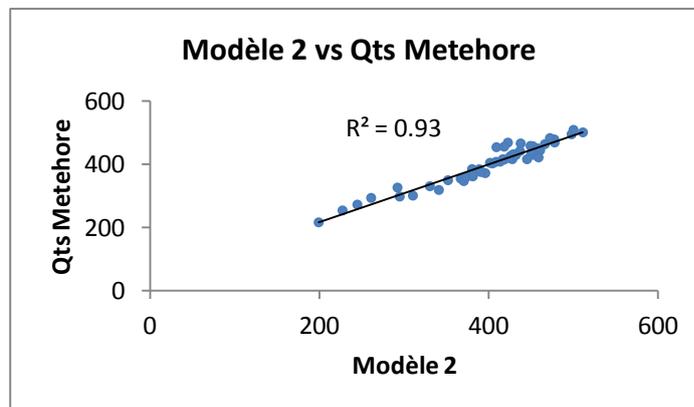
Les statistiques descriptives de chaque paramètre du modèle sont données dans le tableau ci-dessous. Les valeurs des paramètres estimées par le *Modèle 2* sont très similaires aux résultats du *Modèle 1c*. Le paramètre  $\beta$  est estimé à 0,9242 et son intervalle de crédibilité à 95 % est de [0,7702 ; 0,9974]. Ce résultat nous montre que la part des pompes à chaleur régionale est environ de 8%.

**Tableau 4.5** : Statistiques descriptives des paramètres du *Modèle 2* pour 19h.

<b>Modèle 2</b>	<b>Moyenne</b>	<b>Ecart-type</b>	<b>Erreur MC</b>	<b>2,50%</b>	<b>Médiane</b>	<b>97,50%</b>
<b>K</b>	37,97	2,082	0,03549	34,08	37,88	42,3
<b>H</b>	5,052	1,19	0,01196	2,738	5,051	7,407
<b>Ti</b>	19,43	0,3964	0,005727	18,72	19,41	20,28
<b>beta</b>	0,9242	0,06178	0,001158	0,7702	0,9391	0,9974
<b>sigma</b>	19,01	1,944	0,01229	15,68	18,84	23,32

#### 4.3.2.4 Vérification de l'ajustement du modèle

La figure suivant, on présente les données  $Q_{ts}$  ainsi que l'estimation du modèle ( $Q_{ts}^*$ ). On voit que le modèle offre un bon ajustement.



**Figure 4.19** : Ajustement du *Modèle 2* aux données.

Afin de valider l'analyse bayésienne et de vérifier si les informations *a priori* sont en adéquation avec les données, on a calculé « bayesian p value » (Tableau 4.6). Cette valeur est autour de 0,5, ce qui montre que le modèle est bon, c'est-à-dire les données simulées et les données observées sont similaires.

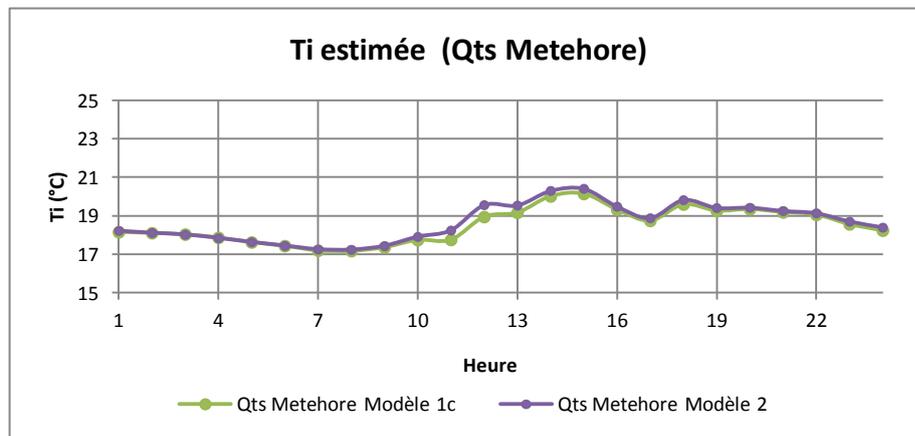
**Tableau 4.6 :** Bayesian p-value et DIC du *Modèle 2*.

	Bayesian p-value	DIC
<b>Modèle 2</b>	0,5279	472,71

Le DIC du *Modèle 2* est très proche du *Modèle 1c*, ce qui montre qu'on ne peut pas faire un choix entre ces deux modèles en se basant seulement sur ce critère.

#### 4.3.2.5 Discussion des résultats

Dans cette partie, on compare les résultats du *Modèle 2* avec ceux du *Modèle 1c*. La température intérieure estimée par les deux modèles est très similaire. On voit une légère différence entre 10h et 13h (Figure 4.20).



**Figure 4.20 :** Comparaison de la température intérieure estimée par le *Modèle 2* et le *Modèle 1c*.

Les paramètres K et H estimés par le *Modèle 2* présentent moins de variation au cours de la journée que les estimations du *Modèle 1c* (Figure 4.21 et 4.22).

Le *Modèle 2* surestime le paramètre K par rapport au *Modèle 1c*. On observe moins de variation pendant la journée. Le *Modèle 1c* trouve des valeurs plus basses pour les heures entre 11h et 15h.

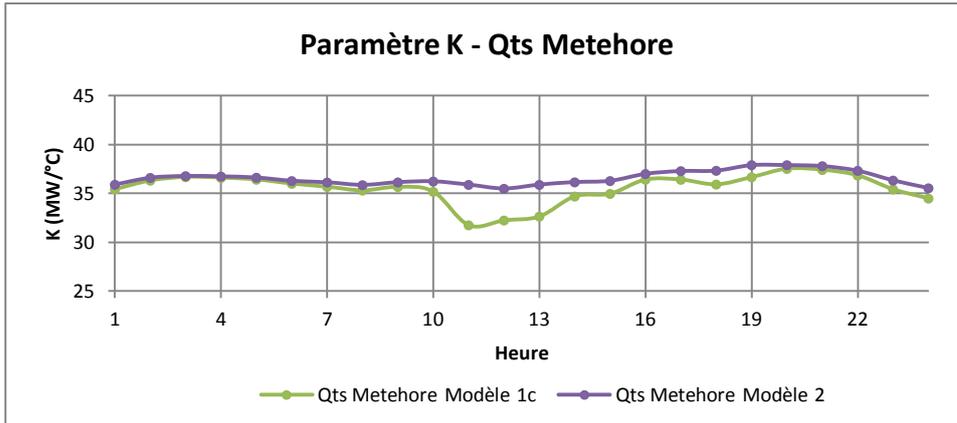


Figure 4.21 : Comparaison du paramètre K estimé par le *Modèle 2* et le *Modèle 1c*.

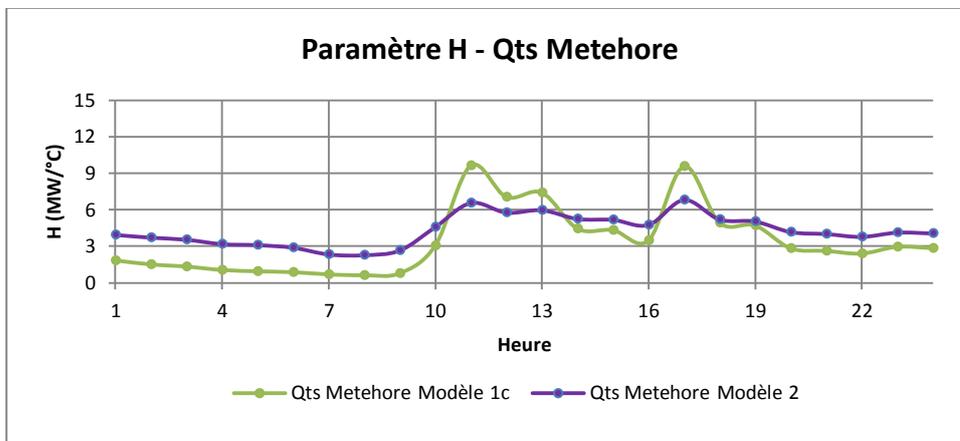


Figure 4.22 : Comparaison du paramètre H estimé par le *Modèle 2* et le *Modèle 1c*.

Les résultats montrent que la part moyenne des pompes à chaleur estimée par le *Modèle 2* est de 10% (Figure 4.23) ce qui est raisonnable, en raison de l'existence de différentes sources de non-linéarité y compris les auxiliaires de chauffage.

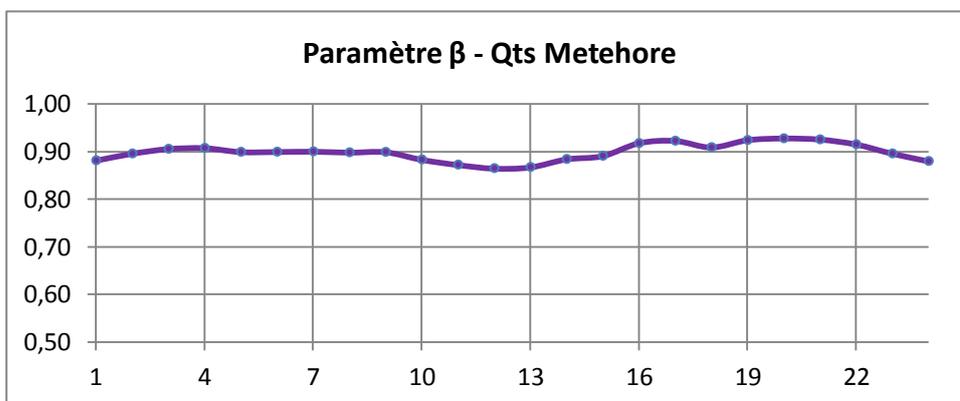


Figure 4.23 : Paramètre  $\beta$  estimé par le *Modèle 2*.

## 4.4 Sensibilité des Inférences

### 4.4.1 Sensibilité des inférences aux distributions a priori

Le choix de la loi a priori comporte une certaine part d'arbitraire, et ce à deux niveaux. Le premier niveau réside dans le choix de la famille de probabilités retenue pour la loi a priori. Le second réside dans le choix des valeurs [154].

Le fait de tester différents *a priori* permet de tester la stabilité de l'analyse bayésienne. Il convient d'examiner dans quelle mesure les choix des distributions *a priori* affectent l'estimation. Dans la pratique, on se contente d'examiner, dans quelle mesure, une perturbation des paramètres modifie l'estimation.

Une analyse de sensibilité a été menée en utilisant d'autres distributions *a priori* pour les paramètres clés (K, H,  $\beta$ , Ti) du *Modèle 2*. L'influence du choix des distributions *a priori* sur les paramètres estimés a été étudiée en utilisant différentes distributions et en comparant les distributions *a posteriori* obtenues pour les paramètres.

#### 4.4.1.1 Sensibilité des inférences aux distributions a priori du paramètre K

On a étudié l'influence de la distribution *a priori* du paramètre K en utilisant quatre distributions différentes :

- **Cas 1** : une loi normale centrée autour de 35 MW/°C de variance 100.
- **Cas 2** : une loi normale centrée autour de 35 MW/°C de variance 0.5.
- **Cas 3** : une loi normale centrée autour de 50 MW/°C de variance 10.
- **Cas 4** : une loi gamma de paramètre (p,  $\theta$ ) avec p=60 et  $\theta=2$  où  
moyenne(K)=p/ $\theta$  et variance(K)=p/ $\theta^2$

Les valeurs moyennes estimées pour chaque paramètre sont données dans le tableau ci-après.

**Tableau 4.7** : Sensibilité des moyennes *a posteriori* de l'inférence à la distribution *a priori* du paramètre K.

Modèle 2	Référence	Cas 1	Cas 2	Cas 3	Cas 4
	dnorm(35,0.1)	dnorm(35,0.01)	dnorm(35,2)	dnorm(50,0.1)	dgamma(60,2)
<b>K</b>	37,97	41,32	35,36	48,76	37,36
<b>H</b>	5,052	4,652	5,611	4,215	5,199
<b>Ti</b>	19,43	19,36	19,67	19,45	19,51
<b>beta</b>	0,9242	0,8502	0,9578	0,664	0,9299
<b>sigma</b>	19,01	19,03	19,18	19,21	19,06

L'analyse de sensibilité réalisée à l'aide de distributions *a priori* plus informatives (cas 2 et cas 4) a révélé une influence modérée du choix des *a priori* du paramètre K. Le choix d'un *a priori* plus vague (plus grande variance - cas 1) a engendré des estimations différentes des paramètres K et  $\beta$ . Quand on choisit une distribution autour d'une valeur plus grande (50), l'estimation des paramètres K et  $\beta$  sont fortement influencée.

Les distributions *a priori* moins informatives (avec une plus grande variance) du paramètre K conduisent essentiellement à des intervalles de crédibilité plus larges traduisant une plus grande incertitude (Tableau 4.8).

**Tableau 4.8** : Intervalles de crédibilité à 95% pour différentes distributions *a priori* du paramètre K.

Modèle 2	intervalle à 95%	
	borne inférieure	borne supérieure
<b>K-réf</b>	34,08	42,3
<b>K-1</b>	35,2	50,3
<b>K-2</b>	34,02	36,67
<b>K-3</b>	42,84	54,76
<b>K-4</b>	33,02	42,55
<b>H-réf</b>	2,738	7,407
<b>H-1</b>	2,27	7,093
<b>H-2</b>	3,561	7,718
<b>H-3</b>	1,671	6,778
<b>H-4</b>	2,857	7,573
<b>Ti-réf</b>	18,72	20,28
<b>Ti-1</b>	18,61	20,25
<b>Ti-2</b>	19,08	20,39
<b>Ti-3</b>	18,58	20,47
<b>Ti-4</b>	18,75	20,39
<b>beta-réf</b>	0,7702	0,9974
<b>beta-1</b>	0,5995	0,9937
<b>beta-2</b>	0,8656	0,9987
<b>beta-3</b>	0,4929	0,8576
<b>beta-4</b>	0,7755	0,9979
<b>sigma-réf</b>	15,68	23,32
<b>sigma-1</b>	15,07	23,25
<b>sigma-2</b>	15,84	23,46
<b>sigma-3</b>	15,87	23,52
<b>sigma-4</b>	15,68	23,28

#### 4.4.1.2 Sensibilité des inférences aux distributions *a priori* du paramètre H

Pour le paramètre H, on a aussi testé deux distributions normales avec différentes variances, plus grande et plus petite que le cas référence, une distribution autour d'une valeur plus élevée et une distribution gamma.

- **Cas 1** : une loi normale centrée autour de 5 MW/°C de variance 10.
- **Cas 2** : une loi normale centrée autour de 5 MW/°C de variance 0.5.
- **Cas 3** : une loi normale centrée autour de 20 MW/°C de variance 5.
- **Cas 4** : une loi gamma de paramètre  $(p,\theta)$  avec  $p=10$  et  $\theta=2$  où  $\text{moyenne}(H)=p/\theta$  et  $\text{variance}(H)=p/\theta^2$

Dans le tableau ci-dessous les estimations du modèle pour trois différentes distributions *a priori* du paramètre H sont représentées. On observe que les valeurs moyennes des paramètres ne sont pas affectées par le choix du paramètre H sauf le cas 3. Le choix de la valeur du paramètre H a un impact fort sur l'estimation du paramètre  $\beta$  et  $\sigma$ .

**Tableau 4.9** : Sensibilité des moyennes *a posteriori* de l'inférence à la distribution *a priori* du paramètre H.

Modèle 2	Référence	Cas 1	Cas 2	Cas 3	Cas 4
	dnorm(5,0.4)	dnorm(5,0.1)	dnorm(5,2)	dnorm(20,0.4)	dgamma(10,2)
<b>K</b>	37,97	37,91	38	36,26	37,98
<b>H</b>	5,052	5,136	5,007	16,8	4,933
<b>Ti</b>	19,43	19,43	19,43	19,67	19,45
<b>beta</b>	0,9242	0,9237	0,9249	0,6436	0,9253
<b>sigma</b>	19,01	19,08	18,94	24,35	19,01

Les intervalles de crédibilité à 95% sont donnés dans le Tableau 4.10. On observe des intervalles de crédibilité du paramètre H sont plus larges pour le cas 3. Les autres paramètres ne sont pas influencés par le choix de la distribution *a priori* du paramètre H sauf  $\beta$  et  $\sigma$ .

**Tableau 4.10** : Intervalles de crédibilité à 95% pour différentes distributions *a priori* du paramètre H.

Modèle 2	intervalle à 95%	
	borne inférieure	borne supérieure
<b>K-réf</b>	34,08	42,3
<b>K-1</b>	33,82	42,36
<b>K-2</b>	34,39	42,28
<b>K-3</b>	30,77	41,87
<b>K-4</b>	34,05	42,35
<b>H-réf</b>	2,738	7,407
<b>H-1</b>	2,094	8,255
<b>H-2</b>	3,712	6,316
<b>H-3</b>	13,59	20
<b>H-4</b>	2,806	7,347
<b>Ti-réf</b>	18,72	20,28
<b>Ti-1</b>	18,72	20,29
<b>Ti-2</b>	18,72	20,27
<b>Ti-3</b>	18,58	20,98
<b>Ti-4</b>	18,73	20,29
<b>beta-réf</b>	0,7702	0,9974
<b>beta-1</b>	0,7722	0,9974
<b>beta-2</b>	0,7742	0,9974
<b>beta-3</b>	0,4552	0,8578
<b>beta-4</b>	0,775	0,9976
<b>sigma-réf</b>	15,68	23,32
<b>sigma-1</b>	15,71	23,36
<b>sigma-2</b>	15,63	23,15
<b>sigma-3</b>	19,76	30,2
<b>sigma-4</b>	15,67	23,21

#### 4.4.1.3 Sensibilité des inférences aux distributions *a priori* du paramètre Ti

On a testé quatre différentes distributions *a priori* comme dans les cas précédents :

- **Cas 1** : une loi normale centrée autour de 15°C de variance 100.
- **Cas 2** : une loi normale centrée autour de 15°C de variance 0.5.
- **Cas 3** : une loi normale centrée autour de 25 MW/°C de variance 10.
- **Cas 4** : une loi gamma de paramètre (p,θ) avec p=30 et θ=2 où  
 $\text{moyenne}(Ti)=p/\theta$  et  $\text{variance}(Ti)=p/\theta^2$

Le choix de la distribution *a priori* du paramètre Ti a une influence plus importante sur le paramètre β que les autres paramètres. On voit comment le choix de la variance du paramètre Ti affecte l'estimation.

**Tableau 4.11** : Sensibilité des moyennes *a posteriori* de l'inférence à la distribution *a priori* du paramètre  $T_i$ .

Modèle 2	Référence	Cas 1	Cas 2	Cas 3	Cas 4
	dnorm(15,0.1)	dnorm(15,0.01)	dnorm(15,2)	dnorm(25,0.1)	dgamma(30,2)
<b>K</b>	37,97	37,86	40,11	37,57	38,02
<b>H</b>	5,052	5,035	5,434	5,005	5,051
<b>Ti</b>	19,43	19,51	18,62	19,6	19,42
<b>beta</b>	0,9242	0,9188	0,9589	0,9161	0,924
<b>sigma</b>	19,01	19,05	19,49	19,09	19

Les intervalles de crédibilité des paramètres  $\beta$  et  $K$  est plus restreints pour le cas 2. (Tableau 4.12). Cela veut dire que si on améliore la loi *a priori* (petite variance) sur le paramètre  $T_i$ , on aura moins d'incertitude sur les paramètres  $\beta$  et  $K$ .

**Tableau 4.12** : Intervalles de crédibilité à 95% pour différentes distributions *a priori* du paramètre  $T_i$ .

Modèle 2	intervalle à 95%	
	borne inférieure	borne supérieure
<b>K-réf</b>	34,08	42,3
<b>K-1</b>	33,93	42,28
<b>K-2</b>	36,67	43,85
<b>K-3</b>	33,59	42,02
<b>K-4</b>	34,23	42,38
<b>H-réf</b>	2,738	7,407
<b>H-1</b>	2,735	7,396
<b>H-2</b>	3,129	7,801
<b>H-3</b>	2,695	7,353
<b>H-4</b>	2,728	7,378
<b>Ti-réf</b>	18,72	20,28
<b>Ti-1</b>	18,77	20,38
<b>Ti-2</b>	18,07	19,19
<b>Ti-3</b>	18,85	20,48
<b>Ti-4</b>	18,72	20,25
<b>beta-réf</b>	0,7702	0,9974
<b>beta-1</b>	0,7656	0,9973
<b>beta-2</b>	0,8643	0,9989
<b>beta-3</b>	0,7578	0,9971
<b>beta-4</b>	0,7715	0,9974
<b>sigma-réf</b>	15,68	23,32
<b>sigma-1</b>	15,07	23,32
<b>sigma-2</b>	15,99	23,98
<b>sigma-3</b>	15,77	23,41
<b>sigma-4</b>	15,67	23,17

#### 4.4.1.4 Sensibilité des inférences aux distributions *a priori* du paramètre $\beta$

Pour analyser la sensibilité des inférences au paramètre  $\beta$ , on a testé deux distributions beta avec différents paramètres de forme que le cas référence et une distribution normale positive autour de 0,5.

- **Cas 1** : une loi beta de paramètres (5,5)
- **Cas 2** : une loi beta de paramètres (25,5)
- **Cas 3** : une loi normale centrée autour de 0,5 de variance 0,5 tronquée entre 0 et 1.

Dans les tableaux ci-dessous les valeurs moyennes et les intervalles de crédibilité des paramètres sont montrées pour trois différentes distributions *a priori* du paramètre  $\beta$ . Le paramètre K est influencé par le choix du paramètre  $\beta$  plus que les autres.

**Tableau 4.13** : Sensibilité des moyennes *a posteriori* de l'inférence à la distribution *a priori* du paramètre.

Modèle 2	Référence	Cas 1	Cas 2	Cas 3
	$\beta \sim \text{dbeta}(1,1)$	$\beta \sim \text{dbeta}(5,5)$	$\beta \sim \text{dbeta}(25,5)$	$\beta \sim \text{dnorm}(0.5,2)I(0,1)$
<b>K</b>	37,97	40,17	38,81	38,09
<b>H</b>	5,052	5,374	5,15	5,045
<b>Ti</b>	19,43	19,77	19,54	19,44
<b>beta</b>	0,9242	0,8051	0,8795	0,9197
<b>sigma</b>	19,01	19,25	19,07	19,01

Donner à  $\beta$  des distributions *a priori* différentes induit des variations moins importantes de la moyenne *a posteriori* et des intervalles de crédibilité des paramètres Ti, H et  $\sigma$  que les paramètres K et  $\beta$ .

**Tableau 4.14 :** Intervalles de crédibilité à 95% pour différentes distributions *a priori* du paramètre  $\beta$ .

Modèle 2	intervalle à 95%	
	borne inférieure	borne supérieure
<b>K-réf</b>	34,08	42,3
<b>K-1</b>	35,95	44,67
<b>K-2</b>	35,06	42,77
<b>K-3</b>	34,19	42,53
<b>H-réf</b>	2,738	7,407
<b>H-1</b>	2,996	7,809
<b>H-2</b>	2,796	7,519
<b>H-3</b>	2,708	7,412
<b>Ti-réf</b>	18,72	20,28
<b>Ti-1</b>	18,99	20,07
<b>Ti-2</b>	18,82	20,36
<b>Ti-3</b>	18,73	20,03
<b>beta-réf</b>	0,7702	0,9974
<b>beta-1</b>	0,6713	0,9203
<b>beta-2</b>	0,7759	0,9559
<b>beta-3</b>	0,7673	0,9971
<b>sigma-réf</b>	15,68	23,32
<b>sigma-1</b>	15,86	23,57
<b>sigma-2</b>	15,76	23,32
<b>sigma-3</b>	15,68	23,25

#### 4.4.2 Sensibilité des inférences aux valeurs des constantes du modèle de COP

La sensibilité des inférences à la valeur de la constante COP a été d'abord étudiée avec trois valeurs différentes : une valeur de référence et deux valeurs qui correspondent à des écarts de  $\pm 50\%$  par rapport à la référence. Les inférences ont été conduites sur trois modèles différents en modifiant la valeur du COP. Les estimations des moyennes *a posteriori* des paramètres du modèle, obtenues avec différentes valeurs de COP sont données dans le Tableau 4.15.

**Tableau 4.15 :** Sensibilité des moyennes *a posteriori* de l'inférence aux valeurs des constantes du COP correspondant à des écarts de  $\pm 50\%$  par rapport à la référence.

Modèle 2	Cas 1	COP Référence	Cas 2
	COP - 50%		COP +50%
<b>K</b>	37,31	37,97	38,05
<b>H</b>	4,713	5,052	5,134
<b>Ti</b>	19,5	19,43	19,42
<b>beta</b>	0,9002	0,9242	0,9321
<b>sigma</b>	19,1	19,01	18,98

On a ensuite étudié l'impact de différentes équations du COP. Pour cela on a utilisé l'équation du COP moyen pondéré du « Scénario basse performance » et du COP moyen pondéré du « Scénario haute performance » calculés dans le chapitre précédent (voir Chapitre 5.8).

**Tableau 4.16 :** Sensibilité des moyennes *a posteriori* de l'inférence aux valeurs des constantes du COP correspondant au COP « basse performance » et au COP « haute performance ».

Modèle 2	Cas 3	COP Référence	Cas 4
	COP-HP		COP-BP
<b>K</b>	38	37,97	37,94
<b>H</b>	5,132	5,052	5,017
<b>Ti</b>	19,42	19,43	19,44
<b>beta</b>	0,9228	0,9242	0,9259
<b>sigma</b>	18,98	19,01	19,02

Les résultats révèlent que pour toutes les valeurs des constantes testées, les inférences ne sont pas sensibles. Les paramètres sont toujours estimés à des valeurs très proches du modèle de référence.

**Tableau 4.17 :** Intervalles de crédibilité à 95% pour différentes constantes du COP.

Modèle 2	intervalle à 95%	
	borne inférieure	borne supérieure
<b>K-réf</b>	34,08	42,3
<b>K-1</b>	33,77	40,95
<b>K-2</b>	34,18	42,56
<b>K-3</b>	34,11	42,34
<b>K-4</b>	34,16	42,29
<b>H-réf</b>	2,738	7,407
<b>H-1</b>	2,394	7,031
<b>H-2</b>	2,802	7,521
<b>H-3</b>	2,794	7,488
<b>H-4</b>	2,702	7,33
<b>Ti-réf</b>	18,72	20,28
<b>Ti-1</b>	18,75	20,45
<b>Ti-2</b>	18,72	20,21
<b>Ti-3</b>	18,73	20,22
<b>Ti-4</b>	18,73	20,3
<b>beta-réf</b>	0,7702	0,9974
<b>beta-1</b>	0,6899	0,997
<b>beta-2</b>	0,8002	0,9976
<b>beta-3</b>	0,7788	0,9972
<b>beta-4</b>	0,7789	0,9974
<b>sigma-réf</b>	15,68	23,32
<b>sigma-1</b>	15,72	23,35
<b>sigma-2</b>	15,68	23,13
<b>sigma-3</b>	15,66	23,24
<b>sigma-4</b>	15,68	23,25

Les intervalles de crédibilité ne sont pas changés pour différentes valeurs des constantes (Tableau 4.16). L'analyse de sensibilité conduite suggère que les résultats obtenus ne sont pas dépendants du choix des constantes du modèle de COP.

Les analyses de sensibilité effectuées n'ont pas révélé de changement significatif dans les intervalles de crédibilité.

#### 4.5 Test du modèle sur des courbes de charge déformées

Dans cette partie, on se place dans l'optique d'une modification de la courbe de charge thermosensible du à une pénétration beaucoup plus importante des PAC. On simule ainsi à partir de la courbe de charge actuelle une courbe de charge hypothétique qui pourrait correspondre à une situation future. Il s'agit donc d'une démarche prospective très simplifiée dont le seul but est de tester la robustesse de notre modèle par rapport à une évolution de la demande.

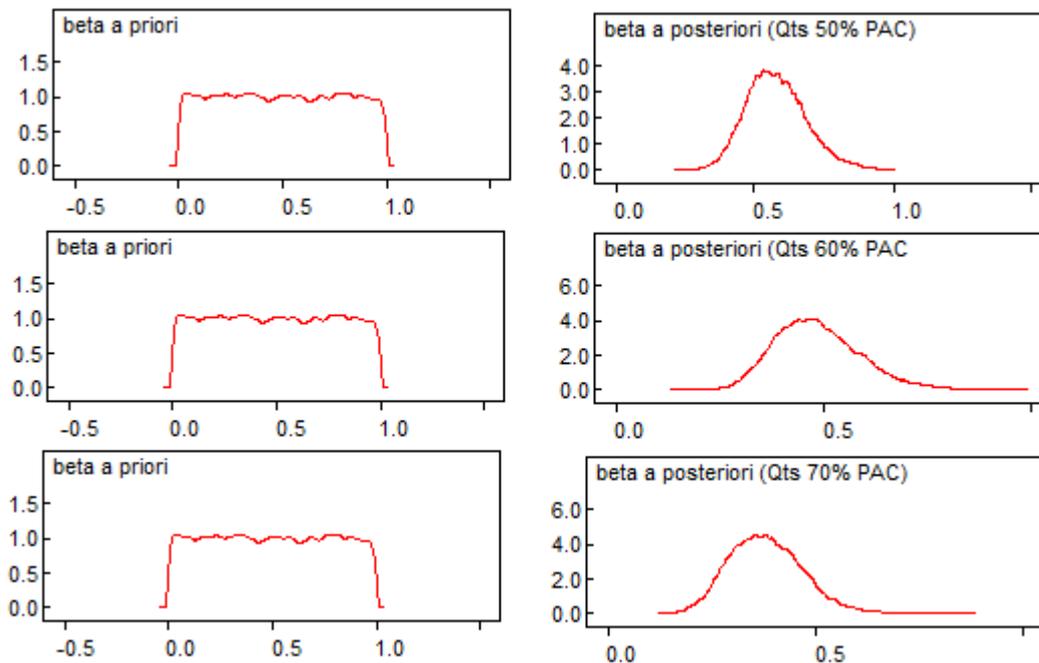
On suppose que :

$$Q_{ts((1-\beta)\% PAC)} = Q_{ts} \cdot \left( \beta + \frac{(1-\beta)}{COP} \right) \quad (4.16)$$

On a testé le modèle avec trois différents  $\beta$  :

- $\beta = 0.5$  c'est-à-dire 50% de chauffage Joule et 50% de PAC
- $\beta = 0.4$  c'est-à-dire 40% de chauffage Joule et 60% de PAC
- $\beta = 0.3$  c'est-à-dire 30% de chauffage Joule et 70% de PAC

Les distributions a priori et a posteriori du paramètre  $\beta$  correspondantes à ces différentes valeurs sont données dans la figure ci-dessous.



**Figure 4.24 :** Distribution a priori et a posteriori du paramètre  $\beta$  pour trois différentes valeurs.

Pour chaque cas, on observe que le modèle est capable de saisir la non-linéarité ajoutée aux données de la consommation avec une surestimation de 8% du paramètre  $\beta$  (Tableau 4.18). Les valeurs moyennes des distributions a posteriori des autres paramètres du modèle sont très similaires pour chacun de ces trois cas. Cela montre que les estimations du modèle sont cohérentes (le parc du bâtiment reste inchangé sauf la non-linéarité des usages change).

**Tableau 4.18** : Statistiques descriptives des paramètres du modèle pour des courbes de charge modifiées avec 50%, 60% et 70% de pénétration de PAC.

50% PAC	Moyenne	Ecart-type	Erreur MC	2,50%	Médiane	97,50%
<b>H</b>	4,373	1,109	0,01181	2,241	4,362	6,594
<b>K</b>	33,89	3,005	0,09384	27,93	33,91	39,76
<b>Ti</b>	19,1	0,4612	0,01035	18,26	19,07	20,07
<b>beta</b>	0,5798	0,1081	0,003636	0,3897	0,572	0,8204
<b>sigma</b>	12,19	1,231	0,007989	10,09	12,09	14,87
60% PAC	Moyenne	Ecart-type	Erreur MC	2,50%	Médiane	97,50%
<b>H</b>	4,454	1,109	0,01161	2,295	4,443	6,658
<b>K</b>	33,42	3,056	0,09675	27,3	33,47	39,27
<b>Ti</b>	19,06	0,4742	0,0109	18,2	19,03	20,06
<b>beta</b>	0,4858	0,1024	0,003454	0,3134	0,4762	0,7162
<b>sigma</b>	10,93	1,113	0,007439	9,025	10,83	13,33
70% PAC	Moyenne	Ecart-type	Erreur MC	2,50%	Médiane	97,50%
<b>H</b>	4,464	1,087	0,01114	2,356	4,463	6,609
<b>K</b>	33,31	2,939	0,09166	27,53	33,34	39,05
<b>Ti</b>	19,02	0,4791	0,0119	18,15	18,99	20,03
<b>beta</b>	0,3815	0,08881	0,003015	0,2266	0,3761	0,5722
<b>sigma</b>	9,634	0,9826	0,00662	7,948	9,55	11,8

#### 4.6 Validation des Modèles

Pour vérifier l'adéquation de l'ajustement du modèle aux données, nous avons effectué les analyses de validation croisée (CV). La validation croisée (cross-validation) est une méthode d'estimation de fiabilité d'un modèle fondé sur une technique d'échantillonnage. Le principe consiste à estimer le modèle sur une partie de l'échantillon puis à calculer l'erreur commise sur une autre partie de l'échantillon qui n'a pas participé à l'estimation des paramètres de ce modèle. Ce calcul est itéré sur plusieurs échantillons de validation afin d'améliorer en moyennant la précision. Différentes façons de constituer des échantillons de validation peuvent être

considérées. On a utilisé la validation croisée K-fold avec  $k=5$ . On divise  $k$  fois l'échantillon, puis on sélectionne un des  $k$  échantillons comme ensemble de validation et les  $(k-1)$  autres échantillons constitueront l'ensemble d'apprentissage. On calcule comme dans la première méthode l'erreur quadratique moyenne. Puis on répète l'opération en sélectionnant un autre échantillon de validation parmi les  $(k-1)$  échantillons qui n'ont pas encore été utilisés pour la validation du modèle. L'opération se répète ainsi  $k$  fois pour qu'en fin de compte chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de validation. L'erreur est estimée en calculant un test, une mesure ou un score de performance du modèle sur l'échantillon de test, par exemple l'erreur quadratique moyenne. La moyenne des  $k$  erreurs quadratiques moyennes est enfin calculée pour estimer l'erreur de prédiction [178].

Pour l'étape de validation croisée, dans le cadre de nos études, des prédictions ont été obtenues pour le modèle calibré avec les données de 19h. On présente dans le Tableau 4.17 l'erreur quadratique moyenne (MSE), l'erreur quadratique moyenne en racine carré (RMSE) et l'erreur moyenne absolue en pourcentage (MAPE) des échantillons. La MAPE moyenne des échantillons est de 4,67%, ce qui montre que le modèle est valide.

**Tableau 4.19** : Erreurs moyennes des échantillons de test.

<b>Modèle 2</b>	<b>MSE</b>	<b>RMSE</b>	<b>MAPE</b>
<b>test1</b>	424,21	20,60	4,98%
<b>test2</b>	640,96	25,32	5,84%
<b>test3</b>	829,71	28,80	6,20%
<b>test4</b>	215,53	14,68	2,62%
<b>test5</b>	351,39	18,75	3,72%
<b>moyenne</b>	492,36	21,63	4,67%

Les données de 19h de l'hiver d'une année différente peuvent être utilisées pour une recalibration du modèle. De ce fait, toute l'information disponible serait utilisée pour renseigner les paramètres du modèle. Pour clore ce chapitre, les limites de ce modèle sont discutées et quelques voies d'amélioration prometteuses sont proposées.

## 4.7 Conclusion

L'effort de modélisation a surtout porté sur l'analyse de la part thermosensible de la consommation. L'objectif poursuivi est de développer un modèle pour la détermination des coefficients de déperdition de chaleur, la température intérieure et la part du chauffage linéaire et non-linéaire. L'accent est donc mis sur l'estimation de ces paramètres par un modèle simple plutôt que sur la construction de modèles complexes et spécifiques pour chaque usage. L'analyse permet d'obtenir la distribution *a posteriori* de toutes les inconnues du modèle. Ces distributions *a posteriori* décrivent l'incertitude de façon explicite.

Compte tenu du caractère convaincant de ces résultats, il nous semble que l'utilisation de l'identification bayésienne permet de dépasser le problème de la non-linéarité des usages lorsque les distributions *a priori* appropriées sont sélectionnées. On pense que le modèle est capable de déterminer la variation de la consommation liée aux caractéristiques du bâtiment, en présence d'usages thermiques non-linéaires. La détermination de la pénétration des pompes à chaleur par le modèle semble intéressante, même si elle reste à être validée par des études de terrain.

Les modèles restent simples et reposent sur des hypothèses simplificatrices fortes. Notamment, les gains gratuits, l'appoint électrique et les chauffe-eau sont ignorés alors qu'ils peuvent avoir une influence importante sur les estimations. Dans la perspective d'améliorer ce travail, chaque année pourrait faire l'objet d'une modélisation plus complexe et spécifique, tenant compte d'autres usages thermosensibles, d'autres variables climatiques comme l'ensoleillement en faisant la distinction entre le secteur résidentiel et le secteur tertiaire. L'utilisation de distributions *a priori* informatives basées sur des panels de mesure ou des données régionales augmentera la précision des modèles. On pourrait aussi utiliser les lois *a posteriori* trouvées pour une année en les supposant lois *a priori* pour l'année suivante. Les lois *a posteriori* d'une région pourraient être utilisées comme *a priori* pour une autre région. Une modélisation de la température intérieure plus adaptée pourrait aussi être proposée. L'aspect dynamique du parc de bâtiment (surface chauffée variable par exemple) n'est pas pris en compte dans le modèle. Afin de prendre en compte cette dépendance temporelle dans les variables, le modèle pourrait être traité comme un modèle dynamique.

