



HAL
open science

**Optimisation des plans de mesure dans un contexte de
garantie de performance énergétique des bâtiments :
Caractérisation des incertitudes de mesure et
propagation au sein de modèles dynamiques**

Fally Titikpina

► **To cite this version:**

Fally Titikpina. Optimisation des plans de mesure dans un contexte de garantie de performance énergétique des bâtiments : Caractérisation des incertitudes de mesure et propagation au sein de modèles dynamiques. Ingénierie, finance et science [cs.CE]. Université d'Angers, 2016. Français. NNT : 2016ANGE0084 . tel-03245912

HAL Id: tel-03245912

<https://theses.hal.science/tel-03245912>

Submitted on 2 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse de Doctorat

Fally TITIKPINA

*Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université d'Angers
sous le label de l'Université de Nantes Angers Le Mans*

École doctorale : Sciences et technologies de l'information et des mathématiques

Discipline : Sciences pour l'ingénieur, section CNU 60, 62

Unité de recherche : Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS)

Soutenue le 06 Décembre 2016

Optimisation des plans de mesure dans un contexte de garantie de performance énergétique des bâtiments Caractérisation des incertitudes de mesure et propagation au sein de modèles dynamiques

JURY

Présidente : **M^{me} Monika WOLOSZYN**, Professeure, Université de Savoie Mont Blanc, LOCIE
Rapporteurs : **M. Bruno PEUPORTIER**, Directeur de recherche, MINES Paristech, CES
M. Marc HIMBERT, Professeur du CNAM, LCM LNE-CNAM
Examineurs : **M. Antoine CAUCHETEUX**, Ingénieur chercheur, Cerema
M^{me} Marjorie MUSY, Chercheure, IRSTV
Invité : **M. Guillaume ANSANAY-ALEX**, Ingénieur chercheur, CSTB
Directeur de thèse : **M. David BIGAUD**, Professeur, Université d'Angers, LARIS
Co-directeur de thèse : **M. Abderafi CHARKI**, Maître de conférences, HDR, Université d'Angers, LARIS

*À mes parents,
Riss & Hamza.*

*«However difficult life may seem,
there is always something you can do and succeed at.»*

Stephen Hawking

Remerciements

En dernière année d'école d'ingénieurs lorsque la plupart de mes collègues se bousculaient pour trouver un sujet de thèse, je me disais «une thèse? Moi, jamais!». Pourtant me voici quelques années plus tard terminant la rédaction de ce manuscrit. Je voudrais pour cela, remercier Kondo Adjallah pour m'avoir donné goût à la recherche et pour avoir vu en moi ce potentiel dont je n'avais pas conscience.

Cette thèse fut une merveilleuse aventure. Une aventure durant laquelle j'ai eu à travailler au sein de la meilleure équipe qui soit!

Je voudrais particulièrement remercier David BIGAUD, mon directeur de thèse. Merci pour la confiance qu'il m'a accordé en me proposant de participer à ce projet de recherche. Merci de m'avoir donné la liberté de m'exprimer à travers ces travaux durant ces trois années de recherche. Son perpétuel questionnement m'a toujours permis d'avoir un regard neuf sur mes travaux et à mieux les valoriser.

Mes remerciements les plus profonds vont à Antoine CAUCHETEUX. Merci Antoine, pour ton encadrement au quotidien, pour tes nombreux conseils, pour ta patience et surtout pour le temps que tu m'as accordé. Merci de m'avoir inculqué les bases de la thermique du bâtiment mais aussi de m'avoir mise à l'aise dès le départ en me mettant dans de bonnes conditions de travail. Et surtout, de m'avoir continuellement encouragé à donner le meilleur de moi-même durant ces trois années. Ce fût un réel plaisir!

Merci à tous les deux! C'est essentiellement grâce à vous que je suis la chercheuse que je suis aujourd'hui. Je m'envole vers de nouveaux horizons professionnels avec des compétences essentielles et surtout pleine de confiance.

La thèse c'est aussi un cadre de travail. Je voudrais aussi remercier l'équipe du Cerema: Constance, Annie, Jordan, Stéphane, Arnaud, Alain, Frédéric, Alexandre, Sébastien, Philippe, Julien mais aussi Sylvain, Joël, Samuel, Sophie, Hervé, Clément, Jean-Louis et Ahmed pour leur accueil et leur sympathie. J'ai adoré partagé ces nombreuses discussions aussi variées les unes que les autres avec vous. Nos parties de Sudoku me manqueront! Je remercie aussi toute l'équipe du Laris pour leur accompagnement notamment Abderafi CHARKI. Un énorme merci à toi, Hassen RIAHI, pour ton apport qui a été si cruciale

durant la dernière ligne droite.

Je tiens aussi à remercier tous les doctorants et les post-doctorants du LARIS avec qui j'ai tissé de nombreux liens d'amitié. Merci à Khaoula, Achraf, Julien, Ken, Khadim, Alejandro, Amine, Nadar, Lamia, Nihad, Ibrahim, Rodrigue, Alban, Sekou, Slim, Rémi, pour avoir partager cette grande palette d'émotions que peut ressentir un doctorant durant ses travaux.

Je voudrais aussi remercier mes amis, ma seconde famille! Mes copines de toujours, Rafatou, Ayoko, Yempab et Mila, qui me soutiennent depuis bientôt dix ans. Merci pour vos conseils les filles, vos encouragements et surtout pour votre patience. Je sais que je n'ai pas toujours été facile. Merci aussi à toi Daniel pour ta disponibilité et pour ce magnifique pôt que tu as organisé! Un grand merci aussi à Merlyne, Malick et Franck. La thèse c'est aussi des rencontres, des rencontres qui nous marquent. Merci Cendrine Mercier pour ta sympathie, ta joie de vivre et pour la personne que tu as été tout au long de cette aventure. Je n'oublierai jamais toutes ces soirées que nous avons partagées au James et au Carré. Puis à toi Rino, je voudrais dire merci! Tout simplement.

Merci à toute ma famille et particulièrement à papa et à maman, pour votre soutien inconditionnel. J'espère que cette thèse sera à la hauteur des efforts et sacrifices que vous faites au quotidien pour nous. Un merci également à mes frères et sœurs Fayçal, Marzukh, Rawya et Fléra pour toutes nos chamailleries! Je voudrais aussi remercier mes grands parents Fléra, Nastasia et Lantam, principalement pour leur soutien moral durant les mois qui ont précédés la soutenance. Merci à mes cousines Hassana & Housseyna, Faydatou, Akem, Meheza, Farida et à mes cousins Izzu-deen et Palakyem pour leurs encouragements quotidiens. Un merci aussi à ma nièce Zahra, qui a fait le déplacement spécialement pour ma soutenance du haut de ses sept mois.

Que les éventuels oubliés me pardonnent...

Je ne pourrai terminer ces remerciements sans penser aux machines de calculs du Cerema, qui ont tournés sans relâche, nuits et jours, ces trois dernières années, pour que ces travaux puissent être livrés dans les meilleurs délais.

Fally Titikpina
À Angers, le 01 Février 2017.

Table des matières

Table des matières	9
1 Introduction générale : Vers une démarche de garantie de performance	13
1.1 Enjeux environnementaux	14
1.2 Problématique	15
1.2.1 Positionnement industriel	16
1.2.1.1 Performance prévue / Performance effective	16
1.2.1.2 La garantie de performance énergétique	17
1.2.2 Objectifs scientifiques	19
1.3 La nécessité d'un plan de mesure et de vérification	22
1.3.1 Principe et fondements	22
1.3.2 Ajustement contractuel	25
1.3.3 Caractérisation de la performance	26
1.4 Construction du modèle	26
1.4.1 La simulation thermique dynamique	26
1.4.2 Couplage physico-probabiliste	27
1.4.3 Déclinaison des modèles de simulation	28
1.5 Objectifs de la thèse	29
2 De la mesure aux données d'entrée des modèles	31
2.1 Un besoin de fiabilité dans les mesures	32
2.1.1 Les différentes sources d'incertitudes	32
2.1.2 Caractérisation d'une mesure : notions d'erreurs systématique et aléatoire, de justesse et de fidélité	33
2.1.3 Méthodes d'estimation d'incertitude : La méthode GUM	35
2.1.3.1 Évaluation type A de l'incertitude	35
2.1.3.2 Évaluation type B de l'incertitude	37
2.1.4 Statistiques et échantillonnage	37
2.1.5 Différence entre erreur et incertitude	39
2.1.6 Intervalle élargi et probabilité associée	39
2.1.7 Variables statiques et variables dynamiques	40
2.1.8 Enregistrement, stockage et prétraitement des données	41
2.1.9 Présentation du cas d'étude	42

2.2	Quantification des incertitudes des grandeurs d'entrée	42
2.2.1	Incetitudes liées aux sollicitations extérieures	44
2.2.1.1	Décomposition de l'incertitude	45
2.2.1.2	Estimation des incertitudes liées aux sollicitations extérieures dans le cas d'un niveau haut (mesures complètes)	45
2.2.1.3	Estimation des incertitudes liées aux sollicitations extérieures dans le cas d'un niveau bas (aucune mesure)	49
2.2.2	Incetitudes liées au renouvellement d'air	49
2.2.2.1	Débits de ventilation	51
2.2.2.2	Efficacité des échangeurs double flux	54
2.2.2.3	La perméabilité à l'air du bâtiment	58
2.2.3	Incetitudes liées aux sollicitations internes	59
2.2.3.1	La température intérieure	60
2.2.3.2	L'occupation	61
2.2.3.3	Les apports internes	64
2.3	Conclusions et limitations	66
2.3.1	Analyse des résultats	66
2.3.2	Conclusion	68
3	Consolidation et réduction des modèles	71
3.1	Analyse de sensibilité	72
3.1.1	Méthodes d'analyse de sensibilité	72
3.1.1.1	Analyse de sensibilité locale et globale	72
3.1.1.2	Analyse de sensibilité par screening	73
3.1.2	Criblage par la méthode de Morris	74
3.2	Propagation d'incertitudes	76
3.2.1	Méthode de propagation d'incertitude proposée par le GUM	76
3.2.2	Méthode de propagation d'incertitude par Monte Carlo	77
3.3	Réduction de modèle et construction d'un méta-modèle	79
3.3.1	Utilisation d'un méta-modèle : Les polynômes du chaos	80
3.3.2	La méthode de régression	82
3.3.3	Construction d'un plan d'expériences	84
3.3.3.1	Le Monte Carlo Classique	85
3.3.3.2	Le Latin Hypercube	85
3.3.3.3	Les suites de Halton	86
3.3.3.4	Les suites de Sobol	88
3.3.3.5	Comparaison et choix de la méthode optimale	88
3.3.4	Post-traitement des coefficients	89
3.3.5	Estimateurs d'erreur et convergence	91
3.3.5.1	Erreur relative	91
3.3.5.2	Leave One Out error	92
3.3.5.3	Comparaisons quantitative et qualitative	93

3.3.6	Algorithme général de construction d'un méta modèle	93
3.4	Modélisation du bâtiment test sous Trnsys	94
3.4.1	Construction du modèle	94
3.4.1.1	Hypothèses de modélisation	95
3.4.1.2	Zonage du bâtiment	96
3.4.1.3	Échanges thermiques avec le sol	97
3.4.1.4	Constitution des parois et ponts thermiques	97
3.4.1.5	Infiltrations	97
3.4.1.6	Les systèmes de ventilation	99
3.4.1.7	Les masques solaires	100
3.4.1.8	Initialisation de la simulation	100
3.4.2	Comparaison Calcul / Mesures	101
3.4.2.1	Signature énergétique	102
3.4.2.2	Indicateurs de performance d'un modèle	103
3.5	Analyse de sensibilité et réduction de modèle sur le bâtiment test	105
3.5.1	Algorithme de simulation pour le couplage physico-probabiliste	105
3.5.2	Identification des variables influentes	107
3.5.3	Réduction du modèle de simulation a posteriori	111
3.5.3.1	Identification du degré optimal	111
3.5.3.2	Comparaison qualitative et quantitative	112
3.5.3.3	La non régularité du modèle et l'effet de seuil	116
3.6	Conclusion du chapitre	117
4	Optimisation du plan de mesure	119
4.1	La place des MV au sein des contrats de performance énergétique	120
4.1.1	La procédure de contractualisation	120
4.1.2	Équilibre entre coût et précision	121
4.2	Méthodologie d'optimisation des plans de mesure	121
4.3	Cas d'étude	125
4.3.1	Utilisation des indices de sensibilité ξ_i de Morris	126
4.3.2	Méthodologie d'optimisation appliquée au modèle Trnsys a priori	126
4.3.3	Intégration des incertitudes des paramètres statiques	128
4.3.4	Ajustement de la performance	129
4.3.5	La dimension stochastique effective	133
4.4	Conclusion du chapitre	134
5	Conclusion et Perspectives	135
6	Annexes	141
6.1	Données météorologiques utilisées pour l'estimation des incertitudes au niveau bas (aucune mesure)	142
6.2	Table de Khi deux	145

6.3	Incertitudes des paramètres statiques	146
6.4	Modélisation du bâtiment test	148
6.5	Instrumentation du bâtiment test	151
6.5.1	Instrumentations in situ	151
6.5.2	Instrumentation de la CTA Bureaux	151
6.6	Consolidation et réduction des modèles	154
Acronymes		157
Liste des tableaux		159
Table des figures		161
Bibliographie		165



1

Introduction générale : Vers une démarche de garantie de performance

Depuis le choc pétrolier de 1973, la maîtrise des consommations énergétiques est devenue une problématique planétaire. Compte tenu de la raréfaction des énergies fossiles, de l'augmentation sans cesse croissante du prix des énergies mais aussi du réchauffement climatique, il est devenu essentiel de réduire la demande énergétique des secteurs les plus énergivores. D'après l'Agence Internationale de l'Énergie (IEA), en poursuivant le rythme actuel des consommations, les réserves de pétroles se tariront d'ici une cinquantaine d'années, les réserves de gaz d'ici une soixante et celles de charbon dans 120 ans alors que la demande énergétique est en perpétuelle croissance (IEA, 2015).

Le secteur du bâtiment est de loin le premier poste de dépense énergétique au sein de l'union européenne : en 2013, il représente 40% de la consommation totale d'énergie finale et est responsable de 24 % des émissions de CO_2 . En France, la part du secteur du bâtiment représente 45% de l'énergie consommée devant les transports (33%), l'industrie (16%) et l'agriculture (3%) (Chiffres clés de l'énergie, 2016). Il est aussi responsable de plus de 22% des émissions de gaz à effet de serre (ADEME, 2013). De plus, compte tenu de l'accroissement continu du parc immobilier, du développement démographique, de l'exigence de confort des usagers et surtout du développement technologique, cette consommation est en accroissement continu et constitue un axe de progrès prioritaire qui est notamment soutenu par les politiques publiques.

1.1 Enjeux environnementaux

De manière générale, un bâtiment est une structure construite par un assemblage de matériaux de construction et un cadre de vie et d'activités pour ses occupants. Il doit donc pour répondre aux besoins de ces derniers, être fonctionnel tout en leur apportant un certain confort thermique, visuel, acoustique mais aussi olfactif ; tout ceci en respectant une qualité d'air et d'hygiène des espaces qu'il propose. Dans ce sens, le bâtiment doit pouvoir trouver un équilibre entre la satisfaction des besoins des usagers et la limitation des consommations énergétiques. En moyenne, en France, un bâtiment consomme $204 \text{ kWh/m}^2/\text{an}$ sur ses différents postes de consommation qui sont : le chauffage, la production d'eau chaude, l'éclairage, ou encore la climatisation ([ADEME, 2013](#)).

Dès lors, plusieurs politiques énergétiques ont vu le jour ces dernières années tant sur le plan international que national. Le 30 Novembre 2015 se tenait à Paris, le sommet international sur le climat : la COP 21. Il y a été confirmé par le Groupe d'experts Intergouvernemental sur l'Évolution du Climat (GIEC) une hausse des températures d'environ $2,5^\circ\text{C}$ d'ici 2100 : le réchauffement climatique devient inévitable. Un accord a donc été proposé (l'Accord de Paris) qui fait suite à plusieurs autres accords internationaux dans cette démarche de transition énergétique notamment le protocole de Kyoto en 1997. La lutte contre le changement climatique est donc devenue une priorité.

La France s'est aussi engagée sur le plan national à diviser par quatre ses émissions de gaz à effet de serre d'ici 2050, à travers la loi POPE (Programme fixant les Orientations de la Politique Énergétique) en 2005. Pour tenir cet engagement, des mesures d'amélioration de l'efficacité énergétique doivent être mises en place ou du moins adoptées au niveau des trois grands secteurs de dépense énergétique à savoir le bâtiment, les transports et l'industrie. Étant le premier poste de dépense, le secteur du bâtiment est donc placé au premier rang et permet comparativement aux autres, la mise en place de projets à fort impact énergétique et environnemental induisant cependant des investissements moins élevés.

Les lois Grenelle Environnement I & II, les réglementations thermiques 2005, 2012 ont ainsi vu le jour. Elles proposent des objectifs précis certes, pour la réduction des consommations énergétiques des bâtiments neufs ainsi que pour le parc existant, mais ambitieux pour l'horizon 2020. Pour les bâtiments neufs, elle vise une réduction de la consommation des bâtiments neufs à $50 \text{ kWh/m}^2/\text{an}$ ainsi qu'une réduction globale de la consommation du parc existant de 38% ; tout cela en préservant un bouquet énergétique équilibré, faiblement émetteur de gaz à effet de serre et contribuant à l'indépendance énergétique nationale ([Réglementation thermique, 2012](#)).

Pour atteindre ces objectifs, les consommations énergétiques des bâtiments construits ou réhabilités doivent observer des diminutions notoires : ces bâtiments ont été qualifiés de basse consommation (BBC). Un travail conséquent est donc effectué en amont de la

construction du bâtiment, lors de la conception pour qu'ils puissent atteindre les performances prévues. Un bâtiment BBC combine ainsi une conception optimisée de sa structure, de ses équipements, de son isolation, pour limiter efficacement les déperditions au niveau de l'enveloppe du bâti mais aussi pour pouvoir récupérer et réutiliser les flux d'énergie sortants.

1.2 Problématique

En phase de conception d'un bâtiment, des études sont réalisées pour dimensionner les équipements nécessaires chargés d'assurer le confort des occupants, tout en respectant les exigences réglementaires mais aussi pour en estimer la performance future. La consommation énergétique d'un bâtiment est alors estimée sous divers scénarios d'usage et lorsqu'il est sous le label BBC, il est considéré comme consommateur de peu d'énergie. Dans le cadre des engagements du protocole de Kyoto, la directive 2002/91/CE européenne de 2002 ([Communauté, 2002](#)) proposa comme définition de la performance énergétique :

La performance énergétique d'un bâtiment correspond à la quantité d'énergie mesurée ou calculée dans le cadre d'une utilisation normale du bâtiment. Elle inclut notamment l'énergie utilisée pour le chauffage, l'eau chaude sanitaire, le refroidissement (éventuellement), la ventilation et l'éclairage.

Un des indicateurs de la performance d'un bâtiment est donc sa consommation énergétique. Elle résume en elle le comportement du bâtiment au fil du temps mais aussi la réponse des équipements techniques aux diverses sollicitations imposées par les occupants et par les variations météorologiques.

Le Centre Scientifique et Technique du Bâtiment (CSTB) a développé une méthode de calcul réglementaire (TH-BCE pour les bâtiments neufs ou TH-CE EX pour les bâtiments existants) ([CSTB, 2012, 2008](#)), dans laquelle sont proposés des valeurs standards pour les variables usuelles et des scénarios conventionnels d'occupation et d'exploitation (température de consigne, apports métaboliques liés à l'occupation, ...). Ces scénarios conventionnels sont disponibles par typologie de bâtiment et permettent de faire un choix suivant l'objectif de l'étude et d'estimer la consommation conventionnelle d'énergie du bâtiment. Ces scénarios ne permettent donc pas de faire une estimation de la consommation réelle mais plutôt une évaluation de cette dernière pour orienter les choix de rénovation pour déterminer ensuite l'économie réalisée grâce aux travaux effectués par rapport à la situation antérieure.

Il est cependant difficile de juger de la fiabilité de ces scénarios par rapport à la spécificité de chaque cas : les valeurs fournies sont alors entachées d'incertitudes pour plus de représentativité. Pour évaluer cette performance, des mesures sont mises en place pendant les

premières années d'exploitation. Elles permettent de confronter les consommations d'énergie prédites à celles réellement perçues lorsque le bâtiment est occupé. Cependant, il a été constaté sur bon nombre de bâtiments instrumentés que la consommation d'un bâtiment après construction diffère de celle prévue au préalable lors de la conception (PREBAT, 2015). Les conditions d'usage constatées sont différentes des scénarios réglementaires et présentent surtout des variabilités importantes. Afin d'éviter ces désagréments, des études sont effectuées pour en déceler les causes mais aussi dans le but de sécuriser les fournisseurs d'énergie et les maîtres d'ouvrage : cette sécurité se traduit par la mise au point de contrat de garantie de performance.

1.2.1 Positionnement industriel

La consommation énergétique réellement perçue par les usagers d'un bâtiment est généralement le double ou le triple (De Wilde, 2014) de celle prédite avant sa construction. Compte tenu des challenges environnementaux et de l'attente des clients pour bénéficier de bâtiments de plus en plus performants, il devient indispensable de pouvoir réduire cet écart. Cette différence existante réduit la crédibilité des professionnels du bâtiment et conduit au scepticisme de la population et des services publics vis à vis des bâtiments à haute performance énergétique et des politiques environnementales adoptées.

1.2.1.1 Performance prévue / Performance effective

Les causes de l'écart entre la performance effective et celles prévue ont été recensées et peuvent être répertoriées suivant trois phases durant le cycle de mise en œuvre d'un bâtiment, selon De Wilde : celles dues à l'étape de conception, à l'étape de construction et celles qui interviennent pendant l'exploitation du bâtiment (De Wilde, 2014; Sun, 2014). Ces causes varient bien sûr d'un bâtiment à un autre et peuvent être spécifiques à la typologie de bâtiment.

En phase de conception, les modélisateurs en bureaux d'études modélisent le bâtiment pour avoir une représentation en conditions réelles et ainsi pouvoir prendre les décisions adéquates. Ils assignent le plus souvent des valeurs déterministes aux variables prises en considération et il est souvent difficile de quantifier la fiabilité des informations sur lesquelles ils se basent. Un mauvais dimensionnement des équipements peut par exemple avoir des incidences sur le chauffage ou la distribution aéraulique du bâtiment.

Les modèles en eux-mêmes peuvent aussi présenter des manquements (dus aux simplifications et aux éventuelles hypothèses) et même en ayant un modèle qui représente le plus fidèlement possible le bâtiment à construire, on est toujours confronté à des incertitudes liées à la prédiction de l'inconnu (Gorse et al., 2016). Marija Trcka et al. expliquent dans leurs travaux (Trčka and Hensen, 2010) qu'à partir d'un certain niveau de détail, le modèle

de simulation devient de plus en plus complexe. Le nombre de variables augmente et la quantité d'informations dont a besoin le modèle aussi. Les incertitudes sur les données d'entrée s'accumulent et contribuent à accroître les incertitudes en sortie du modèle.

Les retours d'expérience fournis par l'Agence Qualité Construction (AQC) (AQC, 2011) démontrent que des défauts et malfaçons interviennent souvent pendant la phase de construction, et dégradent la qualité du bâtiment. Les épaisseurs d'isolants peuvent par exemple ne pas être respectées, des composants non prévus peuvent être ajoutés et non notifiés, ce qui peut accroître les déperditions au niveau de l'enveloppe du bâtiment. Ces disparités sont difficiles à quantifier et à évaluer car les parois sont constituées de plusieurs couches de matériaux isolants. Elles peuvent aussi survenir pendant l'installation des équipements techniques : les bouches de ventilation peuvent par exemple être déplacées ou la tuyauterie mal raccordée.

En phase d'exploitation, les conditions opérationnelles du bâtiment ainsi que l'efficacité de la maintenance qui sera effectuée ne peuvent pas être prévues avec certitude. Le taux d'occupation d'un bâtiment est très difficile à quantifier et le comportement des habitants encore plus à prédire (Caucheteux et al., 2013; Menezes et al., 2012; Hoes et al., 2009). Les scénarios considérés en phase de conception sont établis avec des hypothèses qui malheureusement sont souvent très loin de la réalité. Le développement technologique induit aussi que les charges électriques prévues initialement sont sous-estimées par rapport aux charges réelles. Les systèmes peuvent être mal réglés et l'occupant peut parfois en faire un usage excessif. Des projets ont été menés dans le but d'avoir des retours d'expérience tant sur le plan national qu'international. Pour plus de détails, le lecteur peut se référer aux documents suivant : (PREBAT, 2015; VAD, 2013; BOUGRAIN, 2012; ENERTECH, 2012; RAGE, 2012; AQC, 2011).

1.2.1.2 La garantie de performance énergétique

Au vu de ces retours d'expérience et en se basant sur la définition de la performance énergétique proposée dans la section 1.2, il a été proposé en France, sous forme contractuelle, en Avril 2012 (par un groupe de travail composé de professionnels du bâtiment et de l'immobilier, d'assureurs et de juristes), une définition de la garantie de performance énergétique (GPE) entre les entreprises du bâtiment, de l'énergie et leurs clients (GPE, 2013).

Cette garantie constitue un moyen légal (Directive européenne 2012/27/UE) obligeant le prestataire de services à atteindre les objectifs d'efficacité fixés lors de la livraison des bâtiments. On passe alors d'une obligation de moyens à une obligation de résultats. Ces résultats peuvent concerner les performances intrinsèques des matériaux de construction avant la réception des travaux, on parle alors de garantie de performance énergétique

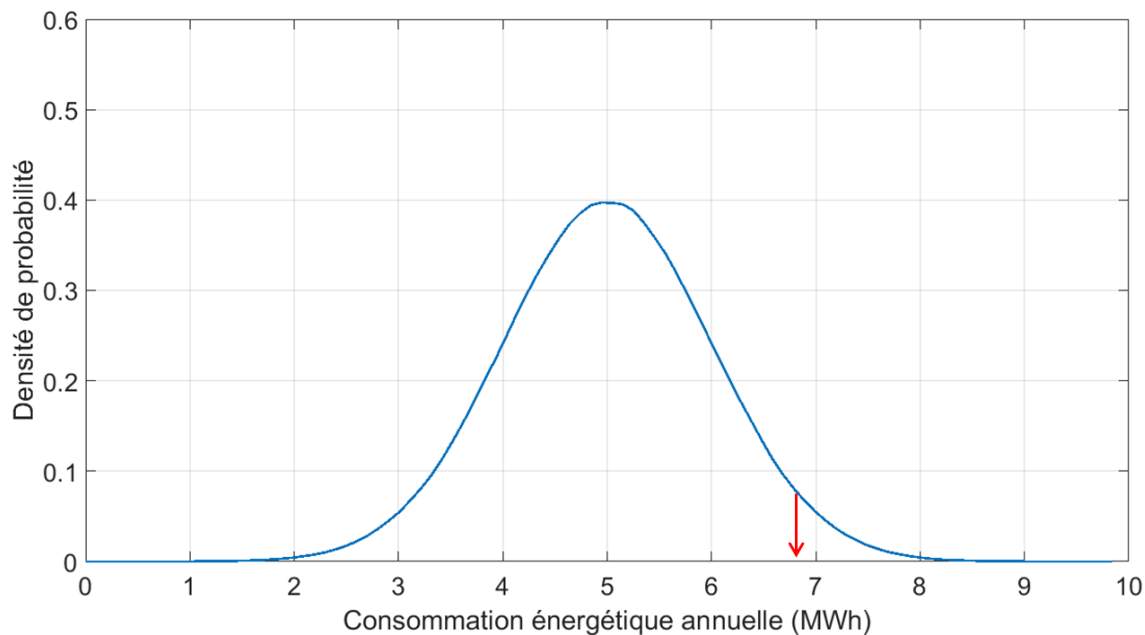


FIGURE 1.1 – Exemple de performance annuelle garantie par un contrat de performance énergétique

intrinsèque (GPEI) ou plutôt porter sur les performances énergétiques du bâti en phase d'exploitation, il s'agit dans ce cas d'une garantie de résultats énergétiques (GRE).

La performance prédite en phase de conception peut donc être fournie avec une marge d'erreur ou une plage de variation à l'intérieur de laquelle la performance réelle du bâtiment a une forte probabilité de se trouver. Elle ne doit donc plus être une estimation ou être traitée comme une valeur déterministe mais plutôt sous forme de probabilité d'obtenir cette valeur. Elle serait dans ce cas garantie avec un niveau de risque associé et permettrait de tolérer une marge due à la présence d'incertitudes dans les prédictions, dans les mesures ou dans l'analyse de donnée effectuée. La performance estimée en phase de conception et celle mesurée en phase d'exploitation peuvent alors être représentées sous la forme de lois de distributions.

Pour illustrer la démarche, considérons l'exemple présenté dans la figure 1.1. La courbe représente toutes les valeurs que sont susceptibles de prendre la consommation énergétique annuelle (en considérant l'ensemble des aléas) du bâtiment sous contrat de garantie de performance. L'entreprise qui s'engage à contractualiser la performance du bâtiment estimera alors que la probabilité est plutôt forte que la consommation énergétique sur une période donnée soit de 5 MWh, mais garantira une performance avec 5% de risque, soit 6.96 MWh compte tenu de tous les aléas cités précédemment. Au delà de 6.96 MWh, elle s'engage à ce que le client soit dédommagé puisque les termes du contrat n'auront pas été respectés.

1.2.2 Objectifs scientifiques

Sur le plan scientifique, l'amélioration de la performance des bâtiments est au cœur des débats. La communauté scientifique (SIMUREX sur le plan national et IBPSA sur le plan international) a compris l'enjeu que représentait la garantie de performance pour atteindre les objectifs environnementaux.

Des travaux sont menés dès la phase de conception pour améliorer la qualité des bâtiments construits. Kibert C. J. (Kibert, 2016) expose les initiatives menées pour disposer de bâtiments de plus en plus performants sur le plan énergétique, environnemental mais aussi dans l'amélioration du confort des occupants. De nombreux efforts sont mis en œuvre pour aider les bureaux d'étude à prendre de meilleures décisions dès la phase de conception. Pour bon nombre d'entre eux, c'est à cette étape du cycle de vie du bâtiment qu'il est important de concentrer les efforts pour arriver à des performances effectives initialement prévues. Echenagucia (Echenagucia et al., 2015) propose d'aider les concepteurs, par une aide à la décision multi-critère, en leur fournissant dès le départ toutes les informations dont ils ont besoin pour le bon dimensionnement des enveloppes. D'autres chercheurs comme Huang Yu, Sun et Trcka (Huang et al., 2015; Yu et al., 2015; Sun et al., 2014; Trčka and Hensen, 2010) misent sur le bon dimensionnement des équipements de ventilation, de chauffage. Pour eux le fait de se baser seulement sur les moteurs de calculs réglementaires, ne permet pas de réaliser des choix optimaux. Il faudrait pouvoir remonter aux bureaux d'études des données de retour d'expérience pour qu'ils puissent anticiper les risques de baisse de performance.

On comprends très vite la nécessité de pouvoir maîtriser les paramètres les plus susceptibles d'agir sur le bon fonctionnement du bâtiment tel qu'il a été prévu en phase de conception (Hopfe and Hensen, 2011). Les deux principaux aspects qui ne peuvent absolument pas être maîtrisés sont les variables météorologiques et le comportement de l'occupant (Anstett-Collin et al., 2015; de Wilde and Tian, 2012; Menezes et al., 2012). Une des principales difficultés est de pouvoir quantifier la variabilité associée à ces entrées dès la phase de conception. La maîtrise de cette variabilité permettrait de mieux appréhender les possibles scénarios susceptibles de se produire et ainsi de concevoir des bâtiments de plus en plus robustes (Anstett-Collin et al., 2015; Tian and De Wilde, 2011).

L'impact de la variabilité climatique inter-annuelle sur la performance énergétique et sur l'adaptabilité des bâtiments aux différents rigueurs climatiques a ainsi fait l'objet de beaucoup d'attention ces dernières années. En France, les travaux de Goffart (Goffart, 2013) ont permis d'étudier l'impact de la variabilité des données météorologiques sur la performance des bâtiments. En concordance avec elle sur le plan international, il tend à se développer des approches de type probabilistes (O'Neill and O'Neill, 2016; Oldewurtel et al., 2012; Tian and De Wilde, 2011; Crawley, 2008) pour pouvoir recréer des fichiers météorologiques plus robustes adaptés à l'emplacement réel d'un bâtiment.

Aussi, la performance du bâtiment pourrait tout à fait décroître tout au long de son cycle de vie (Aksoezen et al., 2015). Il tend à se développer des bâtiments intelligents (Wong et al., 2005; Nguyen and Aiello, 2013) capables de s'adapter au changement climatique, au comportement des usagers mais aussi à la baisse probable de la performance des systèmes. Ces bâtiments sont aussi capables de s'imprégner du milieu et d'ajuster si nécessaire la performance (Yang and Peng, 2001).

En ce qui concerne les modèles d'occupation, les études ont été plus poussées. Elles s'étendent depuis les procédés de mesure au développement de modèles d'estimation de l'occupant pour la création de modules d'occupation adéquats ainsi que leur intégration au cœur des modèles de simulation (Yan et al., 2015). On s'intéresse notamment à la quantification du nombre d'occupants (Amayri et al., 2016; Tahmasebi and Mahdavi, 2016; Mahdavi and Tahmasebi, 2015; Duarte et al., 2013) mais aussi à la maîtrise de leur comportement (Kashif et al., 2013; Peuportier et al., 2013; Azar and Menassa, 2012; Fabi et al., 2012; Menezes et al., 2012; Page et al., 2008) au sein d'un bâtiment. Néanmoins, l'occupant étant en perpétuelle interaction avec son environnement (les portes, les fenêtres, le thermostat, l'utilisation des équipements électriques, ...), il est difficile de mettre sur pied des modèles robustes (Ryu and Moon, 2016).

Certains facteurs comme la qualité de l'air intérieur (Teodosiu et al., 2003), les nuisances sonores provenant de l'extérieur ou le manque d'intimité peuvent emmener un occupant à manipuler ses ouvrants (D'Oca et al., 2014; Warren and Parkins, 1984) ou à agir sur les systèmes de chauffage, de ventilation et de climatisation (CVC) pour améliorer son confort. Masoso (Masoso and Grobler, 2010) a pu démontrer que dans les bâtiments tertiaires, plus de 50 % de la consommation énergétique était engagée pendant les heures où le bâtiment était censé ne pas être occupé. Ce genre de comportement est imprévisible dans les modèles de simulation courants qui ont tendance à modéliser le comportement de manière complètement statique tout au long de l'étude. L'idéal serait d'avoir une interaction dynamique dans les modèles de simulation de telle sorte que les scénarios d'occupation puissent se réajuster en fonction du confort qui règne au sein du bâtiment. Des efforts restent à fournir sur ce plan.

Cela dit, pour construire un modèle d'occupation, il est nécessaire de disposer d'un nombre important de données sur les facteurs qui interagissent avec l'homme ou sur des données d'occupation passées. Les procédés de mesures sont souvent complexes (Amayri et al., 2016; Caucheteux et al., 2013) et aucune précision n'est fournie quand à leur fiabilité (Yan et al., 2015). Ces données permettent d'établir des modèles stochastiques en se basant sur les distributions du taux de présence des occupants (Tahmasebi and Mahdavi, 2016; Chang and Hong, 2013); certains modèles proposent notamment de prendre en compte la dimension spatio-temporelle (Page et al., 2008). Il tend aussi à se développer des

modèles de prévision de charge alliant des stratégies d'optimisation énergétiques adaptées aux contraintes d'exploitation (Berthou, 2013; Robillart, 2015)

Sur un tout autre volet, après avoir mené des actions de performance énergétique, il est nécessaire de pouvoir évaluer la quantité d'énergie économisée. Celle-ci est généralement estimée par la différence entre la consommation mesurée après les actions de performance énergétique et celle qui aurait dû être constatée durant cette même période. Cette dernière ne peut effectivement pas être mesurée directement à cause de la dynamique incertaine de certaines entrées telles que la météo ou l'occupation.

Pour ce faire, on utilise soit des modèles physiques comme ceux qu'on utilise à l'étape de conception, soit des modèles à base de données. Une partie de ces modèles, comme les réseaux de neurones artificiels (ANN) (Benedetti et al., 2016; Magnier and Haghighat, 2010; Neto and Fiorelli, 2008), les réseaux bayésiens (O'Neill and O'Neill, 2016) permettent usuellement d'estimer la variable d'intérêt à savoir la consommation énergétique en fonction des variables d'entrée. On retrouve aussi des modèles inverses, comme l'a notifié l'ASHRAE (ASHRAE, 2013) qui prennent en entrée des données de consommation énergétique provenant de retours d'expérience ou de mesures antérieures sous forme d'a priori pour estimer la meilleure valeur des paramètres d'entrée (Heo and Zavala, 2012; Lee and Braun, 2004).

Pour obtenir un niveau d'incertitude acceptable, en phase de conception ou dans le cadre d'actions de performance énergétique, un assez large panel de données mesurées est nécessaire pour fiabiliser les données d'entrée des modèles, ce qui rend le processus de mesures et de vérification coûteux. Les entreprises disposent généralement d'un budget limité pour les processus de mesure, la mise en place des capteurs, la collecte des données ainsi que leur analyse. *Le développement de nouvelles méthodes de modélisation permettant en partant d'un budget donné, la mise en place d'un plan de mesure garantissant des incertitudes maîtrisées sur la variable d'intérêt, fait partie des verrous scientifiques et industriels à lever.*

Burkhart et Heo ont développé dans leurs travaux (Burkhart et al., 2014; Heo et al., 2013; Heo and Zavala, 2012) un modèle basé sur les processus gaussiens, tenant compte de la corrélation spatio-temporelle entre les variables et permettant d'estimer l'incertitude sur plusieurs variables d'intérêt tout ceci dans un procédé de mesure itératif. Burman et al. (Burman et al., 2014) ont proposé un plan de mesure et de vérification basé sur la calibration d'un modèle de simulation thermique dynamique. La comparaison entre la performance effective constatée et celle estimée est donc réalisée sous des conditions d'exploitation identiques, ce qui peut aider à identifier les causes de l'écart qui sont inhérentes au bâti ou au fonctionnement des systèmes thermiques.

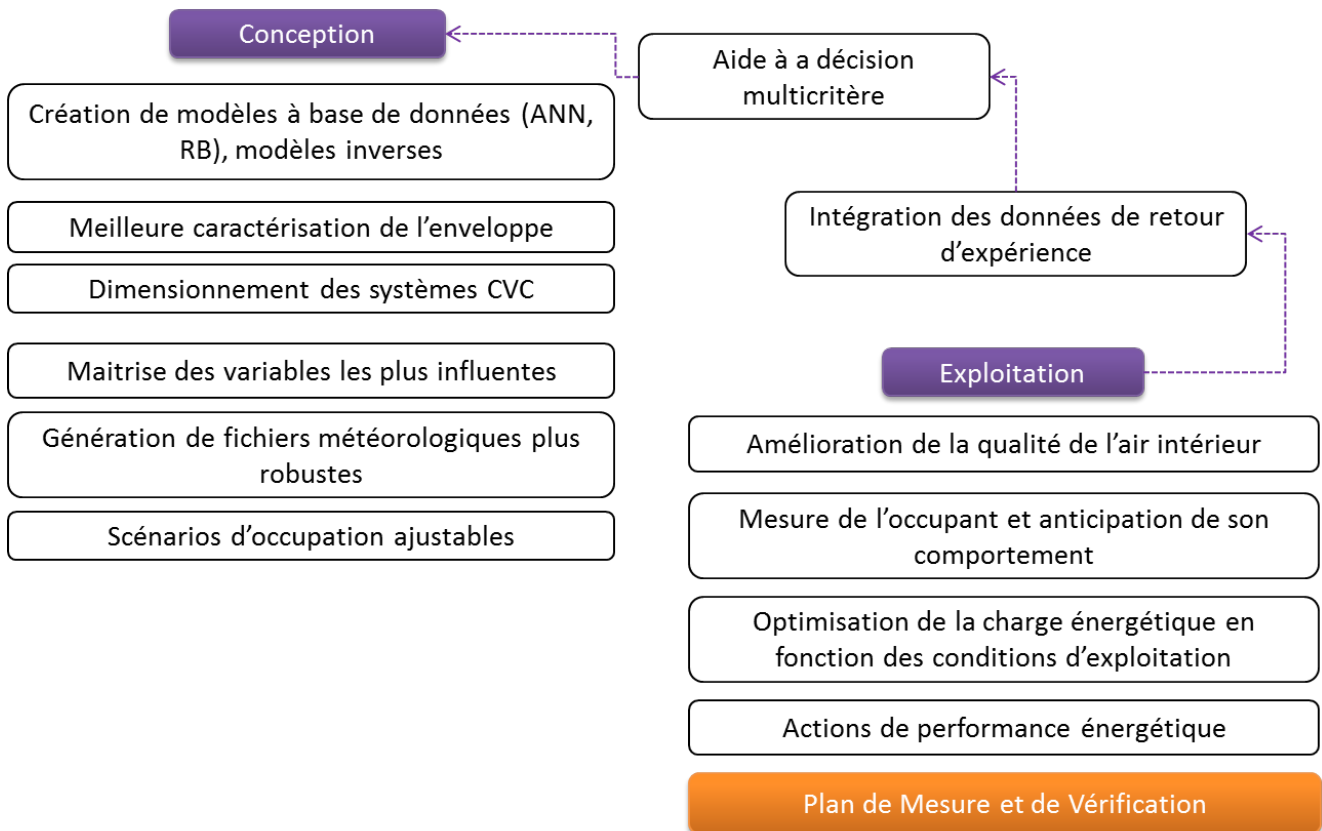


FIGURE 1.2 – Positionnement de la thèse par rapport aux différents champs d'actions d'amélioration de la performance énergétique dans le bâtiment.

Au final, les efforts mis en place pour améliorer la performance des bâtiments sont fournis tout au long de son cycle de vie (figure 1.2). Que ce soit à l'étape de conception ou durant son exploitation, toute l'industrie est mobilisée pour fournir des bâtiments plus robustes énergétiquement. Les champs d'actions sont variés (figure 1.2) et couvrent notamment l'optimisation des plans de mesure et de vérification.

1.3 La nécessité d'un plan de mesure et de vérification

1.3.1 Principe et fondements

Une méthodologie a été proposée en 1995 par le protocole international de mesure et de vérification de la performance (IPMVP) (IPMVP, 2010), qui vise à analyser l'utilisation de l'énergie pour déterminer les économies réalisées via des actions de performance énergétiques (APE). Cette méthodologie sert de base ou de support à la mise au point de tout contrat de performance énergétique.

Ces actions de performance peuvent porter sur l'exploitation du bâtiment (l'amélioration du confort des occupants, la maintenance des équipements techniques) ou sur le bâti

en lui même (composition de l'enveloppe, isolation thermique ou phonique), ou sur l'un et l'autre à la fois. Un plan de mesure et de vérification (MV) est ensuite mis en place et permet d'estimer les gains d'énergies lors du projet d'amélioration de performance, que ce soit dans le cadre d'une rénovation, d'un changement d'équipements ou tout simplement du changement dans le comportement des usagers. Les MV sont généralement effectuées par des propriétaires ou des agents des services publics soucieux de l'amélioration des performances des bâtiments.

Le plus souvent, elles sont sous forme de contrat de performance énergétique (CPE) entre deux ou plusieurs parties prenantes (Duplessis, 2008). Le but de ces contrats est de pouvoir garantir la performance des démarches effectuées dans le cadre du MV et des arrangements sont souvent prévus entre les différents signataires incluant des dédommagements financiers dans le cas où les objectifs ne seraient pas atteints. Ils apportent une certaine garantie au client sur les investissements potentiels fournis dans la démarche d'efficacité énergétique. Par exemple, les économies d'énergies réalisées sur les équipements du bâtiment et non sur les conditions d'occupation, seront versées sous forme de dédommagement financier au client.

Le but d'un MV, dans le cadre de l'IPMVP, est donc de pouvoir mesurer les économies d'énergies réalisées sur la performance énergétique dudit bâtiment, pendant une période donnée sur la base d'un contrat de performance énergétique (CPE). Ces économies correspondent à une différence de consommation énergétique entre une période de référence, pendant laquelle on a pu observer le comportement du bâtiment, et une période de suivi durant laquelle les MV seront mises en place et les économies réalisées. La méthodologie proposée par l'IPMVP consiste donc à analyser le comportement du bâtiment pendant la période de suivi et à en déterminer les économies d'énergie qui en découlent (figure 1.3). L'économie d'énergie réalisée pendant la période de suivi est donc la différence entre la consommation de référence ajustée et celle mesurée après les APE.

Ainsi, en fonction de la précision souhaitée, des pistes d'amélioration identifiées et du bâtiment étudié, on choisit de mettre en place un plan de mesure avec un degré de complexité adapté. Dans l'IPMVP (IPMVP, 2010), ont été déclinées quatre options de MV allant de la mesure de toutes les variables impliquées dans l'estimation de la performance à l'utilisation de compteurs d'énergie pour l'évaluation de la performance (table 1.1).

Cependant, mettre en place un plan de mesure et de vérification requiert un investissement en temps et financier conséquent, et généralement plus les mesures effectuées sont détaillées, plus on augmente en précision et en coût (table 1.1). Ainsi, depuis la prise de conscience de la nécessité d'améliorer la performance énergétique, les premières actions de performance ont été effectuées essentiellement sur l'isolation ou la performance intrinsèque du bâti. Rares sont celles qui ont été faites en effectuant un suivi détaillé sur les

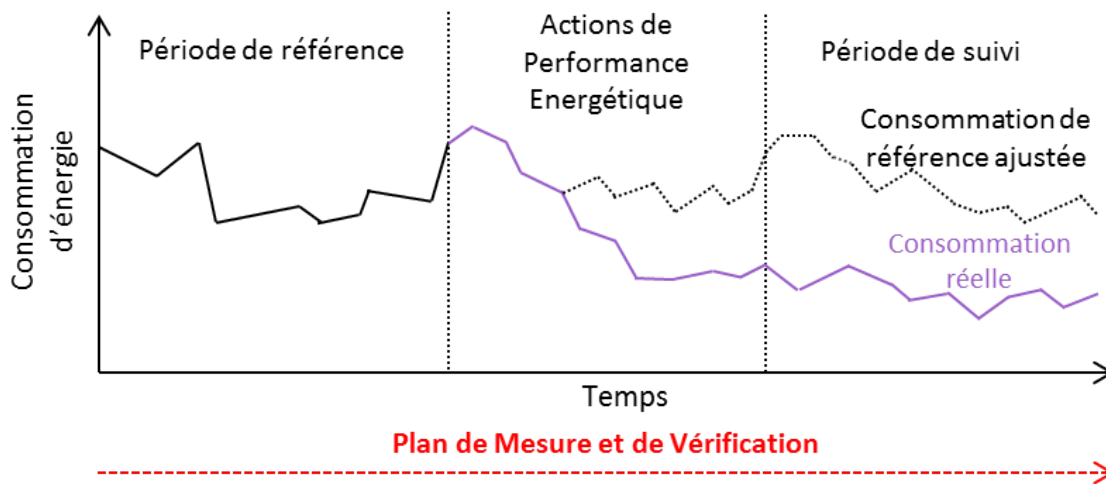


FIGURE 1.3 – Suivi de la performance énergétique du bâtiment basé sur la méthodologie proposée par l'IPMVP (IPMVP, 2010).

factures de consommation énergétique ou avec des mesures dans des conditions opérationnelles. L'ASHRAE (ASHRAE, 2013) préconise la mesure des variables qui sont susceptibles d'être modifiées après les actions de performance énergétique (variables d'ajustement) mais aussi celles qui peuvent potentiellement avoir une influence sur cette dernière (variables influentes).

TABLE 1.1 – Les quatre options de plan de mesure et de vérification déclinées par l'IPMVP

Option	Mesures mis en place	Méthode d'estimation des gains d'énergie
Option A	Mesure des grandeurs clés. Les grandeurs restantes sont estimées.	Les mesures sont effectuées sur de courtes périodes et la performance est estimée à l'aide des grandeurs mesurées.
Option B	Mesure de toutes les grandeurs.	Les mesures sont effectuées sur de courtes périodes et la performance est estimée à l'aide des factures de consommations calorifiques et électriques.
Option C	Mesure de toutes les grandeurs et à chaque niveau de détail.	Les mesures sont continues et la performance est estimée grâce aux mesures mises en place.
Option D	Performance estimée par un modèle de simulation.	Les consommations calorifiques et électriques sont mesurées et comparées pour estimer la performance du bâtiment.

Cela laisse supposer que dans un processus de mesure et de vérification complet (Option C de l'IPMVP), on effectue des mesures de consommation, de variables météorologiques ainsi que de l'occupation et de leur comportement pendant toute la période d'étude. L'ASHRAE recommande que la période soit suffisamment représentative pour pouvoir caractériser le comportement du bâtiment dans toutes les circonstances. Il faut donc disposer de beaucoup de données de mesure pour s'assurer de la qualité des résultats et affiner les estimations. Une fois les mesures effectuées, on peut établir un modèle de régression à partir des paramètres du modèle pour estimer la consommation énergétique ajustée pendant la période de suivi ainsi que l'économie d'énergie réalisée via les factures énergétiques comme représenté à la figure 1.3.

La première alternative pour optimiser le plan de mesure est de réduire le nombre de grandeurs à mesurer mais aussi la période de mesure (Option A ou B de l'IPMVP). Le fait est qu'en réduisant le nombre de points de mesure, l'incertitude sur les résultats augmente et en travaillant sur un bâtiment totalement instrumenté, les coûts de mesures peuvent devenir très élevés. *On optera alors pour l'option C, qui permettrait, en effectuant des mesures à un niveau de granularité élevée, d'estimer la performance avec une incertitude moins élevée mais à un coût très onéreux. L'ASHRAE (ASHRAE, 2013) recommande que le coût total du plan de mesure mis en place ne dépasse pas 10 % des économies d'énergie attendues.*

1.3.2 Ajustement contractuel

Pour effectuer une comparaison adéquate entre la période de référence et la période de suivi, la performance de référence $Y_{pré}$ qui ne peut pas être mesurée directement pendant la période de suivi, est déterminée par ajustement. Elle est notée $Y_{pré}^{aju}$.

On considère l'ensemble des entrées incertaines $X = \{X_1, X_2, \dots, X_N\}$. Supposons que les variables incertaines les plus influentes ont été mesurées pendant la période de référence $X_{pré}$ pour une performance énergétique de référence $Y_{pré}$. De manière analogue, ces mêmes variables ont été mesurées à la suite des actions de performance X_{post} pour une performance énergétique Y_{post} durant la période de suivi (figure 1.3).

L'économie d'énergie réalisée \hat{Y} durant toute la période de suivi est obtenue par :

$$\hat{Y} = \sum Y_{pré}^{aju} - \sum Y_{post} \quad (1.1)$$

Le principe d'ajustement, proposé dans les protocoles de mesure et de vérification existants (FBE, Octobre 2014; IPMVP, 2010; Heo and Zavala, 2012), permet de comparer dans des conditions identiques, la performance en amont des actions de performance énergétique

(APE) et celle de la période de suivi, en ajustant les variables qui ont été mesurées. Les variables d'ajustements sont sélectionnées parmi les variables mesurées en amont $X_{pré}$. Elles peuvent être des variables dynamiques dont l'évolution a été permanente durant toute la durée de l'étude (exemple : les rayonnements solaires) ou être statiques et être liées aux actions de performance énergétique qui ont été effectuées (exemple : débit de ventilation).

Pourra ainsi être considérée comme variable d'ajustement, toute variable mesurable après travaux et ayant une influence sur l'estimation de la consommation du bâtiment. Le choix des variables d'ajustement sera contractualisé. Les scénarios d'occupation, les températures de consigne et les variables météorologiques peuvent par exemple être considérés comme variables d'ajustement.

1.3.3 Caractérisation de la performance

En résumé, la procédure proposée par l'IPMVP permet d'évaluer l'efficacité des actions de performances énergétiques. Elle peut aussi servir de base pour l'évaluation de la performance énergétique dans d'autres contextes puisque cette étape fait partie intégrante des contrats de performance. Que ce soit dans une démarche d'efficacité où l'on chercherait à atteindre des niveaux de performances visées, ou dans le cadre d'opérations d'instrumentations ou de recherche, cette méthodologie peut servir d'appui à tout prestataire de service désirant évaluer la performance d'un bâtiment. C'est notamment sur cette base que sont fondés les travaux effectués dans cette thèse. Nous proposons une démarche, fondée sur l'IPMVP, pour l'identification d'un plan de mesure optimal dans le cadre de l'évaluation de la performance énergétique d'un bâtiment. Un plan de mesure et de vérification ne se limitera donc plus à l'évaluation des économies d'énergies suite à la mise en place d'actions de performance énergétique mais s'étendra aussi à l'évaluation de la performance dans un cadre général.

1.4 Construction du modèle

1.4.1 La simulation thermique dynamique

Dans le secteur du bâtiment et dans d'autres secteurs d'activité (en mécanique par exemple), on utilise la simulation thermique dynamique pour étudier l'évolution de systèmes physiques dans le temps. Ces simulations ont pour objectif de reproduire un ou plusieurs phénomènes physiques, mécaniques ou thermiques, de manière numérique et d'en étudier le comportement sous diverses sollicitations en effectuant une série de calculs à des pas de temps définis. Ainsi, lors de la phase de conception d'un bâtiment, les bureaux d'études utilisent fréquemment des outils de simulation thermique dynamique (STD) pour dimensionner le bâti et les divers équipements qui y seront installés en vue de répondre aux besoins des occupants mais également pour vérifier la conformité des choix qui ont été fait

par rapport aux normes en vigueur ([Réglementation thermique, 2012](#)). Il convient, préalablement à toute analyse, de construire un modèle robuste représentant le plus fidèlement possible la réalité. Ces outils de simulation sont aussi utilisés pour évaluer la performance du bâtiment dans une démarche de suivi ou de rénovation (Option D de l'IPMVP par exemple, confère table 1.1 de la section 1.3.1).

L'utilisateur a ainsi la possibilité de définir et d'interagir sur différents scénarios d'usage. Il peut aussi effectuer des variations autour des valeurs nominales des variables d'entrée en jeu pour en analyser les effets sur la réponse du modèle. Ces variations sont représentatives des incertitudes des variables d'entrée et permettent d'incorporer un attribut probabiliste au modèle physique construit.

1.4.2 Couplage physico-probabiliste

Les incertitudes présentes dans les variables d'entrées se traduisent sous la forme de probabilités qui doivent être prises en compte dans les études effectuées. Les simulations thermiques dynamiques qui sont faites à l'étape de conception d'un bâtiment considèrent que les variables d'entrée sont déterministes. Or le caractère déterministe ne permet pas de prendre en compte toute la plage de variation des variables et donc de représenter la réalité. Cette hypothèse de modélisation est sans nul doute une cause de l'écart qui existe entre la performance prédite et la performance effective.

Dans le but de prendre en compte le caractère stochastique des variables d'entrée au sein du modèle étudié, il est utile d'intégrer au modèle physique à construire, un attribut probabiliste. En effet, l'ensemble des études qui seront faites sur le modèle se feront sur la plage de variation complète des variables d'entrée considérées et tiendront aussi compte de la distribution qui leur sera assignée. L'utilisation d'un modèle physico-probabiliste permet ainsi à partir de la sélection des points représentatifs de chaque variable, d'estimer la variable d'intérêt en sortie du modèle ainsi que sa fonction de distribution (figure 1.4).

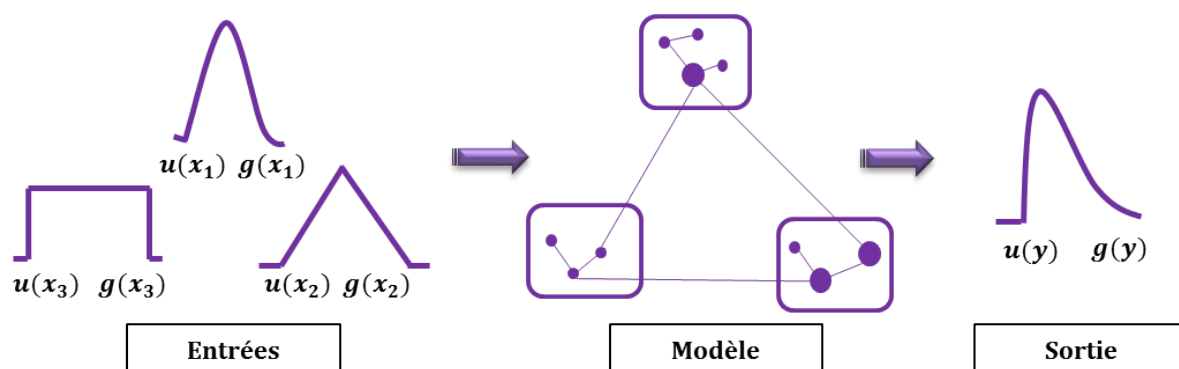


FIGURE 1.4 – Représentation physico-probabiliste d'un modèle de simulation. Chaque grandeur d'entrée x_i est représentée par son incertitude $u(x_i)$ et sa distribution $g(x_i)$, ce qui permet d'obtenir une incertitude $u(y)$ et une distribution $g(y)$ sur la sortie du modèle.

On constate par ailleurs, que les incertitudes sont présentes dans tous les cas de figure : que les variables d'entrées aient été estimées à partir de mesures complètes ou basées sur des connaissances antérieures, elles sont toujours entachées d'incertitudes. Comme a pu le souligner De Wilde dans ses travaux (De Wilde, 2014), la meilleure manière de prendre en compte ces incertitudes est évidemment de les traiter sous forme probabiliste et non déterministe. On obtiendrait ainsi des distributions en sortie de modèle, ce qui permettrait aux concepteurs ou aux industriels de prendre de meilleures décisions. Ces cinq dernières années, les chercheurs ont orienté leurs travaux en tenant compte de la dimension stochastique des variables et un aperçu exhaustif est proposé dans (O'Neill and O'Neill, 2016).

Pour illustrer ce concept, considérons la figure 1.5. La consommation garantie par le contrat de performance énergétique est de 15 MWh. Dans le premier cas de figure (courbe bleue), l'entreprise contractante a estimée suite aux simulations thermiques effectuées sur le modèle physico-probabiliste que la consommation calorifique annuelle serait de 10 MWh. Puis, dans le deuxième cas (courbe verte), une consommation calorifique annuelle de 12 MWh mais cette fois-ci avec une dispersion moins élevée. On constate qu'en garantissant une performance énergétique faible (courbe bleue) avec une incertitude plus élevée, la probabilité est plus forte de tomber sur des valeurs de consommation calorifique annuelle très élevées (> 15 MWh) et donc plus faible de respecter les termes fixés dans un contrat de garantie de performance. Dans le cas contraire, en optant pour une performance plus élevée mais entachée d'incertitudes moindres, on a moins de risque et donc une faible probabilité de dépasser les limites acceptables contractualisées.

Les incertitudes associées aux variables d'entrée jouent ainsi un rôle majeur sur le niveau de prestation offert dans les contrats de garantie. Il convient donc de bien les caractériser afin d'affiner la précision des résultats mais aussi de bien sélectionner le modèle physico-probabiliste à utiliser.

1.4.3 Déclinaison des modèles de simulation

Comme spécifié antérieurement, en phase de conception, les scénarios conventionnels sont utilisés pour estimer la performance future du bâtiment. Ces scénarios sont des cas de figure idéaux qui sont très souvent loin de la réalité. Aujourd'hui, de nombreuses bases de données concernant les usages constatés dans différents types de bâtiment, émergent grâce aux opérations de recherche diversifiées et aux retours d'expérience. Elles constituent une riche source d'informations sur laquelle peuvent s'appuyer les bureaux d'étude pour affiner leurs travaux. C'est dans cette optique que nous déclinons dans cette thèse deux modèles de simulation : *un modèle a priori et un modèle a posteriori*.

Le modèle a priori est le modèle construit et utilisé en phase de conception d'un bâtiment. Les scénarios d'usage et les données d'entrée des variables sont ceux issues de connaissances antérieures, d'opérations de recherches ou de données provenant de

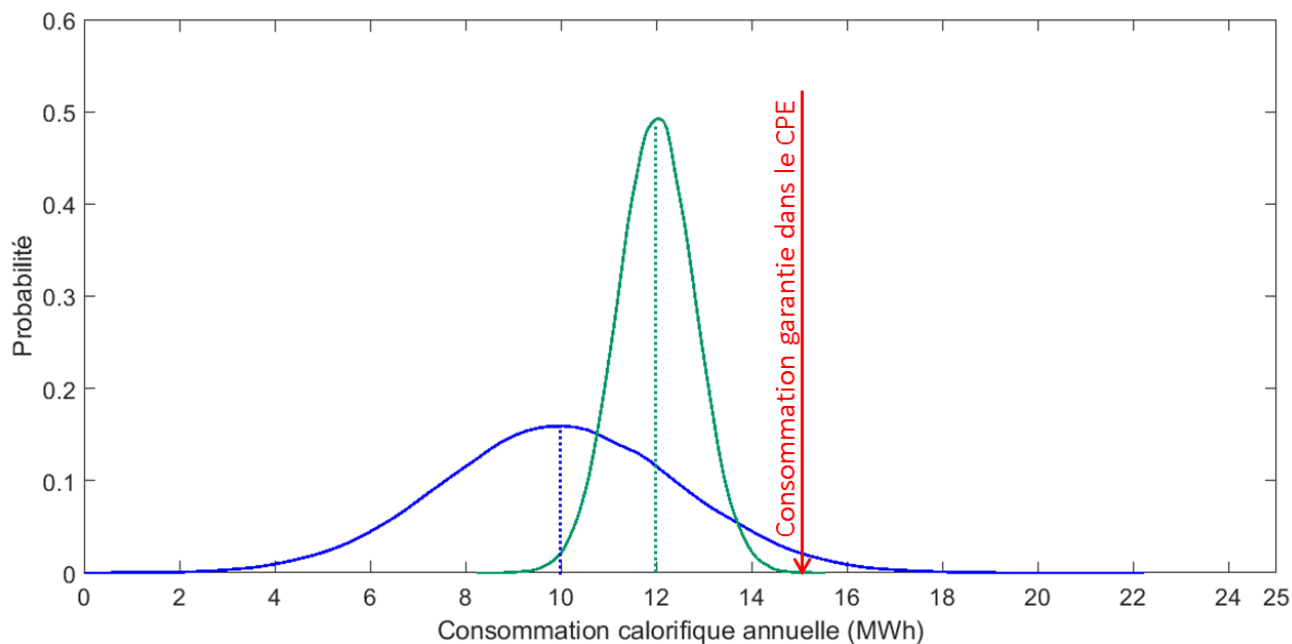


FIGURE 1.5 – Exemple de dispersion de deux distributions de performance énergétique garantie par un contrat de performance

bâtiments similaires. Les incertitudes associées aux variables dans ce modèle sont des incertitudes de non mesures et sont issues de mesures antérieures effectuées sur des bâtiments similaires. Ce modèle permet d’avoir une meilleure représentativité des variables d’entrée que dans le cas d’un modèle réglementaire et ceci sans avoir effectué de mesures.

Le modèle a posteriori quant à lui, permet de représenter le bâtiment dans les conditions réelles d’observation. Les données mesurées de chaque variable sont utilisées pour les études à mener ainsi que les scénarios constatés in situ. Il sera plus fin, plus précis et permettra d’obtenir des résultats factuels.

1.5 Objectifs de la thèse

Au regard des enjeux scientifiques présentés, on comprend que l’optimisation des plans de mesure et de vérification est l’un des principaux champs d’investigation, essentiel dans le processus de garantie de performance énergétique. Différents scénarios de mesures peuvent être envisagés et en fonction des résultats souhaités et du bâtiment étudié, un plan de mesure avec un niveau de complexité adapté sera mis en place.

Cette thèse s’inscrit dans le cadre du projet *ANR-13-VBDU-0006* Outils METHodologiques pour la GARantie de performance énergétique (OMEGA) qui a pour objectif de proposer des outils méthodologiques permettant d’appréhender, d’anticiper et de maîtriser

les risques liés à la mise en place d'un contrat de garantie de performance énergétique. Ces outils permettront de déterminer les variables qui ont une influence sur la consommation énergétique mais aussi de mieux maîtriser les incertitudes sur cette dernière en amont et en aval de l'étude. Ils permettraient ainsi aux entreprises qui souhaitent contractualiser la performance, d'adapter les contrats en fonction du risque visé et des exigences des clients.

Quant à la thèse, elle a pour objectif de proposer une méthodologie d'estimation des incertitudes des potentielles variables d'ajustement permettant d'optimiser la mise en place d'un plan de mesure en fonction du risque visé par un contrat de garantie de performance énergétique en phase d'exploitation du bâtiment. La variable d'intérêt considérée pour l'évaluation de la performance sera le besoin annuel de chauffage. ***Le besoin de chauffage annuel est l'apport d'énergie calorifique nécessaire au bâtiment sur une année complète d'étude pour assurer le confort thermique, de ses occupants. Elle permet en ce sens de rétablir l'équilibre thermique entre les apports internes et solaires et les déperditions causées par les infiltrations, la conduction à travers les parois et même la ventilation.*** Nous proposons ainsi une démarche opérationnelle et rapidement réalisable pour l'évaluation de la performance d'un bâtiment, en tenant compte de la mesure ou non des variables d'entrée. Ces variables d'entrée portent en priorité sur les variables dynamiques mais aussi sur les variables potentiellement ajustables dans un contexte de garantie.

Au chapitre 1 ont été exposés les enjeux environnementaux ainsi que les objectifs scientifiques et technologiques qui ont motivé ces travaux de recherche. Le mémoire est ensuite organisé comme suit : une méthodologie de quantification des incertitudes des variables d'entrée est proposée dans le chapitre 2. Les variables sélectionnées et étudiées sont des variables d'ajustement potentielles. La quantification des incertitudes, effectuée suivant différents cas de figure est basée sur les informations disponibles mais aussi sur le niveau de confiance escompté sur les résultats. Les incertitudes quantifiées serviront ensuite dans le chapitre 3, à hiérarchiser les variables et à construire un modèle physico-probabiliste pour la prise en compte de leur caractère aléatoire. Les méthodologies d'analyse de sensibilité et de réduction de modèle sont utilisées pour réduire le modèle physico-probabiliste et ainsi gagner en temps de calcul. Le modèle physico-probabiliste construit servira ensuite pour la propagation d'incertitudes dans le but d'optimiser les plans de mesure (Chapitre 4).

De la mesure aux données d'entrée des modèles

Introduction

Le chapitre précédent a permis de montrer que pour mettre en place un plan de MV adapté et efficace dans une démarche énergétique, il est important de connaître les objectifs de l'étude (en terme de performance et de précision) et surtout le cadre de la démarche qui est menée. L'utilisateur ne part pas d'une feuille blanche lors de la réalisation de son plan de MV, il peut s'appuyer sur les quatre options proposées par l'IPMVP ([IPMVP, 2010](#)). Cependant, il est important d'avoir une idée générale des mesures qui pourraient être mises en place ainsi que leur impact sur l'évaluation de la performance. Les grandeurs à mesurer varient d'une étude à une autre et dépendent de la typologie des bâtiments. Néanmoins, savoir en tenir compte ou définir les caractéristiques qui leur sont associées, telles que leurs distributions et leurs incertitudes, découle d'une démarche commune.

Il sera donc question dans ce chapitre, de définir selon la nature des mesures¹ qui ont été mises en place ou de l'estimation qui a été faite de chaque grandeur, les incertitudes qui y sont associées en amont de l'étude. La méthode d'estimation des incertitudes est exposée ainsi qu'une démarche simple et pragmatique pour la quantification des grandeurs considérées.

¹On distinguera en effet plus loin deux cas de figure : selon que l'on ait effectué des mesures in situ sur une variable, puis lorsque l'on se base sur des données issues de mesures antérieures sur des bâtiments similaires pour quantifier la variable.

2.1 Un besoin de fiabilité dans les mesures

Pouvoir quantifier une grandeur physique, c'est pouvoir lui attribuer une ou plusieurs valeurs qui contiennent en elles, toutes les informations pour la caractériser (moyenne, écart type, maximum, minimum, ...). Lorsque ces informations prennent l'aspect d'une valeur unique et déterministe, elles peuvent être interprétées comme étant certaines, c'est à dire connues avec exactitude et n'induisant aucun doute sur la valeur attribuée à la variable. Ce n'est malheureusement pas le cas en réalité : toute grandeur est entachée d'incertitudes.

2.1.1 Les différentes sources d'incertitudes

Pour assigner une valeur ou un ensemble de valeurs à la grandeur en question, on se base usuellement sur toutes les informations disponibles. Elles proviennent généralement de diverses sources et peuvent être de natures différentes. On qualifiera de données qualitatives toutes les informations provenant des connaissances ultérieures, de dires d'experts ou tout simplement d'un jugement personnel qui permettent de décrire la variable sans pouvoir la chiffrer. En effectuant des mesures sur cette grandeur, on obtient alors des informations quantitatives que l'on pourrait venir confirmer ou infirmer avec des données quantitatives ; ce qui permet de s'assurer de la précision des résultats.

Dans un cas comme dans l'autre, il est difficile d'affirmer que les connaissances dont on dispose sont certaines. Les données qualitatives peuvent altérer la prise de décision et par là induire un biais sur le résultat final. Dans le cas contraire où la grandeur a été mesurée, dans des conditions de mesures répétables (le même procédé, le même appareil, le même opérateur, le même lieu, ...), on observe une variabilité dans les résultats. Cette variabilité caractérise la dispersion des mesures et fait état de l'incertitude attribuée à la grandeur mesurée (sous la forme de distribution et d'écart type).

Il est donc essentiel, en assignant une valeur donnée à une grandeur, de pouvoir caractériser l'incertitude contenue dans les résultats obtenus à travers le procédé qui a permis de la caractériser. On entendra alors par incertitude tout au long de ce manuscrit, la dispersion attribuée à une grandeur, quelque soit le procédé utilisé. Elle peut prendre la forme d'un écart type, d'une plage de variation au sein de laquelle toutes les valeurs de la grandeur doivent se trouver. On pourra aussi associer à cette incertitude une probabilité associée, comme défini dans la section 2.1.6, pour que la valeur effective de la grandeur ait toutes les chances (généralement 70 à 99 %) de se trouver au sein de cette plage.

Il est fondamental de s'assurer que la quantification des incertitudes assignées aux grandeurs d'entrée soit menée avec précaution et que les résultats obtenus soient fiables car de ces incertitudes dépendront essentiellement la suite des analyses. Cette étape est souvent négligée dans le processus d'analyse d'incertitudes. La plupart des travaux ont

tendance à se concentrer sur la propagation d'incertitudes en elle-même et non sur la quantification minutieuse des incertitudes des grandeurs d'entrée. De récents travaux ont porté sur le sujet (Sun, 2014; Goffart, 2013; Spitz, 2012; De Wit and Augenbroe, 2002). Dans nos travaux, nous nous appliquerons dans ce sens à quantifier l'incertitude associée à la consommation calorifique annuelle à travers la propagation d'incertitude et l'accent sera mis sur la quantification des incertitudes des grandeurs qui influent sur cette consommation. C'est là l'un des points clés de nos travaux.

Les incertitudes peuvent aussi provenir des procédures de mesures et des appareils utilisés pour l'acquisition des données. Elles peuvent être estimées en tenant compte de l'influence des 5M (à savoir Méthode, Main d'oeuvre, Moyen, Matière et Milieu). Les instruments de mesure sont parfois la principale cause de biais dans les données. Le jugement de l'opérateur et les hypothèses faites peuvent avoir une grande influence sur les résultats de mesure et ainsi affecter la quantification d'incertitudes.

Aussi, en dehors des incertitudes inhérentes aux grandeurs d'entrées, d'autres incertitudes telles que celles propres au modèle de simulation peuvent aussi influencer les résultats. Les incertitudes issues de l'analyse ou de l'exploration des données en aval de l'étude sont difficiles à quantifier et proviennent dans la plupart des cas de jugements personnels ou de la maîtrise des outils par l'utilisateur (ce dernier doit avoir les compétences nécessaires pour utiliser les outils et de manière adéquate (De Wilde, 2014)).

Les incertitudes issues du modèle de simulation peuvent être déterminées en effectuant une comparaison des résultats de simulation et des données réelles, dans un processus de calibration. Le but d'un modèle de simulation est de représenter le bâtiment dans les conditions réelles auxquelles il est soumis et ceci le plus fidèlement possible. Une modélisation inappropriée ou l'utilisation de paramètres d'entrée inadéquats tendent à augmenter les incertitudes déjà présentes. Nous avons considérés dans notre étude que ces incertitudes étaient négligeables comparativement aux incertitudes des grandeurs d'entrée. **Une des particularité de cette thèse est l'accent qui a été mis sur la quantification des incertitudes des grandeurs d'entrée.**

2.1.2 Caractérisation d'une mesure : notions d'erreurs systématique et aléatoire, de justesse et de fidélité

Dans le but de caractériser une grandeur physique donnée et de pouvoir lui assigner une valeur, des observations sont effectuées. Ces observations sont faites par la mesure de la grandeur en question ou du phénomène qui l'engendre en le mesurant. Lorsqu'il s'agit d'une grandeur physique statique, en répétant la mesure un certain nombre de fois, on obtient des résultats différents d'une mesure à l'autre. Cette variabilité, due au changement des conditions de l'expérience, caractérise la dispersion de la grandeur.

En mesurant par exemple le débit de ventilation à une bouche d'un bâtiment, une erreur de lecture ou de positionnement de l'appareil de mesure peut faire différer les résultats d'une mesure à une autre. Ces incertitudes sont liées aux erreurs de type aléatoire et de type systématique. Le vocabulaire international de mesure (JCGM, 2008) propose comme définitions :

«L'erreur systématique est la composante de l'erreur de mesure qui dans les mesurages répétés, demeure constante ou varie de façon systématique.»

«L'erreur aléatoire est la composante de l'erreur de mesure qui varie de façon imprévisible.»

«L'erreur systématique est donc égale à la différence entre l'erreur de mesure et l'erreur aléatoire.»

L'erreur de type systématique caractérise la moyenne des écarts entre les valeurs que peuvent prendre la grandeur lors des mesures et sa valeur vraie. Elle est synonyme de biais et fait état de la justesse de la mesure. L'erreur de type aléatoire contient quant à elle la variabilité intrinsèque de la grandeur et permet de caractériser sa dispersion.

Pour visualiser ces concepts, nous reprenons l'exemple de la mesure du débit de ventilation. Sur la figure 2.1 sont représentés les résultats obtenus à la suite de mesures répétées. On constate sur la figure 2.1.a que les résultats de mesure sont très rapprochés et que l'écart entre les mesures est faible : les mesures sont justes² et fidèles³. Il y a donc une forte répétabilité dans les résultats. On associera alors à la mesure du débit de ventilation une erreur aléatoire très faible voire nulle. À l'opposé, on observe sur la figure 2.1.d une dispersion élevée, l'histogramme est étalé, les résultats de mesure ne sont donc pas justes et sont peu fidèles. L'erreur associée à ce résultat de mesure sera constituée d'une composante aléatoire et d'une composante systématique.

À la base, en effectuant des mesures sur une grandeur, on améliore la connaissance a priori sur cette dernière. On s'attend à obtenir des résultats de mesures fidèles et justes des erreurs aléatoires et systématiques très faibles voir inexistantes. Comme indiqué dans le chapitre précédent, l'indicateur de performance choisi étant le besoin annuel de chauffage, l'erreur aléatoire associée à certaines variables aura tendance à se compenser sur une année complète. **Seule l'erreur systématique sera représentative de la grandeur et aura un effet sur la variable d'intérêt en sortie du modèle.**

²La justesse de mesure est l'écart entre la moyenne d'un nombre infini de valeurs mesurées et une valeur de référence (JCGM, 2008).

³La fidélité de mesure caractérise l'écart entre les valeurs mesurées obtenues par des mesurages répétés, du même objet dans des conditions spécifiées (JCGM, 2008).

2.1.3 Méthodes d'estimation d'incertitude : La méthode GUM

Pour assigner une incertitude à une grandeur d'entrée, plusieurs méthodes usuelles sont employées. Le plus souvent, l'incertitude associée aux données mesurées provient non seulement du procédé métrologique mais aussi de la variabilité contenue dans les données. La qualité des résultats est généralement affectée par les procédés utilisés, les capteurs et instruments métrologiques mais aussi par les hypothèses faites durant l'analyse ou le dépouillement des résultats.

Plusieurs méthodes sont utilisées pour estimer les incertitudes de mesures. Parmi elles, la méthode GUM proposée par le Guide to the expression of Uncertainty in Measurement (GUM) (GUM, 1995) reste la plus courante. Elle propose d'évaluer l'incertitude suivant deux cas de figures.

2.1.3.1 Évaluation type A de l'incertitude

On considère la variable X . Un nombre n de mesures a été effectué pour quantifier X dans des conditions expérimentales similaires donnant des résultats de mesures $\{x_1, x_2, \dots, x_n\}$.

Les métrologues proposent, lorsque l'on fait face à plusieurs résultats de mesures, de caractériser leur dispersion par une approche statistique.

En estimant la moyenne \bar{x} qui est le meilleur estimateur :

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad (2.1)$$

L'écart type σ qui caractérise la dispersion des mesures effectuées :

$$\sigma = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.2)$$

Et la dispersion sur la moyenne :

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad (2.3)$$

La moyenne \bar{x} fait alors office de valeur assignée à la grandeur et la dispersion sur la moyenne $\sigma_{\bar{x}}$, d'incertitude associée aux mesures effectuées.

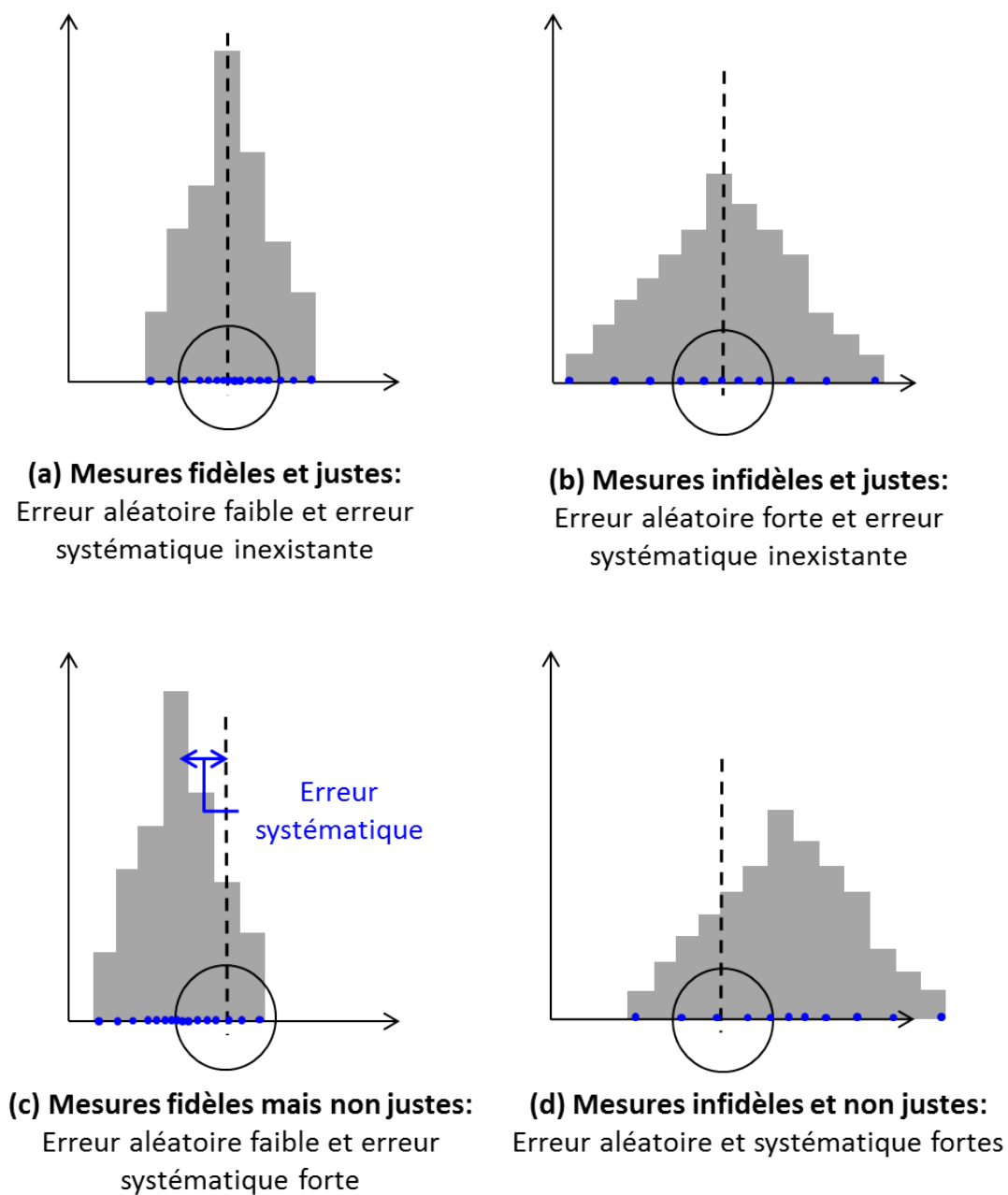


FIGURE 2.1 – Différentes configurations possibles des erreurs aléatoires et systématiques. La ligne en pointillée représente la valeur vraie de la grandeur physique et l'histogramme est construit sur la base des mesures effectuées.

2.1.3.2 Évaluation type B de l'incertitude

Cette méthode est appropriée lorsque la grandeur étudiée n'a pas été quantifiée à partir d'observations répétées. Elle est basée sur les connaissances dont dispose l'opérateur sur le procédé utilisé lors de l'expérimentation, l'étalonnage des appareils de mesure, leur résolution⁴, ..., bref toute information qui pourrait aider à assigner une incertitude à la grandeur mesurée. Elle est donc une combinaison de toutes ces sources d'informations et est déterminée par :

$$u_b = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2} \quad (2.4)$$

Où u_i correspond à une composante de l'incertitude de non mesure.

u_b est en quelque sorte une incertitude composée.

Il est aussi, tout à fait possible, de combiner les deux composantes (incertitudes de type A et B), lorsque l'on dispose d'informations dans les deux cas de figures. On est alors face à une incertitude de type composée telle que spécifiée dans l'équation 2.5.

$$u = \sqrt{u_a^2 + u_b^2} \quad (2.5)$$

Pour conclure, la méthodologie d'estimation des incertitudes de mesure proposée par le GUM est relativement simple et facile à implémenter. Pour estimer l'incertitude associée à une grandeur, il convient tout d'abord de définir les causes d'incertitudes. Si la grandeur a été quantifiée par un mesurage, il convient de déterminer l'incertitude de type A en appliquant les méthodes statistiques présentées précédemment sur les séries de mesure effectuées ; puis lorsque l'on dispose d'informations autres que des mesures répétées, informations sur le procédé de mesure par exemple (résolution, étalonnage de l'appareil de mesure), on évalue l'incertitude de type B. Ensuite, dans le cas où l'on peut évaluer l'incertitude de type A et de type B, on détermine l'incertitude composée.

2.1.4 Statistiques et échantillonnage

Dans le but de quantifier au mieux une grandeur considérée, au lieu de lui assigner une estimation ponctuelle, on lui attribue un intervalle dans lequel il a toutes les chances de

⁴La résolution d'un instrument de mesure est la plus petite variation de la grandeur mesurée qui produit une variation perceptible de l'indication correspondante (JCGM, 2008).

se trouver avec une probabilité contrôlée.

Soit X la grandeur à quantifier et p_X la population de tous les x mesurables. On souhaite caractériser p_X à partir de la mesure d'un échantillon e_X ($e_X \in p_X$) à un instant t donné.

On cherche donc à déterminer un intervalle de confiance de la population p_X pour que la probabilité que toutes les valeurs prises au sein de cette population ne soient pas dans cet intervalle soit égale à p . On prend donc un risque donné p pour que les valeurs n'appartiennent pas à cet intervalle.

La moyenne

On considère une variable aléatoire X tel que, à un instant t donné :

$$X \hookrightarrow N(\mu, \sigma) \quad (2.6)$$

La population p_X est constituée de x_1, x_2, \dots, x_n , n échantillons indépendants et de même loi que X . La moyenne \bar{x} et la variance empirique s , sont données par les équations 2.1 et 2.2.

Lorsque la taille de l'échantillon est inférieure à 30, la variable aléatoire $\frac{\bar{x}-\mu}{s/\sqrt{n-1}}$ suit une loi de Student à $(n-1)$ degrés de libertés. Grâce à la théorie de l'inférence statistique, la valeur moyenne μ , sur la population p_X , peut être approchée par la moyenne de l'échantillon \bar{x} et son écart type estimé s par :

$$\bar{x} - t_{p/2} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{p/2} \frac{s}{\sqrt{n}} \quad (2.7)$$

Où $t_{p/2}$ est lu dans la distribution de Student avec un risque associé de $p/2$ et $(n-1)$ degrés de libertés connus.

L'équation 2.7 donne l'intervalle de confiance de μ . Elle fournit une marge d'erreur sur l'estimation de la moyenne de la population.

La variance

On sait que :

$$\frac{(n-1)s^2}{\sigma^2} \hookrightarrow \chi^2(n-1) \quad (2.8)$$

En encadrant la loi de Khi-deux avec le niveau de risque associé :

$$P\left(\chi_{1-p/2}^2 < \chi^2 < \chi_{p/2}^2\right) = 1 - p \quad (2.9)$$

On obtient :

$$P\left(\chi_{1-p/2}^2 < \frac{(n-1)s^2}{\sigma^2} < \chi_{p/2}^2\right) = 1 - p$$

$$\frac{(n-1)s^2}{\chi_{p/2}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{1-p/2}^2} \quad (2.10)$$

Où $\chi_{p/2}^2$ et $\chi_{1-p/2}^2$ sont donnés par les tables de Khi-deux à $(n-1)$ degrés de liberté.

On obtient ainsi un intervalle de confiance de la variance de la population p_X avec un niveau de risque associé p .

2.1.5 Différence entre erreur et incertitude

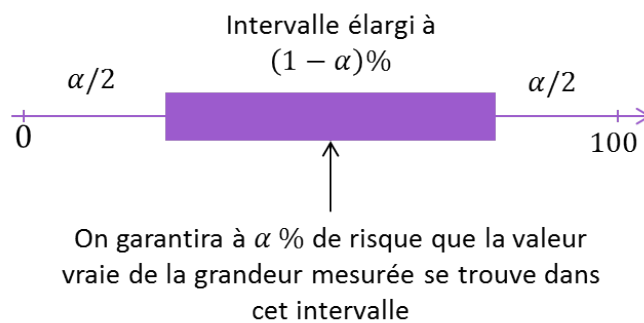
Les notions d'erreur et d'incertitude sont deux termes souvent confondus à tort. Comme nous l'avons défini précédemment, l'incertitude permet de caractériser la dispersion attribuée à la mesure d'une grandeur. À ce titre, elle permet de quantifier l'assurance qu'on a sur des résultats de mesure. Elle est donc imputable à un résultat de mesure en particulier et sera différente lorsque le résultat de mesure est modifié. En d'autres termes, c'est l'opinion qu'on se fait sur l'erreur possible du résultat de mesure.

Dans le cadre de la quantification des incertitudes associées aux grandeurs d'entrée, la notion d'incertitude désignera la représentation de l'erreur que l'opérateur peut se faire sur la mesure de la valeur vraie. Elle sera alors constitué des erreurs à caractères systématiques qui dans le cadre de la caractérisation de la consommation calorifique annuelle variera de manière constante suivant une année d'étude.

2.1.6 Intervalle élargi et probabilité associée

En métrologie, il est associé au résultat de mesure, un intervalle élargi contenant l'ensemble des valeurs vraies des résultats de mesure avec une probabilité donnée (JCGM, 2008). À un risque donné α , cette probabilité, appelée probabilité de couverture $(1-\alpha)$, constitue le degré de confiance associé à cet intervalle. Toutes les valeurs vraies de la grandeur auront ainsi une forte probabilité de se trouver au sein de l'intervalle élargi. Il est représentatif des résultats de mesure et donne un encadrement de la valeur réelle de la grandeur physique.

Ainsi, en augmentant le facteur d'élargissement, l'intervalle élargi augmentera aussi. On assiste à une diminution de la précision et la probabilité d'obtenir des valeurs aberrantes sera plus élevée. Choisir un facteur d'élargissement adapté à un résultat de mesure est

FIGURE 2.2 – Illustration de l'intervalle élargi avec un niveau de risque associé α

donc essentiel car il permet d'affirmer que la valeur vraie de la variable mesurée se trouve à $(1 - \alpha)\%$ au sein de l'intervalle élargi (figure 2.2). Quelques facteurs d'élargissement et leurs probabilités équivalentes sont représentés dans la table 2.1 dans le cas de la mesure d'une grandeur supposée gaussienne. Pour un facteur d'élargissement donné $k = 1.96$ par exemple, la probabilité que les valeurs de la grandeur soient dans l'intervalle élargi sera de 95% avec un risque de 5%.

TABLE 2.1 – Exemple de facteur d'élargissement et probabilités associées pour une variable gaussienne

Facteur d'élargissement k	1	1.64	1.96	2	2.33	3.29
Probabilité de couverture $(1-\alpha)\%$	68.27	90.00	95.00	95.45	98.00	99.9
Niveau de risque $\alpha\%$	31.73	10	5	4.55	2	0.1

2.1.7 Variables statiques et variables dynamiques

Parmi les variables qui influent sur le besoin annuel de chauffage, il y a celles qui sont inhérentes à l'architecture du bâtiment telles que celles qui caractérisent sa forme, la constitution de son enveloppe, les ponts thermiques, le facteur solaire des fenêtres puis celles qui caractérisent les systèmes qui sont installés dans le bâtiment telles que le coefficient d'induction, le rendement d'un générateur ou sa puissance. Toutes ces grandeurs sont statiques ou ont une évolution très faible pendant toute la durée de vie du bâti (sauf en cas de mise en place d'APE ou de réhabilitation) et sont fixées à l'étape de conception (sauf changement éventuel pendant la construction).

D'un autre côté, le bâtiment fait partie d'un environnement dynamique et en permanente évolution. Il est notamment sous l'influence des conditions météorologiques constamment variables et s'adapte également à l'activité des occupants. Les grandeurs physiques qui caractérisent cette évolution temporelle telles que la température extérieure, la vitesse du vent, l'humidité, le rayonnement solaire, le taux d'occupation, les apports internes liés à l'activité des occupants sont considérées comme dynamiques.

La démarche adoptée dans ces travaux de recherche étant orientée vers la mise en place des plans de mesure, les grandeurs d'intérêt seront pour la plupart des grandeurs à caractère dynamique. Les grandeurs statiques ne nécessitant pas de mesures en continu, elles sont rarement considérées comme variable d'ajustement dans une démarche de garantie de performance.

2.1.8 Enregistrement, stockage et prétraitement des données

Lorsque des mesures sont effectuées de manière continue, elles sont d'abord enregistrées et stockées avant leur traitement et leur utilisation. Les capteurs de mesures sont souvent raccordés à des systèmes de gestion et de stockage centralisés. Dans notre cas, l'acquisition et le stockage a été spécifique à chaque mesure effectuée. Pour la mesure des températures intérieures des locaux par exemple, les capteurs étaient équipés de boîtiers enregistreurs capables de stocker plus de 10 000 points de mesure. Sur un pas de temps de 5 minutes, l'on pouvait donc acquérir des données en continu pendant un mois sans dépasser la capacité de stockage des enregistreurs. Les données étaient récupérées sur des intervalles mensuels. D'un autre coté pour la mesure des consommations électriques et de chauffage, des compteurs calorifiques étaient installés sur le site. Les données étaient enregistrées à distance et récupérées sur un poste centralisé, ce qui évitait des déplacements et présentait l'avantage de pouvoir vérifier le bon déroulement des mesures de manière régulière. De ce fait, les simulations trnsys étant effectuées sur une base horaire, il n'était pas nécessaire d'avoir des points de mesure à des pas de temps très faibles. Celles qui pouvaient être effectuées à des pas de temps de 5 voire 15 minutes lorsque le stockage le permettait étaient réalisées. En revanche, les autres équipements qui ne le permettaient pas, recevaient des points sur une base horaire.

En ce qui concerne la synchronisation de toutes les grandeurs mesurées au même pas de temps, elle est effectuée par l'opérateur. Nous avons implémenté un code de calcul qui permettait de gérer les données mesurées et de les mettre au même pas de temps. En effet, les données qui n'étaient pas incrémentées à chaque impulsion et qui étaient à des pas de temps de 5 ou de 15 minutes étaient moyennées pour obtenir des données horaires (exemple de mesure de température intérieure, des données météorologiques). Par contre, celles qui sont incrémentées (mesure de la consommation calorifique et électrique) ont été dans un premier temps ajustées au pas de temps horaire puis dé-cumulées avant leur exploitation.

Aussi, une des difficultés majeures à laquelle nous avons été confronté durant le pré-traitement des données a été la présence de données manquantes. En effet, des défaillances de connexion pouvaient intervenir et les boîtiers d'enregistrement ne plus répondre. Le bâtiment étant aussi en exploitation continue, les équipements pouvaient être manipulés

par les occupants et fournir des points de données aberrants. Une méthodologie basée sur la corrélation des mesures a été développée pour faire face à ce problème (Titikpina et al., 2015).

En résumé, en dehors des protocoles de mesures qui sont mis en place, l'enregistrement, le stockage et le prétraitement des données avant leur exploitation peuvent être des sources d'erreurs supplémentaires qui sont surtout liés à des phénomènes aléatoires que l'on ne maîtrise pas.

2.1.9 Présentation du cas d'étude

Le bâtiment test présenté ci-après est le même pour toutes les études effectuées dans le cadre de cette thèse. Il a été modélisé et instrumenté dans le cadre du projet ANR OMEGA et a servi de support pour cette thèse. Le bâtiment test appelé bâtiment CECP (Centre d'Études et de Conception des Prototypes) est situé aux Ponts de Cé, en France et a été livré en 2012. C'est un bâtiment tertiaire dont la majeure partie est consacrée à un usage de bureaux (2/3 de la surface totale), le reste servant d'atelier ou de support pour la réalisation des travaux du Cerema. Il est essentiellement constitué de bois avec un mur en ossature bois et de la laine de bois pour l'isolation (figure 2.3).



FIGURE 2.3 – Bâtiment CECP

2.2 Quantification des incertitudes des grandeurs d'entrée

Dans le but de pouvoir caractériser avec précision les incertitudes en sortie du modèle sur la grandeur physique considérée, il est indispensable de sélectionner dans un premier temps, les variables d'entrée et les sources d'incertitudes inhérentes à chacune d'entre elles. L'identification de ces sources d'incertitudes va permettre de pouvoir quantifier de manière

fiable les incertitudes qui leur sont associées d'autant plus que les sources d'incertitudes sont de natures diverses et dépendent fortement des grandeurs en jeu.

Pour ce faire, dans la suite du chapitre, les grandeurs physiques ont été scindées en trois catégories :

1. Celles dues aux sollicitations extérieures à savoir la température extérieure, les rayonnements solaires, la vitesse du vent ;
2. Celles dues au renouvellement d'air telles que les débits de ventilation, l'efficacité des centrales de traitement d'air ainsi que la perméabilité à l'air du bâtiment ;
3. Celles dues aux sollicitations internes notamment à l'activité de ses usagers, telles que la température intérieure, le taux d'occupation et les apports internes.

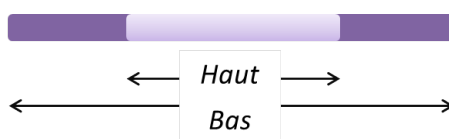


FIGURE 2.4 – Illustration des deux niveaux d'étude considérés pour une grandeur d'entrée donnée X_i

Pour chaque grandeur, il a été défini deux niveaux d'étude : un niveau haut qui correspond à l'option de mesure la plus coûteuse (Option C de l'IPMVP). Dans ce niveau, on considère que toutes les grandeurs sont mesurées et l'incertitude de mesure associée à chaque grandeur d'entrée est estimée.

À l'opposé, on considère un niveau bas où aucune campagne de mesure n'a été menée. Lorsque des mesures ne sont pas effectuées, des valeurs théoriques sont assignées aux grandeurs d'entrée avec des incertitudes relatives fournies arbitrairement (de 10, 20 voire 30%), ce qui peut accroître l'incertitude en sortie du modèle, selon la sensibilité de ce dernier à la variable d'intérêt. Les incertitudes sont estimées à partir de mesures antérieures effectuées sur des bâtiments similaires et de dires d'experts. Elles seront dans ce niveau probablement plus élevées que dans le niveau haut mais la mise en place du plan de mesure sera moins coûteuse (table 2.2 et figure 2.4).

Les incertitudes liées aux grandeurs non mesurées sont estimées à partir de mesures antérieures effectuées sur des bâtiments similaires ou dans le cadre de projets similaires. Seront considérés comme données antérieures, les résultats obtenus lors des recherches menées par le Cerema sur les bâtiments démonstrateurs à basse consommation d'énergie (PREBAT, 2015), et ceux obtenus dans le cadre même du projet ANR OMEGA.

Dans la suite, une méthodologie détaillée est proposée pour l'évaluation des incertitudes associées aux grandeurs d'entrée présélectionnées et suivant les deux cas de figure définis

dans la table 2.2 et à la figure 2.4 puisque de cette quantification dépendent les résultats d'analyse.

TABLE 2.2 – Quantification des incertitudes des grandeurs d'entrée suivant les différents niveaux d'étude

Niveau	Plan de mesure	Quantification des incertitudes
Haut	Mesures complètes	Incertitudes estimées à partir de mesures in situ
Bas	Aucune mesure	Incertitudes estimées à partir de mesures antérieures sur des bâtiments similaires

2.2.1 Incertitudes liées aux sollicitations extérieures

Comme définies plus haut, sont considérées comme sollicitations externes tous les phénomènes physiques extérieurs au bâtiment et qui ont une influence sur son comportement. Le premier, et le plus connu de ces phénomènes est la variabilité météorologique qui, d'une année à l'autre, au gré des saisons et au fil du temps, impose au bâtiment et à ses équipements, le niveau de performance requis pour assurer le confort de ses occupants.

Il est possible de déterminer les incertitudes des variables météorologiques de différentes manières et le choix de la méthode dépend essentiellement de l'objectif de l'étude, qui dans notre cas est de quantifier l'incertitude associée à la consommation calorifique annuelle. Nous nous intéressons donc qu'à l'incertitude annuelle des données météorologiques. Comme l'a souligné Macdonald ([Macdonald and Strachan, 2001](#)), si des mesures ont été effectuées lors de l'année étudiée, alors seules les incertitudes de mesures doivent être prises en compte lors de l'analyse d'incertitudes. Si le but de l'étude est d'estimer une réponse typique du modèle sous certaines conditions, des fichiers météorologiques typiques sont alors générés. Étant basée sur des moyennes annuelles, l'incertitude sur ces moyennes devra alors être estimée. Par contre, si les prédictions sont réalisées pour une année précise, alors l'incertitude associée aux variables météorologiques devra prendre en compte les variations climatiques d'une année à une autre.

Généralement, les fichiers de données météorologiques au pas de temps horaire sont utilisés pour estimer la consommation énergétique du bâtiment lors de simulations. Dans le cadre d'une analyse de sensibilité et de propagation d'incertitudes, la quantification des sources d'incertitudes est une étape clé. Elle permet de définir les caractéristiques propres à chaque variable météorologique suivant la provenance des données utilisées et de leur variabilité d'une année à une autre. Il est néanmoins important de rappeler que seuls les aléas induisant des erreurs systématiques sont considérés car l'erreur aléatoire se compense

sur l'année lorsqu'on considère la consommation calorifique annuelle comme variable d'intérêt. Les variables météorologiques que nous avons choisies de retenir dans ces travaux sont : la température extérieure, la vitesse du vent et les radiations solaires. Les autres grandeurs météorologiques (direction du vent, humidité relative, pression, ...) n'ont pas été considérées. Ceci étant, la démarche adoptée reste la même si l'on souhaite quantifier les incertitudes associées à d'autres variables météorologiques telles que l'humidité relative ou la pression.

Lors d'une simulation thermique dynamique, une valeur unique doit être rentrée au modèle à chaque pas de temps, pour toutes les grandeurs dynamiques notamment les variables météorologiques. Cependant, le modèle d'échange thermique entre une paroi quelconque du bâti et l'extérieur fait état d'une couche limite. Les valeurs des paramètres météorologiques doivent être prises à l'infini c'est à dire à une distance supérieure à l'épaisseur de cette couche limite. Il va pour ainsi dire qu'aucune perturbation autre que celles dues aux échanges avec la paroi considérée ne doit intervenir. La valeur prise par chaque grandeur varie d'une paroi à une autre et doit pouvoir être généralisée dans le cadre de mesures multiples en entrée du modèle STD.

2.2.1.1 Décomposition de l'incertitude

Les différentes sources d'incertitudes présentes dans les variables météorologiques peuvent être scindées en deux. La première concerne la variabilité de chaque grandeur d'une année à l'autre et la seconde est inhérente aux mesures effectuées pour obtenir les données. Pour effectuer une simulation thermique dynamique avec des données mesurées, on utilise les relevés de données météorologiques mais dans le cas où des mesures n'ont pas été effectuées, on utilise des fichiers standards généralement disponibles dans les bases de données météorologiques (Météonorm, Infoclimat). Dans les deux cas de figure, les incertitudes ne sont pas quantifiées de la même manière. Elles sont donc déclinées pour chaque niveau dans les lignes qui suivent.

2.2.1.2 Estimation des incertitudes liées aux sollicitations extérieures dans le cas d'un niveau haut (mesures complètes)

Méthodologie

Pour estimer l'incertitude associée à une mesurande en métrologie, on effectue généralement dans des conditions similaires un nombre répété de mesures (GUM, 1995). Cependant, face à des grandeurs dynamiques comme les variables météorologiques, il devient difficile de répéter les mesures dans des conditions rigoureusement identiques, du fait de la dynamique temporelle. Pour pouvoir estimer l'incertitude liée à la mesure des variables météorologiques, plusieurs stations peuvent être installées sur le site d'expérimentation, pendant une durée suffisamment importante pour recueillir les données nécessaires à l'estimation de l'incertitude.

On considère qu'à chaque emplacement d'une station météo, correspond un échantillon de la grandeur considéré e_X et qu'à elles toutes (toutes les périodes de mesures et tous les emplacements considérés), elles sont représentatives de la population de la grandeur p_X (Section 2.1.4). Ainsi, pour un échantillon donné X_i de la grandeur X mesurée pendant une période Δt_i , sa généralisation sur toute la période de mesure Δt_i est donnée par :

$$X_i = \frac{\int X_i(t) dt}{\Delta t_i} \quad (2.11)$$

On estime ensuite δ_i la différence suivante :

$$\delta_i = X_i - X_* \quad (2.12)$$

Où X_* est la valeur obtenue par la station météo de référence durant la même période de mesure Δt_i .

$$\delta_i = \frac{\int (X_i(t) - X_*(t)) dt}{\Delta t_i} \quad (2.13)$$

La variance de l'échantillon s^2 est ensuite donnée par la variance de tous les δ_i par l'équation ci-dessous :

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (\delta_i - \bar{\delta})^2 \quad (2.14)$$

En encadrant s^2 par la loi de Khi-deux (équation 2.10), on obtient un intervalle de confiance associé à l'estimation de σ^2 , la variance de la population entière. Pour toutes les variables météorologiques, la méthodologie d'estimation d'un intervalle de confiance présentée en section 2.1.4 est appliquée. Les équations 2.7 et 2.10 fournissent un encadrement respectivement de l'estimation de la moyenne et de l'écart type de la population.

Cas d'étude

Les données météo proviennent d'une station implantée sur le site de recherche du Cerema. Ces données proviennent d'observations réelles de l'année d'étude (2015-2016). Deux stations de mesure ont ainsi été installées, l'une fixe et l'autre mobile (table 2.3), à proximité du bâtiment étudié. La station mobile a été déplacée sur sept emplacements différents (figure 2.5). Les données ont été mesurées au pas de temps de 5 minutes et pour une durée de deux semaines à chaque emplacement.

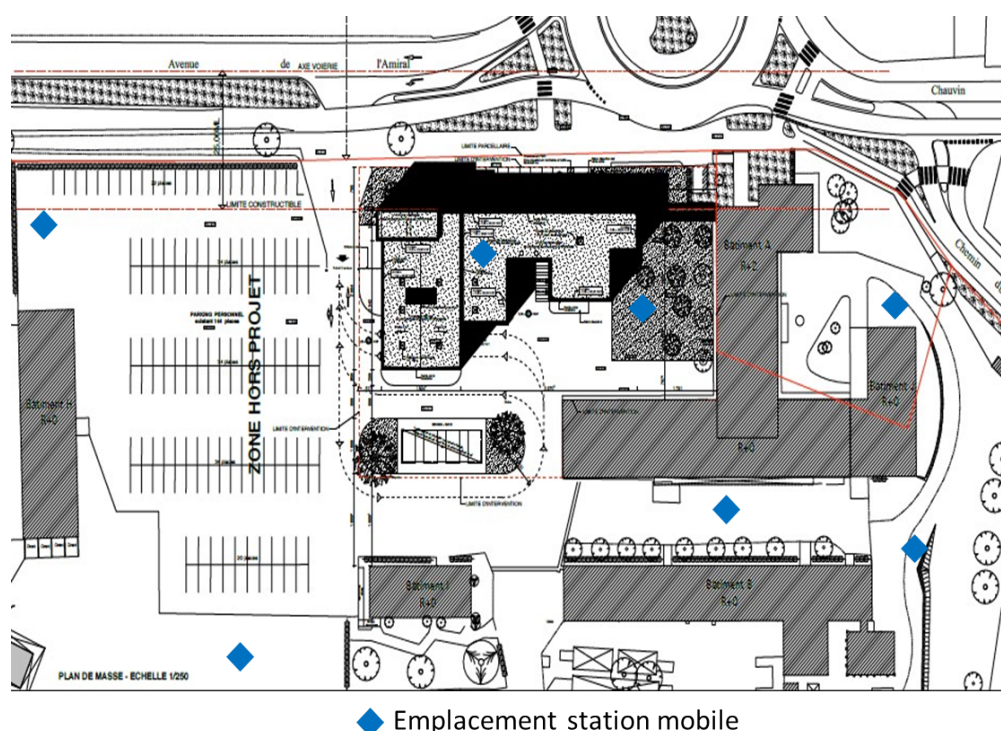


FIGURE 2.5 – Différents points de déplacement des stations météorologiques installées sur site. La station mobile a été déplacée aux différents emplacements notifiés en bleu.

TABLE 2.3 – Illustration des périodes de mesures effectuées aux différents emplacements des stations météo

Période	Station mobile	Station fixe
Δt_1	Station 1	Inchangée
Δt_2	Station 2	Inchangée
\vdots	\vdots	\vdots
Δt_7	Station 7	Inchangée

La température extérieure

La température extérieure en entrée d'un modèle de Simulation Thermique Dynamique (STD) peut être définie comme la moyenne des températures d'air au droit de chaque élément de façade et à une distance suffisamment éloignée pour ne pas être dans la couche limite et à une distance suffisamment proche pour ne pas être sous l'influence d'autres phénomènes thermiques comme le vent, l'effet de bâtiments voisins, ...

On évalue ainsi pour chaque emplacement de la station mobile, la différence avec la station de référence et on encadre ensuite la variance obtenue par la loi de Khi deux à 95% de confiance, en se basant sur les équations 2.10 et 2.14. Les résultats obtenus sont présentés dans les tables 2.4 et 2.5. On associera ainsi dans le cas de mesures, une incertitude à la température extérieure de $0.625\text{ }^{\circ}\text{C}$.

Les rayonnements solaires

Toutes les stations ont été placées avec soin dans des champs dégagés pour ne pas être sous l'effet des masques proches (bâtiments voisins, végétation, . . .), mais tout en tenant compte des masques lointains (montagnes, . . .). L'étude ne portera, pour l'irradiation solaire, que sur les pas de temps où le résultat de mesure est strictement positif c'est à dire en plein jour. Un seuil de $5W/m^2$ a donc été fixé. En deçà des $5W/m^2$, l'irradiation a été considérée comme nulle. Les incertitudes associées aux variables représentant les sollicitations extérieures dans le cas où des mesures ont été effectuées, sont présentées à la table 2.5.

La vitesse du vent

Les mesures effectuées ont été décomposées en deux parties : les vents faibles et les vents forts. Nous ne retiendrons que les vitesses de vent supérieures à 1 m/s, l'anémomètre ayant une résolution de 1m/s, et, en deçà de cela, il n'est pas capable de détecter la présence de vent faible.

TABLE 2.4 – Différence, pour chaque grandeur (température extérieure, rayonnements solaires et vitesse du vent) entre l'emplacement de la station mobile et la station de référence

Emplacement station mobile	$\delta_i(^{\circ}C)$	$\delta_i(W/m^2)$	$\delta_i(m/s)$
1	0.2268	22.64	0.51
2	0.8375	38.43	0.23
3	0.0218	21.84	0.16
4	0.1023	66.64	0.26
5	0.2753	45.77	0.31
6	0.2568	35.17	0.44
7	0.7892	33.11	0.32

TABLE 2.5 – Évaluation de l'incertitude liée à la mesure de la température extérieure, des rayonnements solaires et de la vitesse du vent à 95% de confiance

Variable	Écart type s	Encadrement par la loi de Khi deux de l'écart type σ
Température extérieure ($^{\circ}C$)	0.326	0.625
Rayonnements solaires (W/m^2)	15.32	29.34
Vitesse du vent (m/s)	0.12	0.22

2.2.1.3 Estimation des incertitudes liées aux sollicitations extérieures dans le cas d'un niveau bas (aucune mesure)

Méthodologie

Les fichiers météorologiques couramment utilisés lorsque l'on ne dispose pas de mesures sont issus de la base de données Meteonorm. Meteonorm est une base qui permet de générer des fichiers annuels de météo typiques d'une région particulière, en se basant sur un minimum d'une dizaine d'années d'observation. Ces fichiers météo sont les plus utilisés en simulation. Pour estimer l'incertitude associée à chaque variable météo lors de l'utilisation de fichiers Meteonorm, l'analyse s'est basée sur les données recueillies par Météo France entre 2005 et 2015 à la station météo située à Nantes, ne disposant pas de données suffisantes sur Angers.

Cas d'étude

Pour estimer la variabilité de la température extérieure sur les dix dernières années (2005 à 2015), une base de données constituée des températures minimales, moyennes et maximales de janvier 2005 à Décembre 2015 a été considérée. De ces données ont été estimées la variabilité moyenne de la température extérieure de chaque mois pris individuellement (Annexe 6.1, tables 6.1, 6.2 et 6.3) ainsi, qu'une moyenne de cette variabilité sur une année entière (table 2.6). En procédant de manière similaire à la température extérieure, les caractéristiques des rayonnements solaires et de la vitesse du vent ont été estimées à partir de la base. Elles sont représentées pour chaque variable à la figure 2.6.

Rappelons que la variable d'intérêt considérée étant la consommation calorifique annuelle, l'incertitude associée aux variables météorologiques a été estimée à partir de leur variabilité inter annuelle et non à chaque pas de temps. Par ailleurs, ces données n'ayant pas été recueillies sur le site sur lequel se trouve le bâtiment test (25 km de distance entre les sites), l'incertitude estimée précédemment (table 2.6) doit être corrigée pour pouvoir être utilisée. Cette correction intègre en elle l'effet du micro-climat. La correction s'effectue en intégrant les incertitudes des mesures réalisées sur le site d'expérimentation (évaluation de l'incertitude composée par l'équation 2.5). On obtient ainsi les incertitudes présentées dans la table 2.6. L'écart type expérimental s , celui encadré ensuite par la loi de Khi deux σ et l'incertitude finale sont obtenus respectivement par les équations 2.2, 2.10 et 2.5.

2.2.2 Incertitudes liées au renouvellement d'air

Les variables représentatives du renouvellement d'air qui ont été sélectionnées sont les débits d'insufflation et de reprise des centrales de traitement d'air, dans chaque zone du bâtiment, l'efficacité de ces centrales pour le système aéraulique ainsi que la perméabilité à l'air du bâtiment. D'autres variables (rendement des ventilateurs, puissance des batteries, ...) peuvent être prises en compte dans le cas où d'autres systèmes aérauliques sont

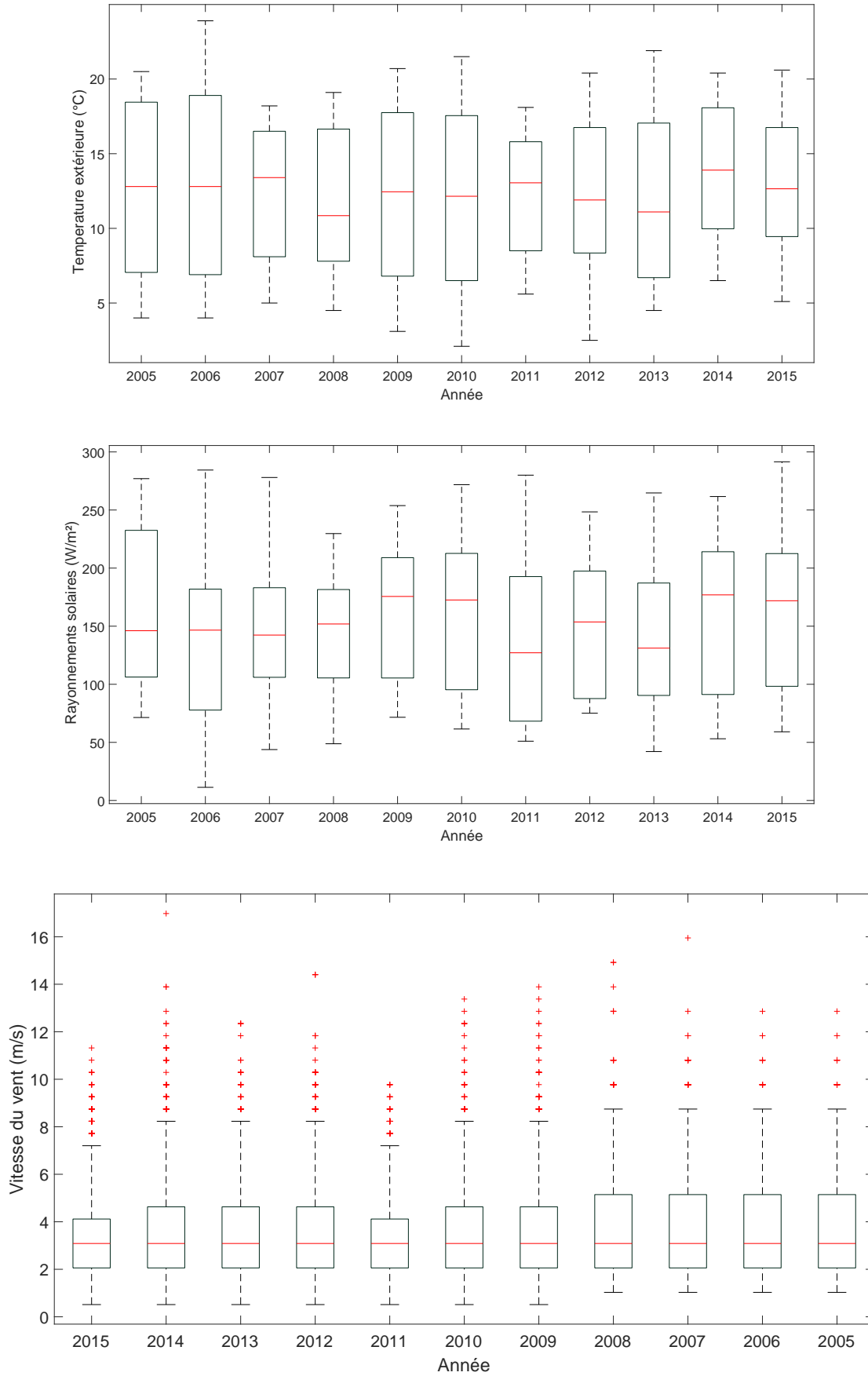


FIGURE 2.6 – Variation de la température extérieure, des rayonnements solaires et de la vitesse du vent de 2005 à 2015. La ligne en rouge représente la médiane des données. Sont représentées de part et d'autre de cette ligne et dans le rectangle, 25 % des données et les deux autres quartiles sont représentés par des pointillés.

TABLE 2.6 – Moyenne de chaque variable météorologique (température extérieure, rayonnements solaires et vitesse du vent) pour chaque année comprise entre les 2005 à 2015 et évaluation de l'incertitude liée aux données de température extérieure de 2005 à 2015 à 95% de confiance et pour 10 degrés de libertés

Année	$\mu_i(^{\circ}C)$	$\mu_i(W/m^2)$	$\mu_i(m/s)$
2005	13.03	163.32	3.71
2006	13.74	161.63	3.81
2007	12.06	143.00	3.61
2008	12.29	155.14	3.39
2009	12.28	140.72	3.45
2010	11.63	164.27	3.21
2011	12.37	164.53	3.36
2012	11.99	147.06	3.50
2013	12.40	148.62	3.46
2014	12.84	143.14	3.16
2015	12.46	167.95	3.46
Écart type expérimental s	0.57	10.23	0.19
Encadrement par la loi de Khi deux σ	0.91	16.3	0.3
Incertitude finale	1.1	33.57	0.37

installés dans le bâtiment étudié (Pompes à chaleur,...).

Le bâtiment étudié est équipé de plusieurs systèmes de ventilation : une Centrale de Traitement d'Air (CTA) (double flux auto-réglable) qui assure le renouvellement d'air dans tous les bureaux du bâtiment, une deuxième CTA qui assure le renouvellement d'air dans la salle de réunion ainsi qu'un système de Ventilation Mécanique Contrôlée (VMC) (simple flux auto-réglable) pour les locaux à usage sanitaire.

2.2.2.1 Débits de ventilation

a) Estimation des incertitudes au niveau haut (mesures complètes)

Décomposition de l'incertitude

Les débits d'air réellement insufflés et extraits dans chaque zone du bâtiment ont été mesurés pour pouvoir prendre en compte les pertes dans les conduits de ventilation et l'efficacité des systèmes de ventilation. Les mesures ont été réalisées avec deux cônes de mesure de débit d'air (le choix des cônes a été fait en fonction des dimensions de chaque bouche) associés à un anémomètre à fil chaud. Le cône de mesure permet de canaliser l'air vers la section aéraulique dans laquelle la sonde de l'anémomètre est positionnée. Un débit moyen a ainsi été relevé pour une durée moyenne de 30 secondes.

L'incertitude liée aux débits d'air est due à plusieurs facteurs dont notamment les pertes

aérauliques dans le réseau de distribution, une mauvaise répartition des débits d'un étage à l'autre, une mauvaise étanchéité des gaines de distribution, le type d'appareil de mesure (bolomètre ou débitmètre), l'opérateur de la mesure, en début ou en fin de réseau. Toutes ces causes soulignent toute la difficulté des mesures de débit d'air. Lorsque le renouvellement d'air dans un bâtiment est important, les déperditions au niveau des systèmes de ventilation sont aussi importantes.

Les sources d'incertitudes identifiées lors de la mesure d'un débit de ventilation peuvent ainsi être décomposées en :

- Incertitudes dues aux instruments de mesure : La sonde étant étalonnée à l'intérieur de chaque cône individuellement, une incertitude de mesure a été associée à chaque relevé effectué en fonction du cône utilisé et du mesurage obtenu. En fonction des résultats obtenus lors des mesures, les points d'étalonnage sont utilisés pour estimer l'incertitude sur la correction des mesures effectuées. Cette correction peut se faire par une simple règle de trois ou par une régression linéaire. La sonde a une incertitude liée à sa précision d' $1 \text{ m}^3/h$ et l'incertitude liée à son étalonnage est de $10 \text{ m}^3/h$.
- Incertitudes propres au débit : Les débits de ventilation du bâtiment varient entre 30 et $240 \text{ m}^3/h$ en insufflation et en extraction. Nous avons remarqué que plus les débits étaient élevés ($>60 \text{ m}^3/h$), plus le débit mesuré était faible par rapport à la valeur nominale annoncée (figure 2.7).

Le but est de caractériser l'incertitude sur la mesure des débits d'air sans distinction sur la valeur nominale ou sur la consigne nominale de chaque bouche de soufflage ou d'extraction.

Méthodologie

Pour chaque bouche prise indépendamment, la différence entre le résultat de mesure et le débit nominal a servi à estimer l'incertitude associée à chaque bouche (Équation 2.15). Il a été constaté que les bouches dont le débit nominal est élevé (généralement supérieur à $60 \text{ m}^3/h$) présentent un écart à la moyenne nettement plus important que les bouches à faible débit de ventilation. Nous proposons donc de scinder les débits en deux classes distinctes : ceux dont le débit nominal est inférieur à $60 \text{ m}^3/h$ (à savoir 41 bouches en extraction et en soufflage confondues) et ceux dont le débit nominal est supérieur à $60 \text{ m}^3/h$ (à savoir 13 bouches en extraction et en soufflage confondues). Les débits nominaux des bouches ainsi que les résultats de mesure sont présentés à la table 2.7.

TABLE 2.7 – Débits nominales et débits mesurées aux bouches de ventilation du bâtiment test (m^3/h)

Débits nominales inférieurs à 60 m^3/h			
Débit nominal	Débit mesuré	Débit nominal	Débit mesuré
30	25	30	30
30	24	30	29
30	27	30	36
30	26	30	28
60	38	30	26
30	25	60	50
60	44	30	26
30	28	30	25
30	28	30	33
30	31	60	43
60	50	60	38
60	51	30	35
30	25	30	37
30	26	30	32
30	29	60	52
30	25	30	28
30	20	30	35
30	23	60	45
30	24	30	34
60	49	30	32
60	49	-	-
Écart type expérimental s			8.72 %
Encadrement par la loi de Khi deux σ			10.55 %
Débits nominales supérieurs à 60 m^3/h			
120	63	150	71
150	24	150	62
65	46	180	53
150	72	150	109
130	44	180	240
120	60	240	130
195	176	-	-
Écart type expérimental s			19.9 %
Encadrement par la loi de Khi deux σ			28.4 %

$$u(\%) = \frac{d_{mi} - d_i}{d_i} * 100 \quad (2.15)$$

Où, d_{mi} est le débit réel (mesurée) et d_i le débit nominal de la bouche.

Au final, l'incertitude u associée à chaque bouche de ventilation sera déterminée indépendamment de son débit nominal et sera de 10.55 % pour les débits nominaux inférieurs à $60 \text{ m}^3/h$ et de 19.9 % pour ceux qui sont supérieurs à $60 \text{ m}^3/h$.

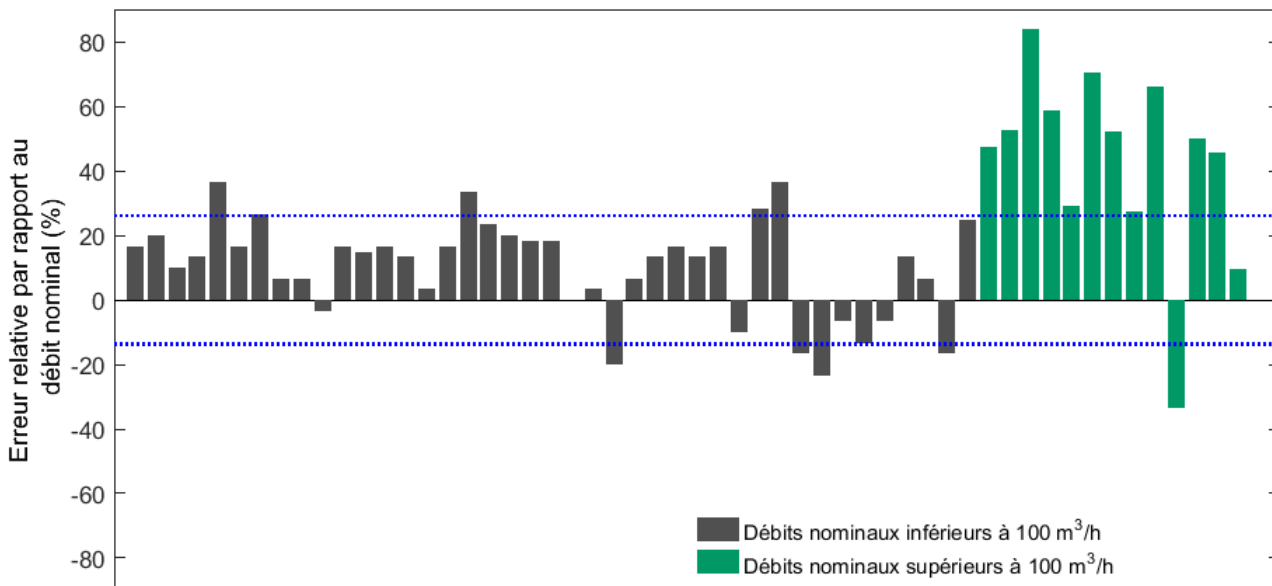


FIGURE 2.7 – Estimation de l'incertitude de mesure des débits de ventilation

b) Estimation des incertitudes au niveau bas (aucune mesure)

Suite à l'étude qu'a menée ARMINES sur l'analyse de sensibilité et d'incertitudes, un intervalle de variation pour les débits de ventilations de $\pm 30\%$. Dans ses recherches, Spitz (Spitz, 2012) a considéré une incertitude sur la mesure des débits de ventilation de $\pm 10\%$. Une incertitude de $\pm 15\%$ a été considérée dans nos travaux.

2.2.2.2 Efficacité des échangeurs double flux

Une efficacité est affectée aux deux centrales de traitement d'air pour l'estimation de la température de l'air insufflé dans chaque zone au détriment de l'efficacité théorique fournie par le constructeur. Toutes des centrales de traitement d'air sont modélisées comme des échangeurs à efficacité constante. La figure 2.8 illustre le principe de fonctionnement d'une

centrale de traitement d'air à double flux. L'air provenant de l'extérieur est réchauffé par la quantité de chaleur extraite de l'air repris dans le bâtiment. Cet air neuf est à son tour réchauffé à la température de consigne avant d'être réinjecté dans le bâtiment.

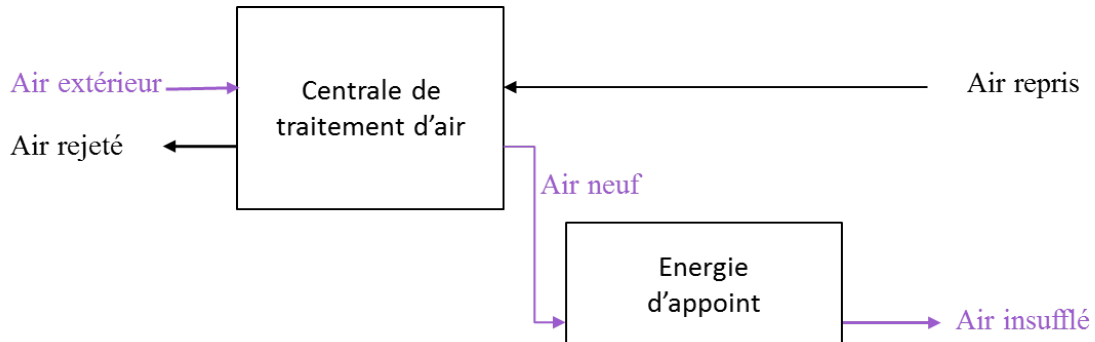


FIGURE 2.8 – Schéma représentatif du principe de fonctionnement d'une des centrales de traitement d'air double flux présente dans le bâtiment test

a) Estimation des incertitudes au niveau haut (mesures complètes)

Méthodologie

Pour estimer cette efficacité, des mesures ont été effectuées dans les gaines aux entrées et aux sorties de la CTA qui assure le renouvellement d'air dans une grande partie du bâtiment.

Décomposition de l'incertitude

Les sources d'incertitude provenant de la mesure de l'efficacité de l'échangeur sont :

- L'incertitude due à la variabilité de l'efficacité des échangeurs au cours du temps ;
- L'incertitude due à la pose des capteurs de température.

L'efficacité d'un échangeur tubulaire double flux à contre courants peut être estimée par :

$$\eta = \frac{\phi}{\phi_{max}} \quad (2.16)$$

$$\begin{aligned} \phi &= c_{pm} \cdot q_m \cdot (T_{ins} - T_{ext}) \\ \phi_{max} &= c_{pmax} \cdot q_{max} \cdot (T_{rep} - T_{ext}) \end{aligned}$$

Où,

η est l'efficacité de l'échangeur considéré ;

c_{pm} et c_{pmax} sont les capacités thermiques des fluides échangés en $J/(kg \cdot ^\circ K)$;

q_m et q_{max} sont respectivement les débits d'air en reprise et en insufflation en m^3/h ;

ϕ et ϕ_{max} sont respectivement le flux effectivement transmis par l'échangeur et le flux de

chaleur maximal qui serait transmis dans des conditions idéales en $J.m^3/h.kg$;

T_{ins} , T_{ext} et T_{rep} sont respectivement la température d'air insufflé dans chaque zone, la température d'air extérieure et la température d'air reprise en $^{\circ}C$.

Le fluide échangé étant de l'air en reprise et en insufflation, la capacité thermique massique est la même en insufflation et en récupération. Le débit maximal est le débit d'insufflation et l'efficacité de l'échangeur est donc estimée finalement par l'équation 2.17.

$$\eta = \frac{q_m \cdot (T_{ins} - T_{ext})}{q_{max} \cdot (T_{rep} - T_{ext})} \quad (2.17)$$

Cas d'étude

Au départ, les capteurs avaient été posés sur les gaines des conduits comme représenté en Annexe 6.5.2, figure 6.4. Les valeurs relevées lors de l'instrumentation étaient donc des données de températures de surface affectées par l'isolation des gaines des conduits de ventilation.

La mesure de l'air insufflé dans les zones du bâtiment a été faite après le caisson de chauffage installé après l'échangeur de la CTA. Ce dispositif permet d'apporter l'énergie complémentaire (énergie d'appoint) celle récupérée dans la CTA, pour chauffer l'air neuf (figure 2.8). Pour cette raison, l'efficacité de l'échangeur η estimée était par moment supérieure à 1 (pendant le fonctionnement du système).

Les capteurs ont été, par la suite, mis à l'intérieur des gaines de ventilation en amont et en aval de la CTA, pour corriger le rendement obtenu précédemment. Les données ont été mesurées pendant une semaine en période hivernale. Après avoir identifié les plages horaires correspondant à chaque régime de fonctionnement (table 2.8), il en a résulté des efficacités différentes pour les régimes 2 et 3, la CTA ayant deux régimes de fonctionnements différents suivant les périodes d'occupation du bâtiment (figure 2.9). Les plages horaires et les régimes de fonctionnement sont présentés en Annexes 6.5, aux tables 6.7 et 6.8.

TABLE 2.8 – Incertitude liée à la mesure de l'efficacité de l'échangeur

Régimes de fonctionnement	Moyenne	Écart type
Régime 2	0.65	0.029
Régime 3	0.62	0.028

En appliquant l'équation 2.2 sur l'efficacité estimée par les résultats de mesure de température (GUM type A), on obtient des écarts type de 0.029 et de 0.028 respectivement pour les régimes 2 et 3. Ces écarts types correspondent à l'erreur due à la variabilité de l'efficacité de l'échangeur dans le temps. Elle caractérise donc une erreur aléatoire or seule

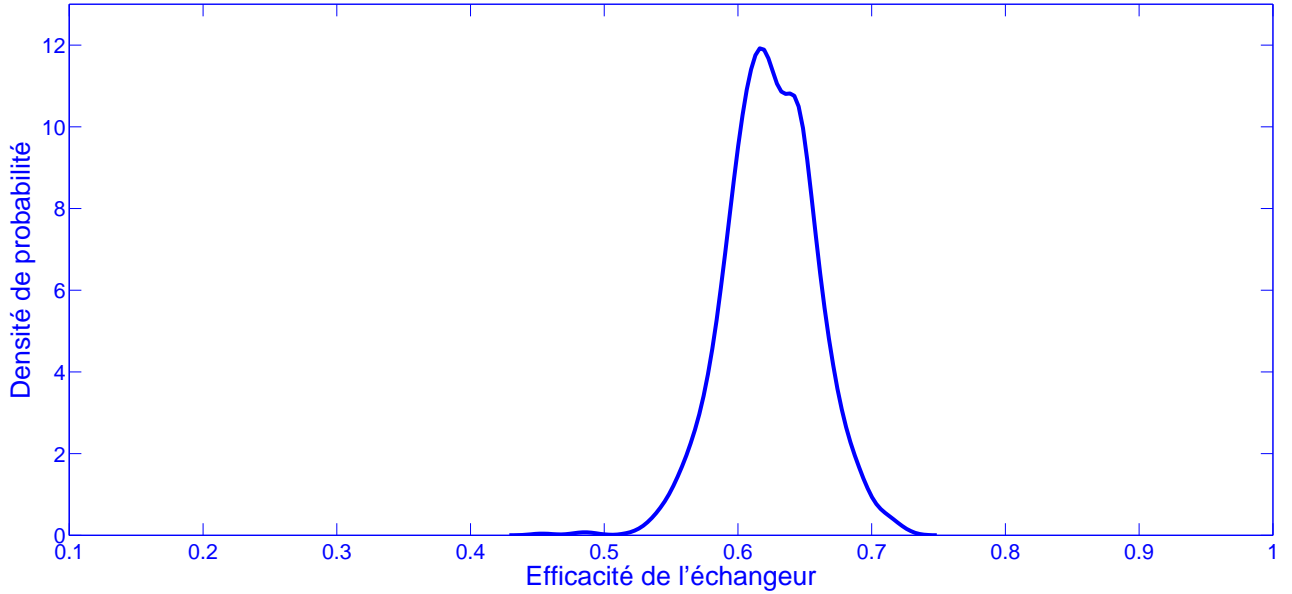


FIGURE 2.9 – Distribution des valeurs de l'efficacité de l'échangeur obtenu par des mesures effectuées pendant les deux régimes de fonctionnement

l'erreur systématique est recherchée.

La variabilité de l'efficacité de la CTA étant faible d'un régime à un autre (table 2.8), nous faisons donc l'hypothèse que la CTA a une efficacité globale de 0.65 et une incertitude u_e (correspondant à un écart type), associée à cette efficacité de 0.029. Ces valeurs seront donc considérées valables pour la deuxième CTA lors de l'analyse d'incertitudes. Par ailleurs, l'incertitude associée à l'efficacité des CTA étant estimée à travers la mesure des températures d'air de chaque bouche, une incertitude de mesure u_p doit être déterminée.

Considérons à présent l'équation 2.17 définie plus haut :

$$\eta = \frac{q_m \cdot (T_{ins} - T_{ext})}{q_{max} \cdot (T_{rep} - T_{ext})}$$

L'incertitude associée à la mesure de l'efficacité de l'échangeur dans le cas de mesure indirecte est donnée par (GUM, 1995) :

$$u^2 = \sum_{k=1}^3 \left(\frac{\delta \eta}{\delta X_k} \right)^2 \cdot u_k^2 \quad (2.18)$$

$$X_k : \{T_{ins}, T_{ext}, T_{rep}\} \quad k \in \{1, 2, 3\}$$

Où u_k est l'incertitude liée à chaque capteur de température.

Ainsi, les dérivées partielles donnent :

$$\frac{\delta\eta}{\delta T_{ins}} = \frac{q_m}{q_{max}} \cdot \frac{1}{T_{rep} - T_{ext}}; \quad \frac{\delta\eta}{\delta T_{ext}} = \frac{q_m}{q_{max}} \cdot \frac{T_{ins} - T_{rep}}{(T_{rep} - T_{ext})^2}; \quad \frac{\delta\eta}{\delta T_{rep}} = -\frac{q_m}{q_{max}} \cdot \frac{T_{ins} - T_{ext}}{(T_{rep} - T_{ext})^2}$$

En émettant l'hypothèse que les températures d'air aux bouches ont été obtenues avec des capteurs de température ayant les mêmes caractéristiques, leurs incertitudes u_k est donc la même ($u_1 = u_2 = u_3$). On déduit de l'équation 2.18 l'équation suivante :

$$u = \frac{\sqrt{2} \cdot q}{q_m} \cdot \frac{u_k}{(T_{rep} - T_{ext})^2} \cdot \sqrt{T_{rep}^2 + T_{ext}^2 + T_{ins}^2 - T_{rep} \cdot T_{ext} - T_{ins} \cdot T_{rep} - T_{ins} \cdot T_{ext}} \quad (2.19)$$

On obtient finalement une incertitude $u = 0.0281$. Elle est liée à la précision des capteurs et à leur étalonnage. Cette incertitude est sensiblement égale à celle estimée précédemment avec la méthode GUM type A. Elle est caractérisée par une erreur qui reste également aléatoire. L'erreur systématique dans le cas de mesures ponctuelles ne peut donc pas être quantifiée car rappelons le, elle est constante ou varie de manière systématique dans le cadre de mesurages répétés (section 2.1.2). On estime dans ce cas de figure que l'incertitude liée à la mesure de l'efficacité de l'échangeur ne peut pas être supérieure à celle estimée. Une incertitude de 2.9% sera alors considérée pour les deux CTA par la suite.

b) Estimation des incertitudes au niveau bas (aucune mesure)

Les efficacités des échangeurs données par les constructeurs sont respectivement de 76.3 % et 89 %. Dans l'étude qu'a menée le Cerema sur les bâtiments démonstrateurs à basse consommation énergétique (PREBAT, 2015), l'efficacité mesurée des échangeurs est le plus souvent supérieure à 70 % et dans les cas où l'efficacité était inférieure à 60 %, des dysfonctionnements existaient dans les bâtiments concernés. En considérant la valeur haute fournie par le constructeur (90 %) et celle généralement constatée sur des bâtiments similaires (70 %), une incertitude de 5 % a été jugée suffisante pour la suite des travaux.

2.2.2.3 La perméabilité à l'air du bâtiment

Construire des bâtiments de plus en plus étanches est un enjeu majeur dans l'amélioration de leur performance énergétique. La perméabilité à l'air est un poste essentiel de déperdition dans les bâtiments à faible consommation énergétique. L'étanchéité du bâtiment est dans la majorité des bâtiments construits aujourd'hui étudiée dans le cadre de la RT 2012 (Réglementation thermique, 2012). Elle est le plus souvent représentée sous la forme d'un débit d'air par m^2 d'enveloppe sous un différentiel de pression de 4 Pascals, le Q_{4Pa} .

a) Incertitudes des incertitudes au niveau haut (mesures complètes)**Méthodologie**

Pour mesurer la perméabilité à l'air des bâtiments, on utilise le principe de la porte soufflante : un ventilateur placé à l'intérieur du bâtiment permet de le maintenir sous un différentiel de pression constant avec l'extérieur. Le débit nécessaire est alors mesuré sous différentes pressions comprises généralement entre 10 et 50 Pa. Le débit sous une pression de 4 Pa est ensuite estimé par régression linéaire.

Décomposition de l'incertitude

Les incertitudes associées à la mesure de la perméabilité à l'air peuvent provenir :

- de la méthode utilisée : pressurisation, dépressurisation,
- du procédé de mesure (instruments de mesure, opérations de mesure : positionnement de la porte soufflante, obstruction des entrées d'air),
- de la dégradation de l'étanchéité du bâtiment au cours du temps.

Cas d'étude

Une mesure de perméabilité à l'air a été réalisée sur le bâtiment CECP à la livraison du bâtiment en septembre 2012. Les résultats de mesures sont représentés dans la table 2.9.

TABLE 2.9 – Mesure de la perméabilité à l'air du bâtiment CECP

Date	$Q_{4Pa}(m^3/h \cdot m^2)$	Incertainitude
Septembre 2012	1.3	0.0475

b) Estimation des incertitudes au niveau bas (aucune mesure)

Une série de mesures de perméabilité a été effectuée sur des bâtiments, dans le cadre de l'opération de recherche PREBAT (PREBAT, 2015) et constituent une base de données riche. Elle concerne les écoles, des locaux sportifs, des restaurants, des commerces, des hôtels mais aussi des bâtiments administratifs et les mesures ont été effectuées dans tout ou une partie du bâtiment. De cette base on tire une moyenne sur la perméabilité de $1.31 m^3/h \cdot m^2$ et un écart type de $1.78 m^3/h \cdot m^2$. Le bâtiment étudié étant un bâtiment administratif et récent, la moitié de cette incertitude soit, $0.89 m^3/h \cdot m^2$, a été considérée pour plus de représentativité.

2.2.3 Incertitudes liées aux sollicitations internes

Au sein du bâtiment, les usagers agissent sur leur environnement direct et aussi sur eux-même pour s'assurer un certain confort. Leur comportement ne suit plus de ce fait les

scénarios conventionnels et idéaux définis par les réglementations en vigueur. Le premier indicateur du confort est la température ambiante. En agissant sur la consigne de température et sur les systèmes, l'occupant a un effet direct sur la consommation d'énergie finale. Les équipements et les appareils électriques utilisés sont aussi des sources d'énergie supplémentaires, qui participent à l'équilibre thermique du bâtiment. Seront alors considérées comme sollicitations internes, la température intérieure, le nombre d'occupants et les apports internes (d'autres variables pourraient également être considérées et quantifier comme par exemple le comportement de l'occupant).

2.2.3.1 La température intérieure

La température intérieure est la température de l'air ambiant d'un local, souhaitée et réglée par ses occupants dans le but d'assurer un certain confort thermique, à travers la température de consigne. La température intérieure est un des paramètres clés influant sur la consommation de chauffage du bâtiment. Pouvoir l'estimer avec précision et pouvoir quantifier son incertitude permettrait de rendre les estimations de performance énergétique plus fiables. Dans bon nombre d'analyses de sensibilité et de propagation d'incertitudes, il lui est associé une plage de variation de ± 1 °C (Spitz et al., 2012), ce qui n'est pas réaliste lorsque des mesures sont effectuées telles qu'exposé ci-dessous.

a) Estimation des incertitudes au niveau haut (mesures complètes)

Communément, dans le cadre d'une STD réalisée sous Trnsys, la grandeur physique modélisée est la température d'air. Cette température d'air, à l'échelle du bâtiment, peut être définie comme étant la moyenne des températures de chaque volume élémentaire à un instant donné t (Équation 2.20). Caucheteux et al. ont assimilé cette problématique à un problème d'inférence statistique en mesurant un échantillon de température en un nombre de points n ; chaque point de mesure donnant la température d'un volume élémentaire du bâtiment.

Méthodologie

On considère le bâtiment avec une infinité de points de mesure de température. On se place alors dans le cadre d'un échantillonnage sans stratification (i.e. la température moyenne du bâtiment est assimilée à la moyenne des températures de l'échantillon). Chacune des zones du bâtiment ayant une consigne de chauffage et un comportement thermique propre, mesurer la température dans un nombre plus élevé de pièces du bâtiment permettrait de réduire l'incertitude sur la température intérieure réelle des locaux.

Soit un bâtiment de volume V , la température du bâtiment T peut être définie comme la moyenne des températures T_i de chaque volume élémentaire v_i à un instant donné t . Ainsi à chaque instant t , la température intérieure sera estimée par :

$$T = \frac{\int T_i \cdot \delta v_i}{V} \quad (2.20)$$

Pour déterminer la température du bâtiment via l'équation 2.20, il convient d'estimer les températures de chaque volume élémentaire v_i à chaque instant donné t . On pourrait ainsi, assimiler chaque volume élémentaire à chaque zone du bâtiment dans laquelle a été réalisée une mesure. Dans le cas où plusieurs capteurs ont été installés dans une même zone, ce volume élémentaire correspondra alors au volume du plus proche voisinage de ces capteurs. Dans le cas contraire où des capteurs n'ont pas été placés dans toutes les zones du bâtiment, la température d'air devra être extrapolée dans les zones où des mesures n'ont pas été effectuées.

Décomposition de l'incertitude

En résumé, l'incertitude associée à la température d'air du bâtiment est composée :

- d'une incertitude due à l'emplacement du capteur dans un volume donné,
- d'une incertitude due à l'instrument de mesure.

b) Estimation des incertitudes au niveau bas (aucune mesure)

Spitz (Spitz, 2012) a soulevé, suite aux mesures in situ mises en place pour ses recherches, la présence de gradients de température notables entre différents étages mais aussi au sein d'une même pièce d'un bâtiment. En réalisant une analyse d'incertitudes sur les maisons INCAS, elle a pu constater qu'il y avait bien un gradient de température au sein d'une même maison, qui pouvait être dû à une mauvaise répartition du système aéraulique entre les étages, mais aussi un gradient de 1 °C au sein d'une même pièce entre le sol et le plafond. Il ressort dans le rapport de capitalisation du Cerema (PREBAT, 2015), qu'il y a un gradient de température de 0 à 3 °C dans les bâtiments tertiaires. Les températures intérieures sont généralement supérieures aux valeurs proposées dans la réglementation thermique, mais présentent une variabilité plus importante dans les bâtiments d'habitation. Aussi, l'étude qu'a réalisée ARMINES sur les bâtiments d'habitation de la FBE (Fondation Bâtiment Énergie) a révélé une variation de température d'air de ± 4.9 °C. Une incertitude de 1.5 °C a donc été adoptée dans ces travaux ce qui correspond à une variation de ± 3 °C.

2.2.3.2 L'occupation

On entend par occupation, le nombre de personnes présentes dans le bâtiment à un instant donné couplé au comportement de ces derniers. L'occupant peut être un poste important de déperdition car il affecte l'équilibre du bâtiment en ouvrant par exemple les portes et les fenêtres, ce qui demande un apport en énergie calorifique supplémentaire, pour

rétablir l'équilibre thermique. La consommation énergétique d'un bâtiment est donc directement liée au taux d'occupation et au comportement de ses usagers (Amayri et al., 2016; Ryu and Moon, 2016; Caucheteux et al., 2013; Nguyen and Aiello, 2013; Menezes et al., 2012; Masoso and Grobler, 2010; Haldi and Robinson, 2011; Hoes et al., 2009). Masoso et al. démontrent dans leurs travaux (Masoso and Grobler, 2010) que plus de 50 % de la consommation totale énergétique est utilisée en dehors des heures d'occupation ce qui est essentiellement due aux appareils laissés en fonctionnement, aux équipements dont les scénarios de consigne sont restés inchangés. Les sources d'incertitudes sont donc multiples et devraient être toutes prises en considération.

Un occupant donné, influence de manière directe ou indirecte l'environnement intérieur du bâtiment. Une personne quelconque apporte par son métabolisme une quantité de chaleur d'environ 100 W au sein de la zone qu'il occupe (apport d'énergie directe). Les appareils dont il fera usage (ménagers dans le cas d'un bâtiment d'habitation et informatiques pour un bâtiment tertiaire) apportent aussi de l'énergie au sein du bâtiment qui sera convertie en apports internes (apport d'énergie indirect). Cet apport d'énergie indirect est dissocié de l'occupation et traité indépendamment sous la forme d'apports internes. On entendra donc par occupation tout au long de cette thèse, le taux d'occupation d'un bâtiment.

a) Estimation des incertitudes liées à la mesure de l'occupation (mesures complètes)

Méthodologie

Pour estimer l'occupation d'un bâtiment, plusieurs estimateurs sont utilisés et sont basés sur les différents protocoles de mesures qui sont mis en place (Caucheteux et al., 2016, 2013). Ils peuvent être basés sur des enregistrements vidéos, de la mesure de concentration en CO_2 , de la détection de mouvements, de l'analyse du bruit ou des données de consommations électriques (Cali et al., 2015; Ansanay-Alex et al., 2016). Ainsi, la combinaison de plusieurs de ces mesures permet d'apporter plus de connaissance sur le taux d'occupation et réduit l'incertitude autour son estimation (Yan et al., 2015). Certains d'entre eux (enregistrements vidéos ou sonores) apportent une information directe mais d'autres estimateurs (la concentration en CO_2 ou la consommation électrique) s'appuient sur la modélisation du mesurage pour fournir l'information nécessaire. Dans ces travaux de thèse, nous nous sommes intéressé à l'estimation de l'occupation d'un bâtiment par la mesure de la consommation électrique.

On définit l'estimateur suivant :

$$(E) : y = ax + b \quad (2.21)$$

Où y est le nombre d'occupants et x la grandeur mesurée (la consommation électrique).

Les coefficients a et b sont estimés pour ensuite déterminer le taux d'occupation. Plusieurs méthodes peuvent être appliquées : un apprentissage peut être réalisé sur une période définie de mesures et les résultats obtenus seront ensuite confirmés par un processus de validation. Durant l'apprentissage, les coefficients a et b pourront être estimés par régression linéaire.

Une alternative à la méthode précédente serait d'identifier les coefficients a et b en émettant des hypothèses sur les consommations électriques effectives liées au taux d'occupation. On considère alors qu'une occupation totale (taux maximal) correspondrait à une consommation électrique maximale et une inoccupation (taux minimal) à une consommation électrique nulle ou minimale fixée. On déduit de ces hypothèses, trois modèles (table 2.10).

TABLE 2.10 – Exemple d'estimateurs pour l'approximation du modèle d'occupation

Estimateurs	x_1	y_1	x_2	y_2	Erreur d'estimation
E_1	0	y_i	x_{max}	y_{max}	e_1
E_2	0	0	x_{max}	y_{max}	e_2
E_3	0	y_i	x_i	y_{max}	e_3

(E_1) : Consommation électrique nulle \rightarrow Occupation fixée y_i ,
 Consommation électrique maximale x_{max} \rightarrow Occupation maximale y_{max} .
 (E_2) : Consommation électrique nulle \rightarrow Occupation nulle,
 Consommation électrique maximale x_{max} \rightarrow Occupation maximale y_{max} .
 (E_3) : Consommation électrique nulle \rightarrow Occupation fixée y_i ,
 Consommation électrique fixée x_i \rightarrow Occupation maximale y_{max} .

Une erreur d'estimation est ensuite déterminée pour chaque modèle en fonction des données d'observation. Elle permettra de sélectionner le modèle le plus performant et la variance des estimateurs u_e sera prise comme étant la composante de l'incertitude associée à l'estimation du nombre d'occupants (Équation 2.22).

$$u_e = \sqrt{\frac{1}{3} \sum_{k=1}^3 (e_k - \bar{e})^2} \quad (2.22)$$

Cas d'étude

$$(E) : y = 0.004 \cdot x \quad (2.23)$$

Les consommations électriques ont servi comme estimateur pour déterminer le taux d'occupation réel du bâtiment. Parmi tous les estimateurs utilisés, il était celui offrant de meilleurs résultats. Des mesures de CO₂ ou des compteurs de présence auraient sans doute fourni de meilleurs estimations mais le bâtiment test n'a pas été muni de ces instrumentations. Les hypothèses définies plus haut ont permis de retenir l'estimateur défini dans l'équation 2.23. L'incertitude sur le taux d'occupation horaire du bâtiment est ainsi déterminée par la composante d'estimation (équation 2.22) par la mesure de la consommation électrique. Cette dernière est estimée par la méthode d'estimation des incertitudes de type A du GUM. L'incertitude finale u est obtenue par encadrement de Khi-deux (équation 2.10) et donne 11 % de la valeur nominale.

b) Estimation des incertitudes au niveau bas (aucune mesure)

Une enquête a été effectuée parmi les usagers du site où est implanté le bâtiment test. Il a été demandé à chaque participant d'estimer le taux d'occupation horaire d'une personne donnée, du bâtiment au cours d'une semaine typique de travail. Cette enquête fournit une estimation de l'occupation du site à chaque plage horaire. On considère par la suite que les résultats fournis par chacun des participants représente un estimateur. Dix estimateurs ont donc été recueillis. En répétant la même méthodologie que dans le cas où des mesures ont été effectuées, on obtient une incertitude d'environ 30 %.

2.2.3.3 Les apports internes

La présence des occupants dans un bâtiment, l'utilisation des équipements et des appareils électriques et l'éclairage constituent des apports internes au bâtiment et peuvent contribuer à réduire la consommation énergétique en chauffage ou en refroidissement. Les apports internes constituent un poste essentiel d'apport de chaleur dans un bâtiment. Un bâtiment tertiaire équipé d'un grand nombre d'appareils électriques verra sa consommation de chauffage diminuée puisqu'une grande partie de la consommation électrique de ces appareils sera restituée en chaleur et constituera un apport calorifique considérable. Rappelons que les apports métaboliques générées par les scénarios d'occupation ne sont pas pris en compte dans les apports internes.

a) Estimation des incertitudes liées à la mesure des apports internes (mesures complètes)

Il est assez complexe de mesurer la quantité de chaleur effectivement restituée par les occupants et les équipements électriques dans un bâtiment. La pratique la plus courante consiste à mesurer les consommations électriques et à émettre une hypothèse forte selon laquelle toute la puissance consommée par les équipements électriques est intégralement restituée au sein du bâtiment. En partant de là, on pourra, en fonction du zonage du bâtiment et de la granularité des mesures électriques qui ont été mises en place, assigner à chaque zone, une puissance dissipée et une incertitude associée. La consommation électrique informe ainsi sur la quantité d'énergie effectivement utilisée par les occupants mais pas sur l'endroit de sa restitution. Un ordinateur peut par exemple être mis en charge dans le bâtiment et être utilisé en dehors lors d'un déplacement professionnel. Le cas inverse est aussi probable.

Décomposition de l'incertitude

Les sources d'incertitudes associées à cette estimation sont :

- La présence effective des occupants au sein du bâtiment,
- La restitution totale de la puissance consommée par les appareils,
- L'uniformité de la répartition de la consommation électrique constatée dans chaque zone,
- Le nombre et la résolution des compteurs électriques.

Méthodologie

Pour l'estimation des incertitudes, nous nous sommes placés dans le cas idéal de la réglementation thermique ([Réglementation thermique, 2012](#)), où tous les occupants sont effectivement présents selon les scénarios d'occupation définis dans un modèle de simulation thermique dynamique. On considère alors que toute l'énergie électrique consommée est entièrement restituée au sein du bâtiment. L'incertitude associée aux apports internes sera donc essentiellement une incertitude de mesure. On considère que la consommation électrique effective totale Y du bâtiment étudié, de surface totale S , est mesurée par n compteurs X donnant chacun une sous-consommation X_k . Chaque compteur a une incertitude associée à la précision de sa mesure notée u_x qui est la même pour tous les compteurs. La consommation électrique totale Y du bâtiment est donc donnée par :

$$Y = \sum_{k=1}^n X_k \quad k \in \{1, \dots, n\} \quad (2.24)$$

Rappelons que l'on cherche à estimer l'incertitude u_i liée à la mesure des apports internes dans une zone i du bâtiment de surface s_i . Les apports internes y_i sont déterminés par :

$$y_i = \frac{s_i}{S} \cdot Y \quad (2.25)$$

De l'équation 2.24, il advient que :

$$y_i = \frac{s_i}{S} \cdot \sum_{k=1}^n X_k \quad (2.26)$$

En utilisant la méthode d'estimation d'incertitude dans le cas de mesures indirectes comme défini dans l'équation 2.18, il en résulte que l'incertitude associée aux apports internes u_i d'une zone i du bâtiment est donnée par :

$$u_i = \frac{s_i}{S} \cdot u_x \cdot \sqrt{n} \quad (2.27)$$

L'équation 2.27 fournit donc une incertitude au prorata des surfaces de chaque zone considérée.

Estimation des incertitudes au niveau bas (aucune mesure)

Dans la RT 2012 ([Réglementation thermique, 2012](#)), il a été défini conventionnellement des apports de chaleur de 14 W/m^2 en période d'occupation. Quand aux retours d'expérience des études menées par le Cerema ([PREBAT, 2015](#)), la puissance dissipée dans tous types de bâtiments confondus varie entre 0 et 9 W/m^2 . l'étude qu'a réalisée ARMINES sur les bâtiments d'habitation de la Fondation Bâtiment Énergie a révélé une augmentation sur les apports internes pouvant aller jusqu'à 20 % avec une valeur moyenne de 2.41 W/m^2 . Une incertitude de 2 W/m^2 a été considérée dans nos travaux, ce qui correspond à une plage de variation de $\pm 4 \text{ W/m}^2$.

2.3 Conclusions et limitations

2.3.1 Analyse des résultats

Le récapitulatif de l'identification des incertitudes attribuées à chaque variable d'entrée est donné dans la table 2.11. Lorsque des mesures n'ont pas été réalisées et que l'objectif de l'étude est par exemple d'aider à la prise de décision en amont de la construction, on attribue généralement des lois de distribution uniformes à chaque variable. Dans le cas contraire, lorsqu'il s'agit par exemple d'effectuer une étude de risques il est préférable d'opter pour des distributions qui représenteraient au mieux les variables de l'étude comme des distributions normales, exponentielles ou gamma, ce qui nécessite la mise en place de mesures complètes. Peu de recherches se sont penchées sur le sujet comme l'a souligné

Tian (Tian, 2013) et généralement, des distributions arbitraires (à dire d'experts) sont attribuées selon des connaissances a priori.

Dans cette thèse, l'intérêt étant porté sur les incertitudes, et non sur les variables en elles-mêmes, il sera considéré que ces incertitudes suivent des lois normales tronquées, afin de garantir la génération de données non aberrantes d'un point de vue physique.

D'une manière générale, on constate que les incertitudes associées aux variables d'entrée sont réduites lorsque des mesures sont réalisées sur ces dernières. En effectuant des mesures sur une quantité, on espère pouvoir réduire son incertitude et aussi pouvoir accroître, sinon améliorer, les connaissances qu'on possédait sur la grandeur mesurée.

TABLE 2.11 – Estimation des incertitudes de mesures et de non mesure liés à chaque grandeur considéré

Grandeurs physiques	Moyenne	Écart type	Unités
Incertitudes au niveau haut (mesures complètes)			
Sollicitations externes			
Température extérieure	-	0.625	[°C]
Rayonnements solaires	-	29.34	[W/m ²]
Vitesse du vent	-	0.22	[m/s]
Distribution aéraulique			
Débits de ventilation BUR	-	10.55	[%]
Débits de ventilation SDR	-	28.4	[%]
Efficacité CTA BUR	-	2.9	[%]
Efficacité CTA SDR	-	2.9	[%]
Perméabilité à l'air	-	0.0475	[m ³ /h/m ²]
Sollicitations internes			
Température intérieure	-	0.38	[°C]
Occupation	-	11	[%]
Apports internes	-	1.23	[W/m ²]
Incertitudes au niveau bas (aucune mesure)			
Sollicitations externes			
Température extérieure	-	1.1	[°C]
Rayonnements solaires	-	33.57	[W/m ²]
Vitesse du vent	-	0.37	[m/s]
Distribution aéraulique			
Débits de ventilation BUR	-	15	[%]
Débits de ventilation SDR	-	15	[%]
Efficacité CTA BUR	-	5	[%]
Efficacité CTA SDR	-	5	[%]
Perméabilité à l'air	-	1.78	[m ³ /h/m ²]
Sollicitations internes			
Température intérieure	-	1.5	[°C]
Occupation	-	30	[%]
Apports internes	-	2	[W/m ²]

Dans certains cas, les mesures ne sont pas directement effectuées sur la variable d'intérêt. On améliore certes la connaissance que l'on avait mais la fiabilité des connaissances obtenues est quelque peu remise en question puisque les résultats ont été obtenus de manière indirecte. En prenant l'exemple de l'estimation de l'incertitude associée à l'efficacité des échangeurs (section 2.2.2.2) par la mesure des températures d'air dans les bouches d'extraction et de soufflage, on se rend compte que le résultat obtenu est complètement dépendant des mesures effectuées. Le fait d'avoir une mesure erronée peut induire un fort biais sur les résultats. D'autre part, toujours dans le cas des échangeurs, les mesures n'ont pas été effectuées sur les débits effectifs de soufflage et de reprise dans les bouches. On a considéré que l'incertitude sur ces débits était nulle donc que la connaissance sur ces dernières était certaine et que les quantités d'air réellement insufflé ou repris par la CTA sont égales au débit nominal fournit par le constructeur. Sans nul doute, cette hypothèse peut être fortement remise en cause car il peut y avoir des pertes dans les réseaux de ventilation ou une baisse de rendement au niveau du ventilateur. L'incertitude réelle associée à l'efficacité de la CTA serait donc plus élevée que celle estimée.

En dehors des sources d'incertitude, le procédé de mesures joue aussi un grand rôle dans les résultats obtenus. En prenant l'exemple des variables météorologiques, les incertitudes ont été quantifiées sur la base de plusieurs stations météo implantées sur site. Ce sont donc des incertitudes de mesure. La durée des mesures effectuées d'une station à une autre, la saisonnalité (le fait que les stations ont enregistré des données à des saisons différentes), le nombre de stations peuvent être des facteurs d'incertitudes supplémentaires. Puisqu'il est question dans cette thèse d'évaluer la variabilité inter-annuelle et non des conditions météorologiques, l'estimation des incertitudes devient très vite délicate.

Au final, la quantification des incertitudes des variables d'entrée est effectuée en fonction des moyens dont on dispose, des mesures mises en place et des informations que l'on a en amont ou a priori de l'étude. Bien souvent, les mesures sont mises en place en fonction de l'objectif de l'étude et compte tenu des investissements financiers et en temps qu'elles induisent, on est souvent contraint dans le choix du procédé de mesure. On rejoint au final la problématique du plan de mesure et de vérification. On se rend aussitôt compte que pour maîtriser les incertitudes sur la consommation énergétique finale et donc les risques liés à la mise en place d'une garantie de performance, l'étape de quantification des incertitudes est fondamentale et doit être réalisée avec rigueur.

2.3.2 Conclusion

Une attention particulière a été portée sur la quantification des incertitudes des variables d'entrée dans ce chapitre, suivant deux situations : lorsque des mesures ont été effectuées pour quantifier l'incertitude associée (incertitude de mesure au niveau haut) et lorsque l'on se base sur des connaissances antérieures, pour attribuer une incertitude aux variables en

jeu (incertitude de non mesure au niveau bas). Les différentes méthodologies qui ont été proposées sont basées sur les ressources disponibles (par le biais d'estimateurs, des incertitudes composées) mais aussi sur la fiabilité des résultats obtenus (estimation du niveau de confiance voulu par encadrement de Khi-deux).

Au vu des résultats de la quantification et de l'ampleur des mesures, il s'avère nécessaire de pouvoir hiérarchiser les variables selon leur influence respective sur la variable d'intérêt (consommation annuelle de chauffage). Une fois la hiérarchisation effectuée, il faut ensuite considérer le caractère stochastique des variables et modéliser celles-ci par des champs aléatoires. Ces derniers feront alors l'objet de discrétisation et de propagation à travers un modèle physique créé sous Trnsys, pour en étudier l'impact sur la consommation annuelle de chauffage. Le chapitre suivant propose une méthodologie de création d'un modèle de simulation qui couple la physique du bâtiment et le caractère stochastique des variables d'entrée.

Dans la suite des travaux on considérera les incertitudes suivant les deux niveaux présentés dans la table 2.11. Les annotations suivantes seront adoptées pour toutes les variables d'entrée.

TABLE 2.12 – Annotation assignée à chaque variable d'entrée

Notation	Variable
X_1	Température extérieure
X_2	Rayonnement solaire
X_3	Vitesse du vent
X_4	Débits de ventilation de la CTA BUR
X_5	Débits de ventilation de la CTA SDR
X_6	Efficacité de la CTA BUR
X_7	Efficacité de la CTA SDR
X_8	Perméabilité à l'air
X_9	Température intérieure
X_{10}	Occupation
X_{11}	Apports internes

Consolidation et réduction des modèles

Introduction

Le chapitre 2 nous a permis de caractériser les incertitudes des variables d'entrée des modèles suivant deux cas de figure. Pour obtenir une bonne estimation de la variable d'intérêt en sortie du modèle la prochaine étape consiste à construire un modèle de simulation. Ce modèle permettra de représenter le bâtiment étudié le plus fidèlement possible et servira de base aux analyses de sensibilité et d'incertitudes menées. Ces dernières permettront d'identifier les variables les plus influentes sur la sortie du modèle mais aussi d'étudier l'effet des incertitudes associées à ces variables sur la sortie du modèle de simulation. Ces analyses sont le plus souvent coûteuses en temps de calcul. Une des solutions est de construire une surface de réponse $\hat{\mathcal{M}}$, qui serait une approximation du modèle initial \mathcal{M} en tout point des variables d'entrées. Nous présentons dans un premier temps les méthodologies d'analyse de sensibilité et d'incertitudes de manière exhaustive puis celles de réductions de modèles existantes. Puis nous développerons les méthodes retenues et appliquées au cas réel.

3.1 Analyse de sensibilité

Réaliser une analyse de sensibilité sur les variables d'entrée d'un modèle permet de déterminer les variables ayant une influence importante sur la réponse du modèle. Elle permet aussi, selon la méthode employée, de quantifier la participation de chaque variable à la variation de cette même réponse. En identifiant ainsi les variables auxquelles la réponse du modèle est sensible et celles ayant peu voire pas d'impact, il est possible d'orienter l'analyse d'incertitude qui va suivre en se concentrant sur les variables les plus pertinentes (Saltelli et al., 2000). Plusieurs méthodes existent pour réaliser une analyse de sensibilité et peuvent être répertoriées suivant leur approche ou leur finalité. De nombreux auteurs ont proposé de répertorier ces méthodes suivant plusieurs critères (Iooss and Saltelli, 2015; Tian, 2013; Heiselberg et al., 2009; Frey et al., 2003; Hamby, 1994) :

- L'approche : statistiques, mathématiques ou graphiques ;
- L'analyse : impact local ou global ;
- La finalité : qualitative ou quantitative.

Un état de l'art des méthodes d'analyse de sensibilité appliquées au domaine du bâtiment a été effectué par Tian (Tian, 2013). Il explique la méthodologie de chacune d'entre elles et soulève leurs points forts ainsi que leurs limitations. D'autres comparaisons ont aussi été effectuées sur différents critères notamment par Nguyen et al. (Nguyen and Reiter, 2015), quant à la hiérarchisation des variables, à leur effet respectif, à leur interaction mutuelle ainsi qu'au temps de calcul ; ou par Marrel (Marrel et al., 2011) quant au criblage des variables influentes ou à la décomposition de la variance.

De nombreuses recherches en performance énergétique des bâtiments ont été menées en réalisant des analyses de sensibilité (Rivalin, 2016; Anstett-Collin et al., 2015; Spitz et al., 2012; Hopfe and Hensen, 2011; Tian and De Wilde, 2011; Heiselberg et al., 2009) et le choix de la méthode appropriée dépendra essentiellement de la finalité de l'étude. Elle peut aussi dépendre du coût en temps de calcul du modèle physique utilisé, du nombre de variables traitées et de la maîtrise des différentes méthodes. Un bref aperçu des méthodes d'analyse de sensibilité locale et globale est donné et la méthodologie d'analyse de sensibilité par screening est décrite dans les sections qui suivent.

3.1.1 Méthodes d'analyse de sensibilité

3.1.1.1 Analyse de sensibilité locale et globale

Les méthodes d'analyse de sensibilité locale permettent de visualiser l'influence d'une variable choisie sur la réponse du modèle en imposant une variation autour de sa valeur nominale. Ces méthodes reposent ainsi sur l'analyse d'une variable à la fois (One At a Time ou OAT) pour déterminer la contribution de la dite variable à la variation de la

réponse du modèle. On suppose ainsi qu'il existe une relation linéaire entre les entrées et la sortie du modèle. Les interactions ne sont donc pas étudiées pour éviter les effets de compensation et garantir la linéarité du modèle. Les analyses de sensibilité locale sont simples à mettre en place.

Les méthodes d'analyse de sensibilité globale (Marrel et al., 2011; Saltelli et al., 2008) sont similaires aux méthodes par analyses de sensibilité locale à la seule différence que toute la plage de variation de la variable est explorée. Elles s'appuient sur la loi de distribution associée à la variable et permet la variation de toutes les variables impliquées à la fois. Elles permettent ainsi de réaliser une analyse quantitative pour attribuer la part de la variabilité de la réponse du modèle induite par une variable donnée en décomposant leur variance ou en réalisant une régression linéaire sur la sortie du modèle obtenue. Elles sont coûteuses en temps de calcul et les résultats ne sont pas optimaux lorsque la relation entre les variables d'entrées n'est pas linéaire.

Les méthodes de Sobol, FAST, de Monte Carlo sont les méthodes d'analyse de sensibilité globale les plus courantes. En général, les méthodes d'analyses de sensibilité globale sont les plus adaptées pour déterminer les variables les plus influentes sur la performance énergétique des bâtiments. Cependant les analyses de sensibilité locale restent très populaires (Tian and De Wilde, 2011) car elles restent moins coûteuses en temps de calcul et permettent d'avoir un aperçu qualitatif rapide du classement des variables.

3.1.1.2 Analyse de sensibilité par screening

En vue d'identifier les variables les plus influentes, de les hiérarchiser et de ne retenir que celles qui sont vraiment significatives, on réalisera une analyse de sensibilité par la méthode de screening. La méthode d'analyse de sensibilité par screening permet d'avoir des informations qualitatives (classement par ordre d'importance) sur la sortie du modèle (Campolongo et al., 2007).

Elle peut être effectuée de différentes manières. La méthode la plus simple pour étudier la sensibilité des paramètres d'entrée d'un modèle consiste à faire varier une variable à la fois sur toute sa plage de variation et à en évaluer l'impact (méthode OAT). Aussi, les plans d'expériences, permettent d'assigner aux variables différentes valeurs à chaque itération, afin d'évaluer leurs effets respectifs sur la réponse du modèle avec une grande précision tout en évaluant leurs interactions. Cette méthode est plus adaptée lorsque peu de variables ont un impact considérable sur la sortie du modèle mais elle s'avère être très coûteuse en temps de calcul.

Il est également possible de réaliser une analyse de sensibilité en regroupant les variables afin de tester l'influence de chaque groupe de ces variables sur le modèle considéré

en réalisant du screening de groupe. Les groupes les moins impactant sont ainsi éliminés progressivement et les plus influents analysés. Parmi les méthodes de screening, la méthode de Morris a été retenue dans le cadre de nos recherches pour sa robustesse. Elle nous permettra de réaliser l'analyse de sensibilité précédant la propagation d'incertitudes qui sera menée plus loin.

3.1.2 Criblage par la méthode de Morris

Pour réaliser une analyse de sensibilité par screening, la méthode la plus efficace et la plus robuste est sans doute la méthode de Morris (Morris, 1991). Elle permet de classer les variables selon leur influence en ciblant ceux qui ont un effet négligeable sur la sortie du modèle tout en identifiant les effets non linéaires.

Lorsque le modèle étudié présente un grand nombre de variables en entrée ou qu'il nécessite un temps de simulation assez long, la méthode de Morris permet de faire un premier tri parmi les variables et de les hiérarchiser. Elle tire ses racines de la méthode OAT classique, où on fait prendre à chaque variable une valeur au sein de sa plage de variation pour ensuite en vérifier l'impact sur la sortie du modèle (Daniel, 1973). La méthode de Morris a subi plusieurs évolutions notoires notamment avec Campolongo et Cropp (Campolongo et al., 2007; Cropp and Braddock, 2002).

Elle est moins coûteuse en temps de calcul (comparativement à la méthode OAT et aux plans d'expérience) et ne nécessite pas la connaissance des distributions des variables. Seules leurs plages de variation et leurs valeurs moyennes sont nécessaires. Néanmoins, un des inconvénients de cette méthode est qu'elle ne permet pas de réaliser par la même occasion une analyse d'incertitude étant donné qu'elle ne tient pas compte des distributions.

Principe de la méthode

Soit un modèle théorique \mathcal{M} dont l'ensemble des entrées incertaines est donné par $X = \{X_1, X_2, \dots, X_N\}$. La réponse du modèle Y est donnée par l'équation suivante :

$$Y = \mathcal{M}(X_1, X_2, \dots, X_N) \quad (3.1)$$

La plage de variation, d'une variable X_i considérée, est d'abord discrétisée en q niveaux équiprobables selon un échantillonnage uniforme au sein des plages définies des valeurs possibles de chaque variable. Pour une trajectoire définie r_k , on attribue aléatoirement une valeur à une variable X_i^k , parmi ses niveaux q , toutes les autres variables restant fixes par ailleurs et, on évalue la réponse spécifique du modèle :

$$P^{(k)} : Y_k = \mathcal{M}(X_1, \dots, X_i^k, \dots, X_N) \quad (3.2)$$

Au niveau suivant de la trajectoire $P^{(k)}$, un niveau est assigné à la variable suivante tout en maintenant les valeurs des autres variables à leur niveau précédent et l'effet résultant de cette modification par rapport à la simulation précédente de la trajectoire est estimé (Équation 3.2). Ce processus est ainsi répété jusqu'à ce que toutes les variables soient modifiées. Ce qui revient à effectuer un nombre $N + 1$ de simulations. On effectue ainsi un nombre r de trajectoires aléatoires en faisant varier une variable à la fois selon un plan OAT classique.

Ainsi, des $(N + 1) \cdot r$ simulations, on peut déduire sur l'ensemble des trajectoires effectuées les éléments statistiques définis dans les équations 3.3, 3.4 et 3.5 qui permettent de caractériser la sensibilité du modèle \mathcal{M} , aux variables d'entrées X . On pourra déterminer l'importance d'une variable X_i donnée, en estimant :

(i) Son effet élémentaire :

$$\Delta_i^{(k)} = \frac{\mathcal{M}(P_{i-1}^{(k)}) - \mathcal{M}(P_i^{(k)})}{\delta} \quad (3.3)$$

(ii) Son effet moyen ainsi que son effet global :

$$\mu_i = \frac{1}{r} \cdot \sum_{k=1}^r \Delta_i^{(k)} \quad \text{et} \quad \mu_i^* = \frac{1}{r} \cdot \sum_{k=1}^r |\Delta_i^{(k)}| \quad (3.4)$$

(iii) L'écart type de ses effets élémentaires (dispersion) :

$$\sigma_i = \sqrt{\frac{1}{r-1} \cdot \sum_{k=1}^r (\Delta_i^{(k)} - \mu_i)^2} \quad (3.5)$$

Ces indices vont permettre de rendre compte des interactions et des effets linéaires ou non des variables sur la sortie du modèle. On préférera utiliser μ_i^* à μ_i pour éviter un effet de compensation.

On construit ensuite un graphe permettant de représenter les variables en fonction de leurs influences (figure 3.1) : on calcule pour ce faire la distance entre le point considéré et l'origine du repère par l'équation 3.6. Les variables ayant une moyenne μ_i importante et une forte dispersion σ_i seront considérées comme influentes et à effet non linéaires et par opposition celles ayant un effet élémentaire μ_i et une dispersion σ_i faibles seront considérées comme ayant un effet négligeable sur la sortie du modèle (Saltelli et al., 2000).

$$\xi_i = \sqrt{\mu_i^{*2} + \sigma_i^2} \quad (3.6)$$

Ces indicateurs permettent une interprétation aisée de l'analyse de par leur simplicité. Cependant, on obtient un classement par ordre d'importance sur la sortie du modèle et

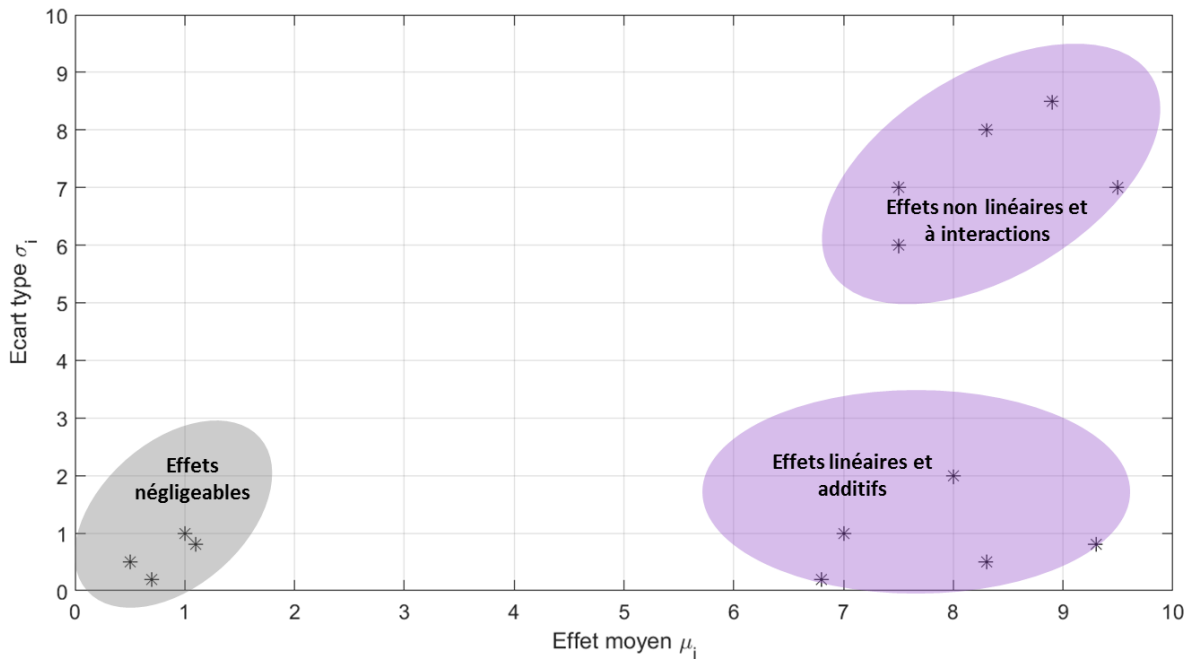


FIGURE 3.1 – Représentation des résultats de l’analyse de sensibilité par la méthode de Morris

non la quantification de leur contribution réelle à la variabilité de cette dernière. Cette information qualitative peut être considérée comme satisfaisante pour identifier les variables les plus influentes dans le but de construire un méta-modèle fiable et robuste et ensuite réaliser une analyse d’incertitudes.

3.2 Propagation d’incertitudes

La propagation d’incertitudes consiste à propager les incertitudes des variables d’entrée à travers un modèle de simulation pour obtenir l’incertitude sur la variable de sortie souhaitée. Elle permet de prendre en considération les caractéristiques inhérentes aux variables d’entrée (moyenne, écart type et distribution), pour avoir une bonne représentativité de la sortie du modèle (figure 3.2). Plusieurs méthodes permettent d’effectuer une propagation d’incertitudes. Parmi elles, la méthode GUM ou méthode du cumul quadratique et la méthode de Monte Carlo sont présentées ci-dessous.

3.2.1 Méthode de propagation d’incertitude proposée par le GUM

Le GUM (GUM, 1995) propose une méthodologie de propagation d’incertitudes, appelée méthode du cumul quadratique basée sur la décomposition de la variance des entrées. Lorsque la variable d’intérêt Y dont on souhaite quantifier l’incertitude n’est pas mesurée

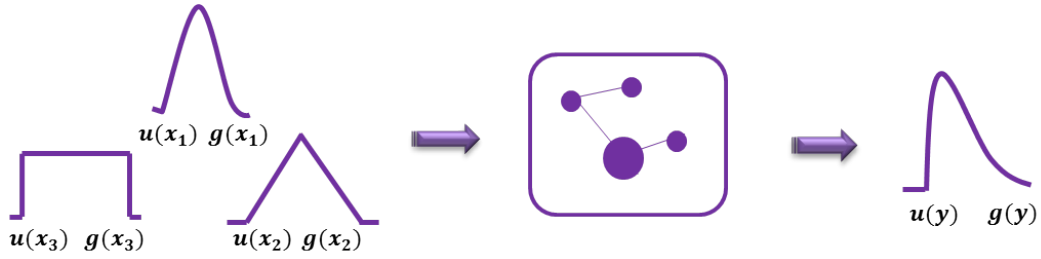


FIGURE 3.2 – Illustration d'une propagation d'incertitudes à travers un modèle de simulation

directement et est plutôt déterminée par l'intermédiaire des mesures effectuées sur d'autres variables $\{X_1, X_2, \dots, X_N\}$, il est possible d'en déterminer l'incertitude par la relation fonctionnelle \mathcal{M} qui les lie (équation 3.7).

$$Y = \mathcal{M}(X_1, X_2, \dots, X_N) \quad (3.7)$$

Elle permet de déterminer la meilleure estimation de la variable d'intérêt Y ainsi que l'incertitude qui lui ait associée $u(Y)$ à partir de l'estimation des incertitudes des variables d'entrée, $\{X_1, X_2, \dots, X_N\}$. L'incertitude de la variable de sortie Y est obtenue en propageant les incertitudes des variables d'entrée : la méthode GUM est donc dite méthode de propagation d'incertitudes. Lorsque les variables en entrée du modèle sont corrélées, l'incertitude associée au résultat de mesure $u(Y)$, est donnée par l'équation 3.8.

$$u^2(Y) = \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \right)^2 \cdot u^2(x_i) + 2 \cdot \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left(\frac{\partial f}{\partial x_i} \right) \cdot \left(\frac{\partial f}{\partial x_j} \right) \cdot u(x_i, x_j) \quad (3.8)$$

Où, x_i et x_j sont des estimations des paramètres X_i et X_j et $u(x_i, x_j)$ est la covariance associée aux variables X_i et X_j . Le modèle peut être ensuite linéarisé par un développement de Taylor au premier ou au second ordre. Les dérivées premières et secondes sont ensuite calculées. Ainsi, pour effectuer une propagation d'incertitudes avec la méthode GUM, l'opérateur n'a pas besoin de réaliser des calculs compliqués. La méthodologie est clairement expliquée dans le guide GUM et est facile à appliquer.

3.2.2 Méthode de propagation d'incertitude par Monte Carlo

Soit un modèle théorique \mathcal{M} dont l'ensemble des entrées incertaines est donné par $X = \{X_1, X_2, \dots, X_N\}$. La réponse du modèle Y est donné par l'équation suivante :

$$Y = \mathcal{M}(X_1, X_2, \dots, X_N) \quad (3.9)$$

Un supplément au GUM, le GUM S1 (BIPM et al., 2008), propose une méthodologie

de propagation d'incertitudes basée sur la distribution des variables d'entrée. Elle permet la propagation des fonctions de densité de probabilité des variables d'entrée à travers le modèle \mathcal{M} (équation 3.9). L'estimation des incertitudes par la méthode de Monte Carlo se fait donc par la propagation des distributions des variables en entrée à travers le modèle dans le but d'obtenir la distribution de la variable d'intérêt Y .

Lorsque les distributions sont définies pour les variables d'entrée, un nombre de simulation M est choisi. Généralement, plus M est élevé, meilleur sont les résultats. La propagation d'incertitudes par la méthode de Monte Carlo est réalisée comme suit. Après avoir défini le modèle \mathcal{M} sur lequel les incertitudes seront propagées :

1. Une distribution $p(X_i)$ est affectée à chaque variable d'entrée X_i en fonction des incertitudes qui lui sont associées ;
2. Un nombre M de Monte Carlo est choisi ;
3. M tirages $X_{i1}, X_{i2}, \dots, X_{iM}$, de chaque variable d'entrée X_i sont effectués et propagés dans le modèle \mathcal{M} pour obtenir M valeurs Y_1, Y_2, \dots, Y_M de la variable d'intérêt Y (équation 3.10).

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_M \end{bmatrix} = \mathcal{M} \left(\begin{bmatrix} X_{11} & X_{21} & \dots & X_{N1} \\ X_{12} & X_{22} & \dots & X_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1M} & X_{2M} & \dots & X_{NM} \end{bmatrix} \right) \quad (3.10)$$

La méthode de Monte Carlo est ainsi appelée méthode de propagation des distributions. Elle est aussi facile à implémenter et fournit des informations concernant la distribution de la variable d'intérêt. Cependant, elle a aussi des inconvénients ; le temps de simulation augmente très rapidement lorsque M est très élevé et l'attribution des fonctions de distributions aux entrées peut être difficile lorsque les données ne sont pas précises ou que l'on ne dispose pas d'assez d'informations. Usuellement, une approximation est faite à partir des données et des informations disponibles. Aussi, en utilisant cette méthode, l'opérateur ne fait pas de distinction entre les différentes sources d'incertitudes (quantitatives et qualitatives).

On détermine ainsi la moyenne associée à l'estimation de la quantité d'intérêt :

$$\tilde{Y} = \frac{1}{M} \sum_{i=1}^M Y_i \quad (3.11)$$

Ainsi que l'incertitude qui lui est associée :

$$u^2(\tilde{Y}) = \frac{1}{M-1} \sum_{i=1}^M (Y_i - \tilde{Y})^2 \quad (3.12)$$

3.3 Réduction de modèle et construction d'un méta-modèle

Le nombre de simulations à réaliser pour effectuer une analyse de sensibilité ou une propagation d'incertitudes, est souvent très important. Ces analyses deviennent très vite coûteuses en temps de calcul si le modèle \mathcal{M} est très complexe ou a été réalisé à un niveau de granularité affinée. Pour réduire les temps de calcul, il est possible de construire un méta-modèle ou une surface de réponse, $\widehat{\mathcal{M}}$ qui est une approximation du modèle existant \mathcal{M} et qui aurait un comportement similaire sur un espace prédéfini. On construit ainsi un modèle représentatif à partir d'un certain nombre de simulations du modèle original.

Plusieurs méthodes permettent d'effectuer des réductions de modèle telles que les polynômes du chaos (Blatman, 2009; Riahi, 2013), les Processus Gaussiens (GP) mais aussi les réseaux de neurones artificiels (ANN), . . . , et le choix de la méthode dépend essentiellement de l'objectif de l'étude (Chen et al., 2006). Cependant, quelle que soit la méthodologie choisie, le méta-modèle $\widehat{\mathcal{M}}$ devra permettre d'être évalué plus rapidement que le modèle initial \mathcal{M} .

En considérant l'exemple des méta-modèles construits par processus gaussiens (GP), on obtient en aval de l'étude un niveau de confiance sur les résultats obtenus grâce au méta-modèle mais ses paramètres doivent être estimés avec précision. Les modèles réduits par GP, comme lors de toute réduction, sont construits par une approximation de la réponse du modèle sur la plage de variation des variables d'entrée. Tous les points d'une même variable étant associés à une probabilité différente, cette approche est peu efficace dans le cadre d'une analyse d'incertitudes. Heo et Burkhart ont utilisé les méta-modèles par processus gaussien dans une démarche de mesures et de vérification pour estimer les gains énergétiques dans des bâtiments tertiaires (Heo et al., 2013; Burkhart et al., 2014). Il est nécessaire d'utiliser une méthodologie qui tient compte des entrées aléatoires : un méta-modèle stochastique.

De manière générale, la construction d'un méta-modèle nécessite moins d'appels au modèle initial \mathcal{M} qu'en réalisant une analyse de sensibilité ou une propagation d'incertitudes. Le coût de calcul en construction reste cependant acceptable car il est réduit de plus de 90% par rapport au modèle initial après la construction, ce qui permettrait de gagner en temps de calcul lors des analyses effectuées. L'opérateur pourrait ainsi effectuer des recherches supplémentaires et plus approfondies pour un coût en temps de calcul équivalent. Ainsi la qualité d'un métamodèle est jugée sur sa capacité d'approximation du modèle initial mais aussi sur ses aptitudes à effectuer des prédictions en un temps de calcul faible.

De nombreuses recherches se sont aussi basées sur la construction de méta-modèle dans le domaine de la thermique du bâtiment. R. Merheb (Merheb, 2013) a utilisé la réduction de modèle pour effectuer une étude d'incertitudes sur le besoin annuel de chauffage et l'inconfort d'été à l'échelle du bâtiment. M. Rabouille (Rabouille, 2014) a utilisé la réduction de modèle par polynômes du chaos dans ses travaux pour étudier la performance des bâtiments réhabilités en simulation thermique dynamique. L. Rivalin (Rivalin, 2016) a étudié l'applicabilité de la réduction de modèles par polynômes du chaos dans une démarche de garantie de performance énergétique. Dans la suite, la méthodologie de construction d'un méta-modèle par polynômes du chaos est présentée.

3.3.1 Utilisation d'un méta-modèle : Les polynômes du chaos

La construction d'un méta-modèle par polynômes du chaos est basée sur la projection des réponses du modèle initial dans la base de polynômes multivariés définis par les entrées du modèle. L'orthogonalité de la base permet ainsi de fournir une seule et unique représentation du modèle.

Considérons le modèle \mathcal{M} introduit précédemment à l'équation 3.1, ainsi que ses entrées incertaines $X = \{X_1, X_2, \dots, X_N\}$. Les variables d'entrée du modèle étant affectées d'incertitudes, les polynômes du chaos permettent de représenter le modèle dans un espace probabiliste (Wiener, 1938) sous la forme :

$$\hat{\mathcal{M}}(X) = \sum_{k=0}^{\infty} \beta_k \psi_k(X) \quad (3.13)$$

Où β_k sont des coefficients déterministes et ψ_k les polynômes multivariés de l'expansion polynomiale.

La réponse du modèle est décomposée dans une base orthogonale. L'orthogonalité étant définie par rapport à la densité de probabilité de chaque variable d'entrée, on considère donc un espace euclidien fait de polynômes orthogonaux, l'espace de Hilbert.

Soit deux polynômes, ψ_l et ψ_m définis dans cet espace. Le produit scalaire de ψ_l et ψ_m est défini par :

$$\langle \psi_l, \psi_m \rangle = \int \psi_l(X) \psi_m(X) p(X) dX \quad (3.14)$$

Où p est la fonction poids. $\forall \psi_l$ et ψ_m orthogonaux,

$$\langle \psi_l, \psi_m \rangle = 0 \quad (3.15)$$

Différentes familles de polynômes peuvent être utilisées pour caractériser les variables d'entrée. La première famille est celle des polynômes d'Hermite pour les variables d'entrée gaussiennes (chaos homogène) mais d'autres familles de polynômes ont suivi et ont permis de caractériser d'autres distributions (table 3.1). Lorsqu'une famille de polynômes ne peut pas être associée à la distribution d'une variable d'entrée, une transformation isoprobabiliste (qui consiste à transformer point par point une densité quelconque en son équivalente gaussienne par exemple) peut être effectuée.

TABLE 3.1 – Type de polynômes en fonction des distributions des variables d'entrée

Polynômes	Expression	Distribution	Domaine de définition
Hermite	$(-1)^p e^{x^2/2} \frac{d^p}{dx^p} (e^{-x^2/2})$	Normale	\mathbb{R}
Legendre	$\frac{1}{2^p} \sum_{k=0}^p \binom{k}{p}^2 (x-1)^{p-k} (x+1)^k$	Uniforme	$[-1, +1]$
Laguerre	$\frac{e^{-x}}{p!} \frac{d^p}{dx^p} (e^{-x} x^p)$	Gamma	$\left[0, +\infty\right[$

Pour construire le méta-modèle, l'approche la plus courante consiste à sélectionner des polynômes ψ_K de telle sorte que la somme des degrés des polynômes n'excède pas un degré choisi p . L'utilisateur prédéfinira alors un degré maximum p en fonction de la précision d'approximation souhaitée. Il est important de choisir minutieusement l'ordre du développement polynomial. En choisissant un ordre trop faible, la réponse du modèle ne serait pas parfaitement approximée et en choisissant un ordre trop élevé, on augmente le nombre de coefficients à estimer donc le nombre de simulations à réaliser. Le choix de l'ordre de l'approximation polynomiale a donc une influence sur la précision des résultats mais aussi sur le temps de calcul.

Pour conclure, on construit donc en premier les polynômes orthogonaux associés à chaque variable d'entrée et on estime les coefficients en exploitant l'orthogonalité de la base. Les coefficients peuvent être estimés par différentes méthodes. Les méthodes les plus courantes parmi les méthodes non intrusives¹, sont celles par projection et par régression.

La méthode de projection consiste à rechercher un polynôme de plus petit degré p qui permet d'obtenir pour chaque valeur prise par les variables d'entrée tout au long de leur plage de variation, la réponse du modèle correspondante. On choisit alors minutieu-

¹Une méthode est dite non intrusive lorsqu'elle ne nécessite que la connaissance des variables d'entrée et de leurs correspondances respectives en sortie du modèle pour être utilisée.

sement les points auxquels l'interpolation sera réalisée en effectuant un échantillonnage par la méthode de Sobol par exemple (section 3.3.3). La deuxième approche consiste à estimer les coefficients du méta-modèle, pour un degré p prédéfini, en minimisant l'erreur d'approximation appelée méthode de régression. Les coefficients sont estimés au sens des moindres carrés en tenant compte de la troncature². La méthode de régression semble être la plus adaptée pour réduire le nombre d'appels au modèle initial. Néanmoins, un des points critiques de cette méthode réside dans le choix des points pour la réalisation des plans d'expériences (section 3.3.3).

3.3.2 La méthode de régression

Le principe de la méthode de régression est d'estimer les coefficients du polynôme en effectuant une approximation de la réponse du modèle par une expression analytique. On considère pour cela le modèle \mathcal{M} comme une boîte noire. La régression peut être effectuée de deux manières différentes.

a) Méthode de collocation

Cette méthode consiste à construire les polynômes en réalisant une approximation de la réponse du modèle sur des points appelés points de collocation. Ainsi, pour une variable d'entrée donnée X_i , des points de collocation $\{x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(n)}\}$ sont choisis de façon optimale au sein de sa plage de variation. Une interpolation de la réponse du modèle Y est ensuite faite sur ces points pour en obtenir une approximation.

Cependant, en réalisant une approximation du modèle aux points de collocation, la solution obtenue peut s'avérer différente du modèle initial (figure 3.3). Aussi le nombre de combinaisons des points de collocation augmentant très rapidement et étant déterminé par la taille du polynôme, sa précision en est altérée. En augmentant ainsi le nombre de points de collocation, on se retrouve à réaliser une régression.

b) Méthode de régression

Cette deuxième approche est basée classiquement sur la minimisation de l'écart entre la réponse Y du méta-modèle et celle du modèle initial. On considère pour cela l'approximation \tilde{Y} de la réponse du modèle à un ordre p défini par :

$$\tilde{Y}(X) = \sum_{k=0}^p \beta_k \psi_k(X) \quad (3.16)$$

On prend l'exemple d'une variable aléatoire X gaussienne centrée réduite. On considère comme précédemment n points $\{x^{(1)}, x^{(2)}, \dots, x^{(n)}\}$ sur la plage de variation de la variable

²La troncature est utilisée pour limiter le développement d'une fonction à un certain ordre.

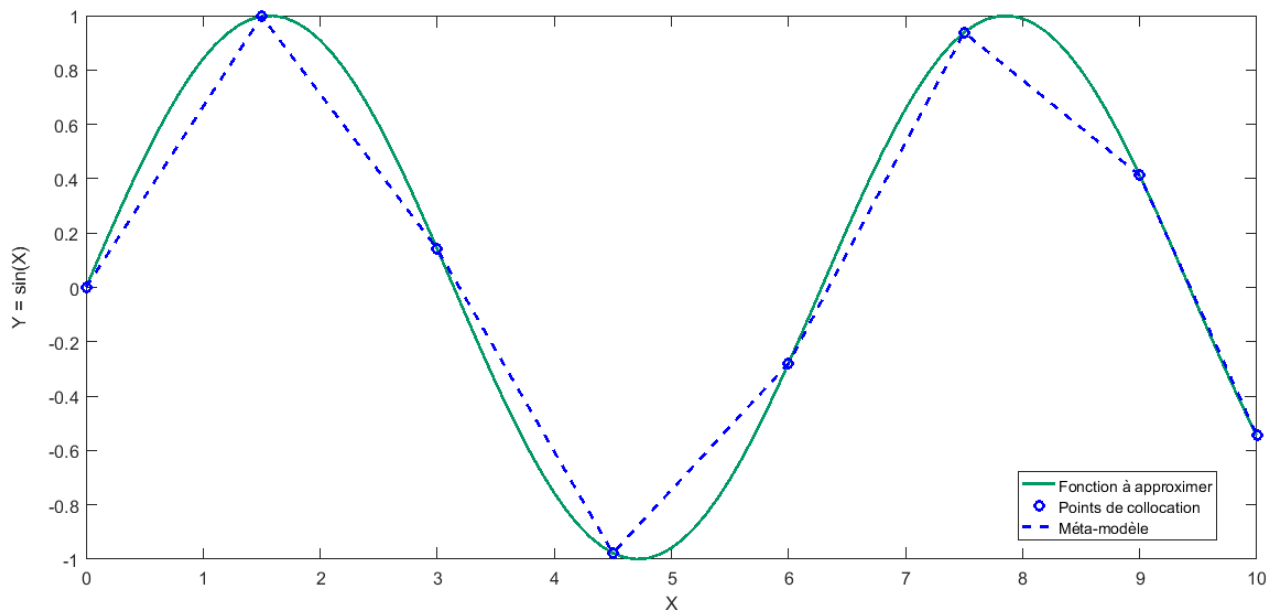


FIGURE 3.3 – Approximation par la méthode de collocation de la fonction $\sin(x)$. La fonction est bien approximée aux points de collocation mais le méta-modèle construit révèle une mauvaise approximation aux autres endroits de la courbe.

X.

En minimisant l'écart entre \tilde{Y} et Y , on obtient par la méthode des moindres carrés ΔY :

$$\Delta Y = \sum_{i=1}^n \left(Y^{(i)} - \tilde{Y}^{(i)} \right)^2 \quad (3.17)$$

Où $Y^{(i)}$ et $\tilde{Y}^{(i)}$ sont respectivement les réponses et les approximations du modèle au point $x^{(i)}$.

En remplaçant l'équation 3.16 dans l'équation précédente, on obtient :

$$\Delta Y = \sum_{i=1}^n \left(Y^{(i)} - \sum_{k=0}^p \beta_k \psi_k(x^{(i)}) \right)^2 \quad (3.18)$$

Pour estimer les coefficients β_k , la méthode de régression revient à résoudre l'équation 3.18 et à minimiser par rapport aux coefficients β_k l'expression de ΔY .

Le minimum est obtenu pour les valeurs de β_k vérifiant :

$$\frac{\delta \Delta Y}{\delta \beta_l} = 0 \quad \forall l \in \{0, \dots, p\} \quad (3.19)$$

Ce qui revient à résoudre au final l'équation matricielle de taille $p \times p$:

$$\begin{pmatrix} \sum_{i=1}^n \psi_0(x^{(i)})\psi_0(x^{(i)}) & \cdots & \sum_{i=1}^n \psi_0(x^{(i)})\psi_p(x^{(i)}) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n \psi_p(x^{(i)})\psi_0(x^{(i)}) & \cdots & \sum_{i=1}^n \psi_p(x^{(i)})\psi_p(x^{(i)}) \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_p \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n Y^{(i)}\psi_0(x^{(i)}) \\ \vdots \\ \sum_{i=1}^n Y^{(i)}\psi_p(x^{(i)}) \end{pmatrix} \quad (3.20)$$

Les valeurs des coefficients $\{\beta_0, \beta_1, \dots, \beta_p\}$ s'obtiennent en résolvant l'équation matricielle 3.20. Ces coefficients sont donc obtenus, en fonction du degré de polynômes choisi p , au sens des moindres carrés.

Les points servant à la construction du méta-modèle n'étant pas totalement indépendants les uns des autres, il est important d'utiliser des algorithmes adaptés pour la résolution du système matriciel (3.20). Ce problème est généralement résolu en utilisant des méthodologies robustes telles que l'algorithme Décomposition en Valeurs Singulières (SVD) (Björck, 1996). Elle permet la décomposition en une matrice orthogonale, une matrice diagonale qui contient les valeurs singulières ainsi que la transposée d'une matrice diagonale.

3.3.3 Construction d'un plan d'expériences

Le choix du plan d'expériences est important car il permet de définir la combinaison des points lors de la construction du méta-modèle (figure 3.7). La principale difficulté réside dans le choix de la méthode d'échantillonnage de chaque variable prise indépendamment pour la construction du plan d'expériences. Plusieurs méthodes permettent de générer les points sur lesquels le méta-modèle sera construit. Cependant, pour une bonne couverture de l'espace de chaque variable, il convient de choisir la méthode d'échantillonnage la mieux adaptée. Pour sélectionner les points qui seront évalués au sein de la plage de variation de chaque variable, on réalise un échantillonnage. Que ce soit pour réaliser une analyse de sensibilité ou d'incertitudes, on se base sur les valeurs prises par la variation des grandeurs d'entrée considérées pour en évaluer l'effet sur la sortie du modèle. Le résultat obtenu dépend donc fortement des points choisis lors de l'échantillonnage.

Une exploration cohérente de l'espace de chaque grandeur doit être faite afin d'obtenir des résultats fiables. Les points d'évaluation doivent être bien répartis au sein de l'espace de la grandeur considérée pour une bonne représentativité. Ainsi, deux propriétés essentielles sont à prendre en compte lors d'un échantillonnage. Les grandeurs doivent être échantillonnées indépendamment les unes des autres pour éviter une corrélation des

entrées étudiées (l'effet de chaque grandeur étant évalué indépendamment à l'entrée du modèle). Les points échantillonnés doivent aussi garder en eux les informations statistiques transmises par les lois de distributions et leurs caractéristiques avant l'échantillonnage.

Plusieurs méthodes d'échantillonnage ont été développées dans ce sens (Saltelli et al., 2008). Elles se basent sur une génération aléatoire ou quasi-aléatoire des points d'évaluation. Les méthodes d'échantillonnage de type aléatoire permettent de générer des points par un processus complètement lié au hasard. À l'opposé, les méthodes d'échantillonnage quasi-aléatoires permettent de générer des suites de points aléatoires dans un processus parfaitement contrôlé de manière à obtenir une meilleure représentativité. Les points échantillonnés sont répartis le plus uniformément possible au sein de la plage de variation : d'où leur dénomination de méthodes quasi-aléatoires.

Dans la suite, nous comparons deux méthodes d'échantillonnage aléatoires (Le Latin Hypercube Sampling et la méthode de Monte Carlo classique) ainsi que deux méthodes d'échantillonnage quasi-aléatoires (Les suites de Halton et les suites de Sobol). Pour des informations supplémentaires sur les méthodes existantes, le lecteur pourra se référer à (Franco, 2008).

3.3.3.1 Le Monte Carlo Classique

La méthode d'échantillonnage selon Monte Carlo consiste à effectuer un nombre choisi n de tirages suivant la densité de probabilité et les caractéristiques de la grandeur choisie. n simulations sont ensuite réalisées sur ces échantillons afin d'obtenir la densité de probabilité de la grandeur de sortie. Aussi, plus n est élevé, plus le temps de calcul est long mais plus la densité de probabilité obtenue se rapproche de la réalité. Cette convergence est indépendante du nombre de grandeurs d'entrée ce qui est d'autant plus intéressant lorsque l'on est face à un modèle nécessitant un échantillonnage sur un nombre important de grandeurs.

3.3.3.2 Le Latin Hypercube

L'échantillonnage par Latin Hypercube a initialement été développé par McKay en 1979 (McKay and Beckman, 1979). Le but de cette méthode est de subdiviser les données en plusieurs groupes de manière à représenter au mieux la plage de variation de la grandeur et ceci en effectuant un nombre réduit de tirages. La plage de variation est ainsi subdivisée en q intervalles $\left\{ \left[0, \frac{1}{q}\right], \left[\frac{1}{q}, \frac{2}{q}\right], \dots, \left[\frac{q-1}{q}, 1\right] \right\}$ avec une même probabilité. Une valeur est ensuite tirée une fois aléatoirement au sein de chaque sous intervalle. Et même avec un nombre limité de tirages, les points échantillonnés couvriront toute la plage de variation. Un échantillonnage par Latin Hypercube permet d'obtenir une meilleure couverture de l'espace des grandeurs que la méthode de Monte Carlo classique. Cependant, la

bonne couverture de l'espace s'évalue par la mesure de la régularité de la distribution par discrédance (Franco, 2008). Elle permet de juger de la répartition homogène des points échantillonnés sur l'espace entier en estimant la proportion de points contenus dans un sous ensemble de q par rapport au nombre total de points échantillonnés : plus faible est la discrédance, meilleure est la répartition des points.

L'échantillonnage par Latin Hypercube présente néanmoins deux inconvénients non négligeables. Le premier est la diagonalisation des points échantillonnés. Ceux-ci peuvent être placés sur une même diagonale de l'espace à recouvrir ce qui entraîne des sous espaces vides et non représentés par l'échantillonnage (figure 3.4). De plus, lorsque l'on remarque cette discrédance, il n'est pas possible de réaliser des tirages complémentaires pour combler les espaces vides ; il faut refaire un échantillonnage supplémentaire. Il devient donc dans ce cas plus intéressant de recourir à des méthodes d'échantillonnage où le processus de tirage est maîtrisé comme dans le cas d'échantillonnages quasi-aléatoires. Elles sont faciles à implémenter et ont un temps d'exécution très faible. Les suites de Halton et de Sobol sont des méthodes d'échantillonnage où le processus de tirage est contrôlé.

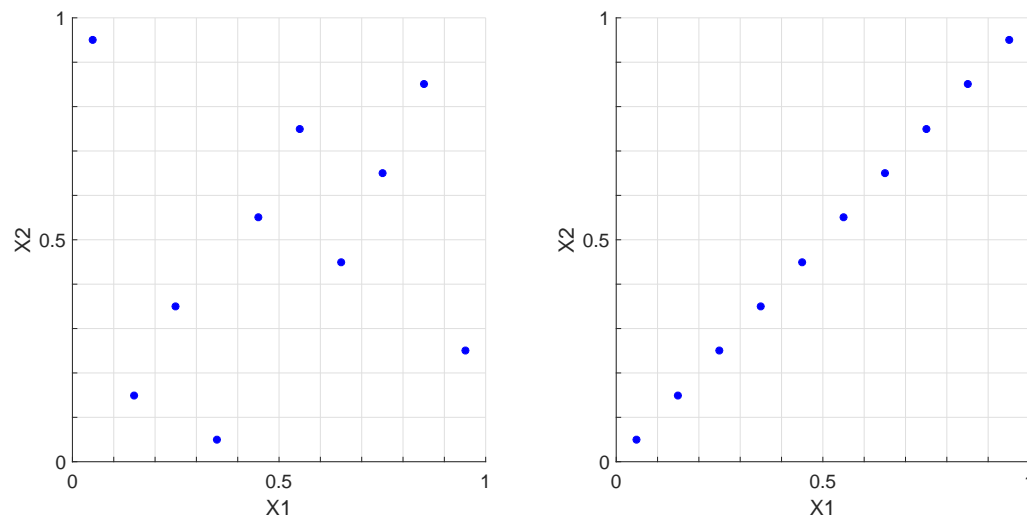


FIGURE 3.4 – 10 points échantillonnés par Hypercube Latin en dimension 2. On remarque sur la figure de gauche, que les points échantillonnés recouvrent bien l'espace des deux variables. À l'opposé sur la figure de droite, les points échantillonnés sont alignés et laissent des régions de l'espace vides donc non représentés.

3.3.3.3 Les suites de Halton

Les suites de Halton sont connues pour être des méthodes d'échantillonnage à faible discrédance. Elles constituent une généralisation à plusieurs dimensions des suites de Van der Corput³ (Franco, 2008).

³ Tout entier n dans une base b donnée, a une représentation en terme de cette base. La suite de Van der Corput en base b est donnée par :

Soit b la base de décomposition. La suite de Halton pour tout entier $b \in \{2, 3, 4, \dots\}$ en dimension p est définie par :

$$u_n = (\phi_{b_1}(n), \dots, \phi_{b_p}(n)) \quad \forall n \in \{0, 1, \dots\} \quad (3.21)$$

Où ϕ_{b_i} désigne la suite de Van der Corput en base b_i . Ainsi, en fixant la base b_i , on effectue une division entre les nombres premiers et la base choisie de manière à obtenir une suite de nombres compris entre 0 et 1.

On obtient une représentation plus homogène de l'espace des paramètres. Cependant en dimension élevée p , la répartition des points devient irrégulière. Les suites produites deviennent répétitives et engendrent ainsi des sous espaces vides comme représentés à la figure 3.5.

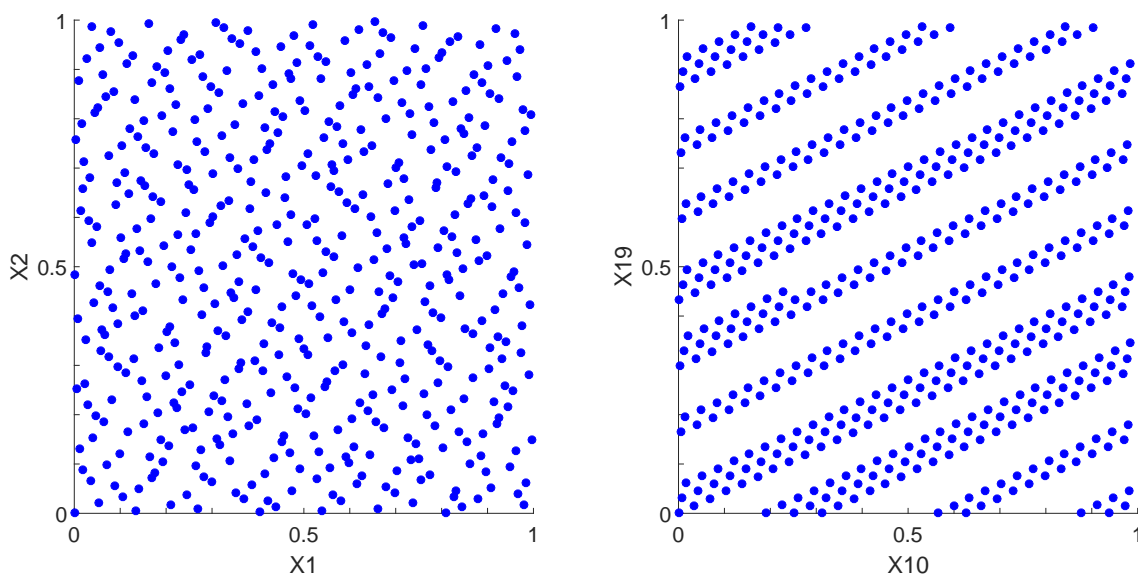


FIGURE 3.5 – Échantillonnage de 500 points avec les suites de Halton réalisé en dimension 20 sur les sous espaces de dimensions 1 et 2 à gauche et de dimension 10 et 19 à droite. On remarque que pour des dimensions élevées, des sous espaces restent vides et non représentés.

$$\phi_b(n) = \sum_{j=0}^{\infty} a_j \cdot b^j$$

3.3.3.4 Les suites de Sobol

Les suites de Sobol permettent de préserver l'uniformité de la distribution étudiée même en dimension élevée contrairement aux suites de Halton. Une suite de Sobol en dimension 1 et définie en base 2 est donnée par :

$$u_n = \frac{a_0}{2} + \frac{a_1}{2^2} + \dots + \frac{a_n}{2^{n+1}} \quad (3.22)$$

Où les a_j sont les coefficients obtenus de l'expansion en base 2.

On remarque toutefois sur la figure 3.6 qu'en dimension élevée, des sous-espaces vides ont tendance à se former. De plus, en augmentant le nombre de points échantillonnés, ces derniers ont tendance à se regrouper.

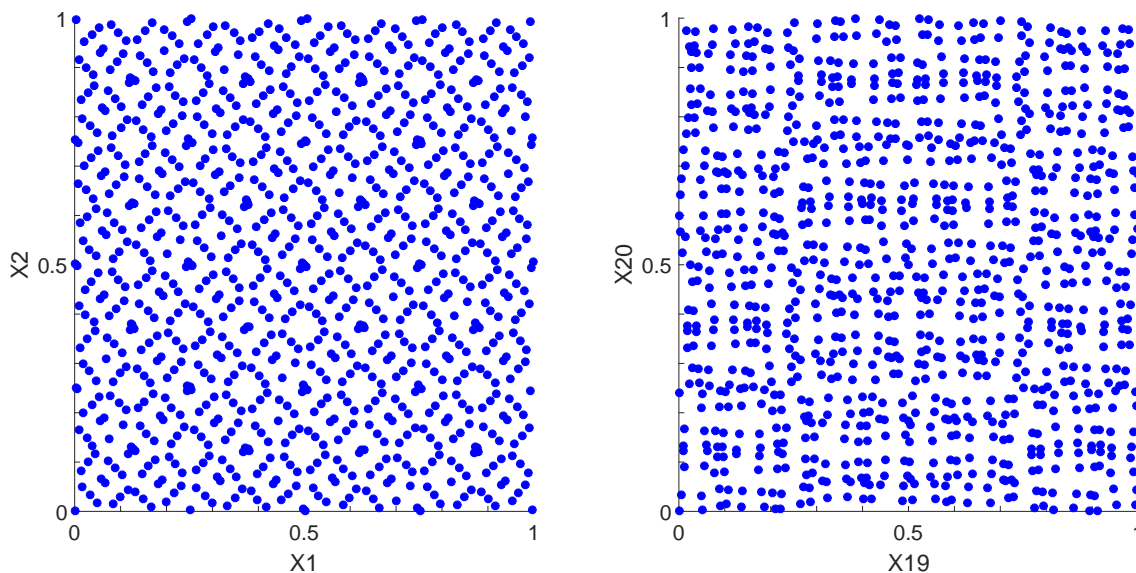


FIGURE 3.6 – Échantillonnage de 1000 points avec les suites de Sobol réalisé en dimension 20 sur les sous espaces de dimensions 1 et 2 à gauche et de dimensions 19 et 20 à droite.

3.3.3.5 Comparaison et choix de la méthode optimale

Toutes ces méthodes ont été abondamment utilisées dans le domaine du bâtiment et de la thermique. Que ce soit avec des méthodes d'échantillonnage aléatoire classique (Huang et al., 2015; Gang et al., 2015), échantillonnage par Hypercube Latin (de Wilde and Tian, 2012; Sun et al., 2014), méthodes d'échantillonnage quasi-aléatoires (Rabouille, 2014) pour diverses applications (rénovation, changement climatique, dimensionnement des systèmes CVC).

Saltelli et al. (Saltelli et al., 2008) ont proposé une comparaison de toutes les méthodes suivant les avantages et inconvénients de chacune d'elles selon les cas d'application et préconisent un choix judicieux de la méthode à employer en faisant un compromis entre le temps de calcul effectif et la précision des résultats. Cependant la méthode la plus simple et la plus couramment utilisée est le Monte Carlo classique mais elle reste néanmoins la méthode probabiliste qui nécessite le plus grand nombre de tirages pour recouvrir entièrement l'espace des paramètres. Ceci, contrairement aux méthodes d'échantillonnage quasi-aléatoire où la convergence est plus rapide.

Les résultats d'estimation sont étroitement liés à la méthode d'échantillonnage, il est donc important de choisir la méthode optimale de manière à obtenir de faibles disparités. Blatman et al. (Blatman et al., 2007) ont effectué une comparaison des méthodes évoquées et ont démontré que les méthodes d'échantillonnage quasi-aléatoires sont les plus performantes. Dans la suite, les suites de Sobol seront utilisées pour l'échantillonnage des variables aléatoires présélectionnées pour leur performance en terme de couverture de l'espace.

3.3.4 Post-traitement des coefficients

À partir du méta-modèle défini dans l'équation 3.13, on obtient après avoir estimé les coefficients $\hat{\beta}_k$ par la méthode de régression, l'expansion polynomiale de degré p du type suivant :

$$\hat{\mathcal{M}}(X) = \sum_{k=0}^p \hat{\beta}_k \psi_k(X) \quad (3.23)$$

On peut ensuite effectuer des post-traitements pour vérifier la cohérence du méta-modèle obtenu. Cette étape est importante et nécessite l'utilisation d'estimateurs performants. À partir de la densité de probabilité de la réponse polynomiale, on peut effectuer une analyse en tendance centrale par l'évaluation des moments statistiques et une analyse de fiabilité pour évaluer la cohérence des queues de distribution du modèle polynomial obtenu.

Analyse des moments statistiques

Les moments statistiques peuvent être obtenus en exploitant l'orthogonalité de la base et par une combinaison des coefficients $\hat{\beta}_k$. À partir de l'équation 3.23, la moyenne et la variance de $\hat{\mathcal{M}}$ sont données respectivement par :

$$\mu_{\hat{\mathcal{M}}} = \hat{\beta}_0 \quad (3.24)$$

$$\sigma_{\hat{\mathcal{M}}}^2 = \sum_{k=1}^p \hat{\beta}_k^2 \quad (3.25)$$

On peut aussi estimer les coefficients d'asymétrie $\delta_{\hat{\mathcal{M}}}$ et d'aplatissement $\kappa_{\hat{\mathcal{M}}}$ si l'on souhaite étudier l'étendu de la distribution de la variable de sortie notamment les queues de distribution.

Estimation de la densité de probabilité

Plusieurs méthodes permettent d'obtenir la densité de probabilité de la sortie du méta-modèle $\hat{\mathcal{M}}$. On peut soit réaliser des simulations de Monte Carlo, soit dériver la fonction de répartition et calculer la densité de probabilité de manière numérique, soit la déterminer en se basant sur les résultats des méthodes fiabilistes de type FORM. La méthode de Monte Carlo appliquée au méta-modèle, a été choisie car elle est la plus facile à implémenter. Elle permet d'effectuer à moindre coût (évaluation très rapide du méta-modèle), une comparaison qualitative avec la réponse du modèle initial

En effectuant un nombre M de simulations de Monte Carlo sur le méta-modèle $\hat{\mathcal{M}}$ construit, on obtient M réponses \hat{Y} .

$$\hat{Y}^{(j)} = \hat{\mathcal{M}}(X^{(j)}) \quad (3.26)$$

La probabilité de la réponse est estimée par :

$$\hat{f}(\hat{Y}) = \frac{1}{Nh} \sum_{j=1}^N K\left(\frac{Y - \hat{Y}^{(j)}}{h}\right) \quad (3.27)$$

Où K représente le noyau Kernel et h la largeur de la bande passante. Les réponses peuvent être représentée sous forme d'histogramme en regroupant les valeurs les plus proches les unes des autres en classes et la densité de probabilité de la réponse est estimée par la fonction $\hat{f}(\hat{Y})$.

Obtention des indices de sensibilité de Sobol

Considérons l'équation 3.13 de l'expansion polynomiale. Selon Sudret (Sudret, 2008), il est possible de réarranger les coefficients de l'équation en fonction des variables $X = \{X_1, X_2, \dots, X_N\}$ dont elles dépendent pour obtenir les indices de sensibilité de Sobol. Cette méthode d'analyse de sensibilité globale est deux à trois fois plus rapide que la méthode d'évaluation traditionnelle des indices de Sobol. Les indices de sensibilité sont obtenus analytiquement ce qui induit un temps de calcul quasiment nul contrairement au modèle physique initial qui aurait nécessité un nombre important de simulations de Monte

Carlo. En considérant les indices non nuls $\{i_1, \dots, i_N\}$, l'expansion polynomiale peut donc se réécrire :

$$\hat{\mathcal{M}}(X) = \beta_0 + \sum_{i=1}^N \sum_{a \in \mathcal{I}_i} \beta_a \psi_a(X_i) + \sum_{1 \leq i_1 \leq i_2 \leq N} \sum_{a \in \mathcal{I}_{i_1, i_2}} \beta_a \psi_a(X_{i_1}, X_{i_2}) + \dots + \sum_{a \in \mathcal{I}_{1, \dots, N}} \beta_a \psi_a(X) \quad (3.28)$$

Une fois que le polynome a été réécrit sous la forme de l'équation 3.28. Les indices de sensibilité d'ordre 1 à N , S_{i_1, \dots, i_N} peuvent être déduits ainsi que les indices totaux S_i^T (équations 3.29 et 3.30).

$$S_{i_1, \dots, i_N} = \frac{1}{\text{Var}[\hat{\mathcal{M}}(X)]} \sum_{a \in \mathcal{I}_{i_1, \dots, i_N}} \beta_a^2 \quad (3.29)$$

$$S_i^T = \frac{1}{\text{Var}[\hat{\mathcal{M}}(X)]} \sum_{a \in \mathcal{J}} \beta_a^2 \quad (3.30)$$

Où $\mathcal{I}_{i_1, \dots, i_N}$ et \mathcal{J} représentent l'ensemble des indices non nuls tels que :

$$\mathcal{I}_{i_1, \dots, i_N} = \{a \in \mathbb{N} : 0 \leq |a| \leq p, \forall k \in \{1, \dots, N\} \setminus \{i_1, \dots, i_N\}, a_k = 0\} \quad (3.31)$$

$$\mathcal{J} = \{a \in \mathbb{N} : 0 \leq |a| \leq p, a_i \neq 0\} \quad (3.32)$$

Ainsi, chaque indice de sensibilité fournit une information sur la part de la variabilité de la sortie du modèle qui est due à l'incertitude de chaque variable d'entrée $\{X_1, X_2, \dots, X_N\}$. Les indices de premier ordre S_i donnent l'influence de chaque variable prise individuellement tandis que ceux d'ordres supérieurs S_{i_j} indiquent la sensibilité des variables qui dépendent les unes des autres. Les indices de sensibilité totaux S_i^T quand à eux évaluent l'effet total de chaque variable, interactions comprises.

3.3.5 Estimateurs d'erreur et convergence

3.3.5.1 Erreur relative

Une fois que le modèle polynomial est construit, il reste à le valider par comparaison avec le modèle physique initial. Cette validation passe par l'estimation d'indicateurs d'erreur performants. La première approche consiste à évaluer la distance moyenne entre le modèle et son approximation polynomiale. Cette distance est ensuite adimensionnée par la valeur

du modèle initial pour donner l'erreur relative $E^{(i)}$.

On considère un plan d'expériences $X = \{X^{(1)}, X^{(2)}, \dots, X^{(n)}\}$ de n réalisations des entrées incertaines du modèle physique et $\mathcal{M}(X^{(i)})$, respectivement $\hat{\mathcal{M}}(X^{(i)})$, les réponses du modèle, respectivement du méta-modèle, correspondantes aux n évaluations.

On définit dans un premier temps une erreur relative $E^{(i)}$:

$$E^{(i)} = \left| \frac{\mathcal{M}(X^{(i)}) - \hat{\mathcal{M}}(X^{(i)})}{\mathcal{M}(X^{(i)})} \right| \quad (3.33)$$

L'erreur relative $E^{(i)}$ est évaluée à tous les points auxquelles ont été construit le méta-modèle. On définit ensuite une erreur globale E de construction qui correspond à l'erreur maximale sur tous les points du plan d'expérience (équation 3.34).

$$E = \max(E^{(i)}) \quad (3.34)$$

Il peut cependant arriver d'obtenir dans certaines situations une erreur relative E très faible alors que l'approximation polynomiale est en réalité mauvaise. Ceci s'explique par le fait que l'estimateur d'erreur n'est évalué qu'aux points du plan d'expérience qui ont servi à construire le méta-modèle tels que représentés dans la figure 3.3. Un estimateur d'erreur supplémentaire est alors nécessaire.

3.3.5.2 Leave One Out error

Pour évaluer l'approximation du méta-modèle sur des points qui ne font pas partie du plan d'expérience initial, on effectue une validation croisée. Le processus de validation croisée consiste à considérer deux plans d'expérience complètement indépendants, à construire le méta-modèle à partir du premier plan d'expérience et à ensuite valider l'approximation polynomiale obtenue sur le second. Cependant, la construction du méta-modèle nécessitant déjà un grand nombre d'appels au modèle, il devient contraignant d'effectuer des simulations supplémentaires, ce qui augmenterait considérablement le temps de calcul.

On peut alors avoir recours à la méthode du *Leave One Out error* (Blatman, 2009). Elle consiste à construire le méta-modèle sur les $n - 1$ points du plan d'expériences construit et à l'évaluer sur le point restant. L'erreur est ensuite évaluée en estimant l'écart entre la réponse du modèle obtenue avec le méta-modèle construit $\mathcal{M}(X^{(i-1)})$ et celle obtenue sur le point de validation $\mathcal{M}(X^{(i)})$ (équation 3.35).

$$\Delta^{(i)} = \left| \frac{\mathcal{M}(X^{(i)}) - \hat{\mathcal{M}}(X^{(i-1)})}{\mathcal{M}(X^{(i)})} \right| \quad (3.35)$$

Aucune simulation supplémentaire n'est alors nécessaire et l'erreur sur l'ensemble du plan d'expérience est obtenue grâce à :

$$E_{LOO} = \max(\Delta^{(i)}) \quad (3.36)$$

L'erreur E_{LOO} (équation 3.36) est souvent connue sous le nom de PRESS (Predicted Residual Sum of Squares). Elle permet d'évaluer la consistance du méta-modèle en se basant sur la plan d'expérience initial sans pour autant effectuer des simulations supplémentaires. Plusieurs auteurs ont aussi proposé d'améliorer les deux indicateurs d'erreurs E et E_{LOO} pour réduire leur sensibilité et améliorer leurs estimations (Blatman, 2009).

3.3.5.3 Comparaisons quantitative et qualitative

On utilise les estimateurs d'erreurs présentés dans la section précédente pour évaluer quantitativement l'approximation polynomiale réalisée. Une estimation de l'indicateur d'erreur choisi est effectuée sur la réponse du modèle et on espère obtenir une convergence en un nombre M de simulations donné. On peut aussi évaluer la qualité du méta-modèle en comparant la réponse du modèle initial et celle de l'expansion polynomiale par scatterplot. On localise ainsi les points aberrants et on vérifie l'alignement par rapport à la première bissectrice de la représentation faite. Il est aussi possible de comparer les distributions obtenues. On pourra ainsi juger du risque de sous ou de sur-évaluation de la sortie choisie.

Pour conclure, la méthode du chaos polynomial permet de construire un méta-modèle au comportement semblable au modèle physique initial, en tout point, mais plus léger en temps de calcul sur un espace défini des variables d'entrée. Le temps de simulation du méta-modèle dépendra de la complexité du modèle initial (modélisation des systèmes, zonage, ...). Il permet à l'utilisateur d'étudier le comportement des sorties en fonction des différentes sollicitations imposées en simplifiant l'approche de simulation. Il permet aussi d'obtenir les indices de sensibilité globaux sans effectuer des simulations supplémentaires contrairement au modèle initial. Les interactions entre les variables peuvent aussi être étudiées à moindre coût de simulation.

3.3.6 Algorithme général de construction d'un méta modèle

En résumé, pour construire un méta-modèle sur les polynômes du chaos par la méthode de régression, on se base sur la minimisation de l'écart entre la réponse exacte du modèle et celle estimée et ceci pour un ensemble de points d'un plan d'expériences réalisé en amont

sur la plage de variation des variables d'entrée du modèle. Dans le plan d'expériences sont représentées toutes les combinaisons possibles de chaque variable. On choisit ensuite la taille du plan d'expériences souhaitée et en utilisant une transformation isoprobabiliste, on détermine les valeurs des variables d'entrée sur lesquelles le modèle sera évalué. On construit ainsi l'équation matricielle (3.20) à résoudre pour déterminer les coefficients β et l'expansion polynomiale. Les erreurs relatives et en validation croisée sont ensuite déterminées et si le seuil d'erreur fixé est dépassé, on passe au degré p suivant pour obtenir une meilleure précision (figure 3.7).

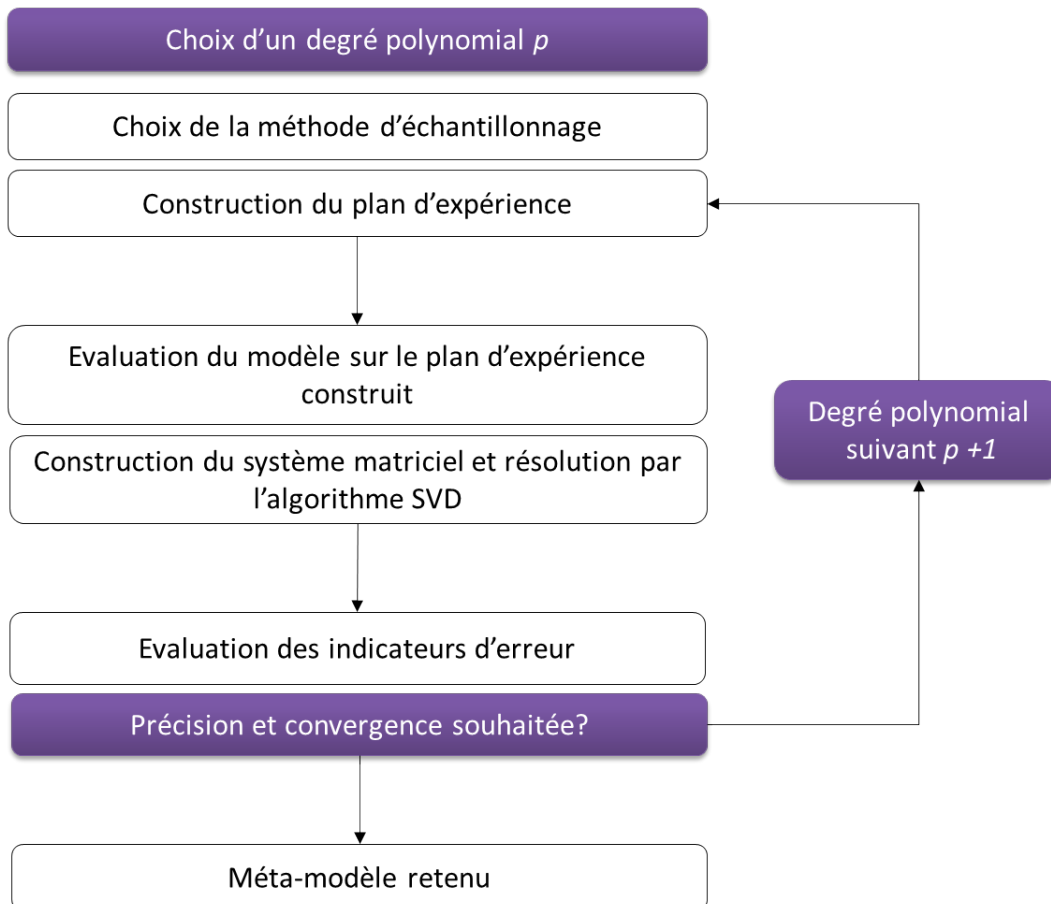


FIGURE 3.7 – Algorithme proposé de construction d'un méta-modèle par polynômes du chaos

3.4 Modélisation du bâtiment test sous Trnsys

3.4.1 Construction du modèle

Le bâtiment du CECP du Cerema qui a servi de cas d'étude pour les travaux a été modélisé sous Trnsys V17 (Transient System Simulation Tool version 17), le logiciel de Simulation Thermique Dynamique (STD) qui a été jugé approprié pour les recherches menées, du fait de la possibilité de prise en compte des systèmes. En règle générale, la STD permet

d'estimer les consommations d'énergie du bâtiment modélisé et le confort des occupants (Crawley et al., 2008), en tenant compte de la structure du bâti, des systèmes énergétiques (y compris les appareils électriques non thermiques) sous différentes sollicitations imposées (comportement des occupants, variations climatiques).

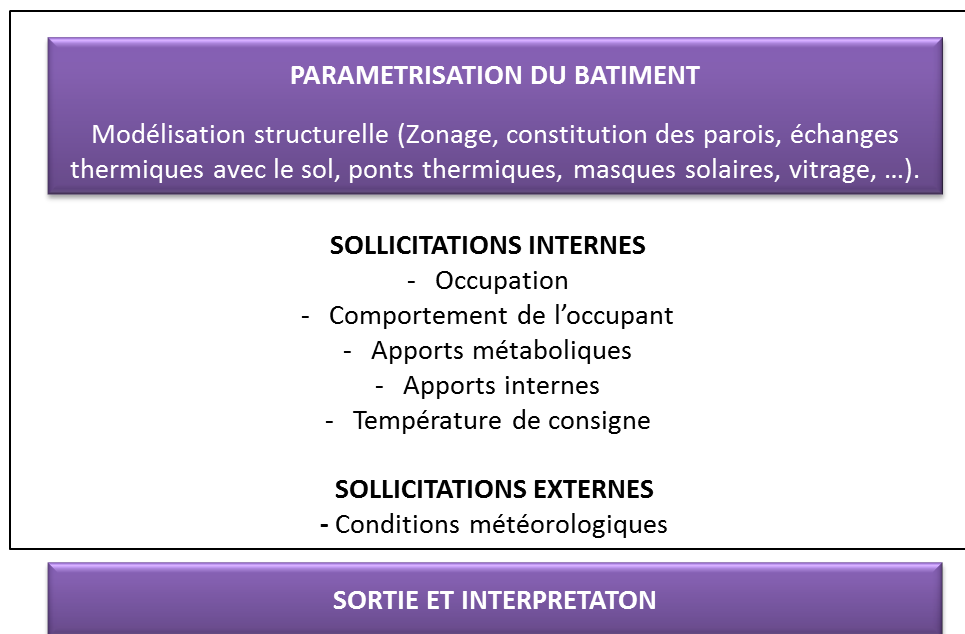


FIGURE 3.8 – Décomposition schématique d'un modèle de simulation thermique dynamique

Le bâtiment est modélisé en zones multiples. Les échanges inter-zones ainsi que les échanges radiatifs, convectifs et conductifs sont pris en compte. Puis, Trnsys considère chaque zone thermique comme un nœud d'air et modélise chaque paroi par une résistance thermique et une masse (Figure 3.9). Le système multi-zone est ensuite résolu par la méthode des fonctions de transfert en effectuant un équilibre entre les différents postes de dépenses et d'apports énergétiques (les infiltrations, les déperditions à travers l'enveloppe, la ventilation, les apports solaires et internes, et l'inertie). Ce bilan est effectué à chaque pas de temps de simulation pour estimer la quantité d'énergie à apporter pour maintenir une température de consigne et ainsi maintenir le système en équilibre.

3.4.1.1 Hypothèses de modélisation

La complexité des simulations nécessite de faire des choix de modélisation. Ces choix se formalisent sous la forme d'hypothèses à chaque poste de dépense ou d'apport énergétique. Ainsi, depuis le choix du zonage du bâtiment à la prise en compte des systèmes énergétiques, l'utilisateur devra faire des choix justifiés et fiables car le modèle en lui-même peut être une source d'incertitude supplémentaire. Comme l'a souligné M. Trčka (Trčka and Hensen, 2010) dans ses travaux, à partir d'un certain niveau de détail, le modèle de

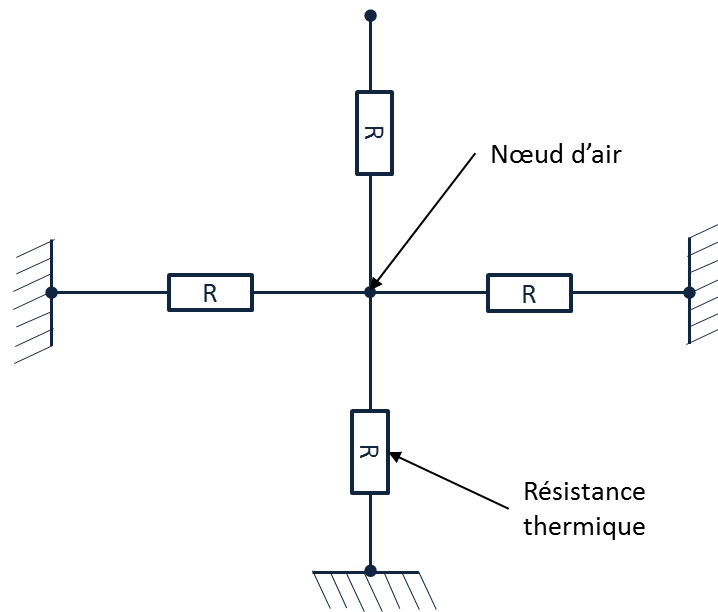


FIGURE 3.9 – Illustration de la modélisation thermique d'une zone à trois surfaces

simulation intègre plus d'incertitudes qu'un modèle simplifié.

3.4.1.2 Zonage du bâtiment

Pour modéliser un bâtiment sous Trnsys, sa structure physique doit être tout d'abord reproduite sous un logiciel de DAO (figure 3.10). Les informations contenant sa géométrie et sa configuration sont ensuite exportées sous la forme d'un fichier .idf. Pour construire ce fichier idf., le bâtiment a été modélisé en attribuant une zone à chaque local. Le bâtiment test a une surface d'environ 1500 m^2 , la surface de la partie atelier comprise. Quarante zones⁴ ont été saisies sous Sketchup 3D (Annexes figures 6.2 et 6.3). Les ouvertures extérieures (portes, fenêtres et baies vitrées) ont aussi été représentées.

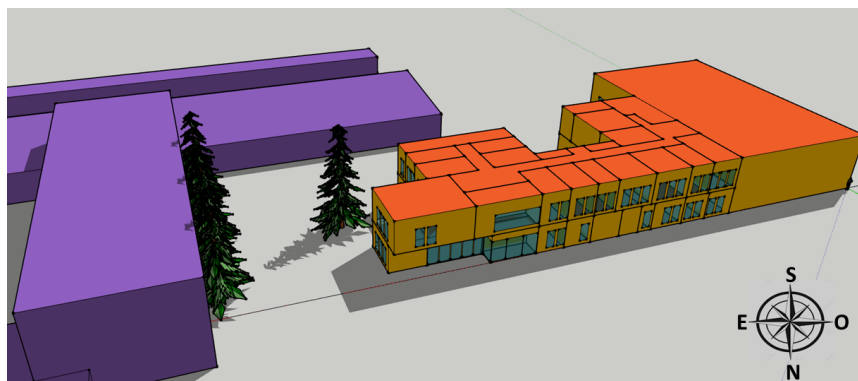


FIGURE 3.10 – Modélisation du bâtiment du CECP sous Sketchup 3D ainsi que des masques lointains.

⁴Une zone thermique en STD représente un volume au sein duquel on considère que tous les volumes élémentaires ont un comportement thermique identique.

Par ailleurs, le bâtiment n'est pas complètement orienté plein nord (décalage de 5° par rapport au nord). Aussi, pour pouvoir représenter les masques solaires, proches comme lointains, une orientation a été créée pour chaque façade du bâtiment (Annexes figure 6.1). Ainsi, comme masques solaires lointains, les deux bâtiments présents sur le site du Cerema, ont été modélisés ainsi que la végétation présente du côté est.

3.4.1.3 Échanges thermiques avec le sol

Les échanges thermiques avec le sol sont modélisés de manière simplifiée. Le sol et le plancher sont modélisés en un seul composant constitué de tous les matériaux du plancher bas plus 0.5 m de sol et une couche virtuelle. Elle est dimensionnée suivant la norme NF EN ISO 13370. La température de la couche virtuelle est égale à la température moyenne extérieure et est considérée constante toute l'année.

3.4.1.4 Constitution des parois et ponts thermiques

Les caractéristiques des parois sont présentées à la table 3.2. Un pont thermique est une rupture de la continuité dans l'enveloppe du bâtiment qui se manifeste par une baisse de la conductivité thermique du mur ou du plancher (figure 3.11). Elles représentent une des principales causes de déperditions. Six types de ponts thermiques ont été déclinés et représentés dans chaque zone (table 3.3). Les coefficients de transmission thermiques linéiques sont ceux utilisées lors de la conception du bâtiment.

TABLE 3.2 – Conductivité des parois du bâtiment CECP

N	Type	Coefficient de transfert thermique ($W/m^2 \cdot ^\circ K$)
1	Mur extérieur bureaux	0.165
2	Mur extérieur atelier	0.24
3	Mur adjacent	0.689
4	Plancher intermédiaire	4.406
5	Plancher sur terre plein	0.192
6	Plancher haut	0.15

3.4.1.5 Infiltrations

Le bâtiment test a été modélisé sous Contam pour la prise en compte des infiltrations et de la dynamique aéraulique. Contam est un logiciel d'analyse du déplacement de l'air dans un bâtiment constitué de plusieurs zones. Il permet de modéliser les échanges aérauliques au sein du bâtiment mais aussi entre ce dernier et l'extérieur.

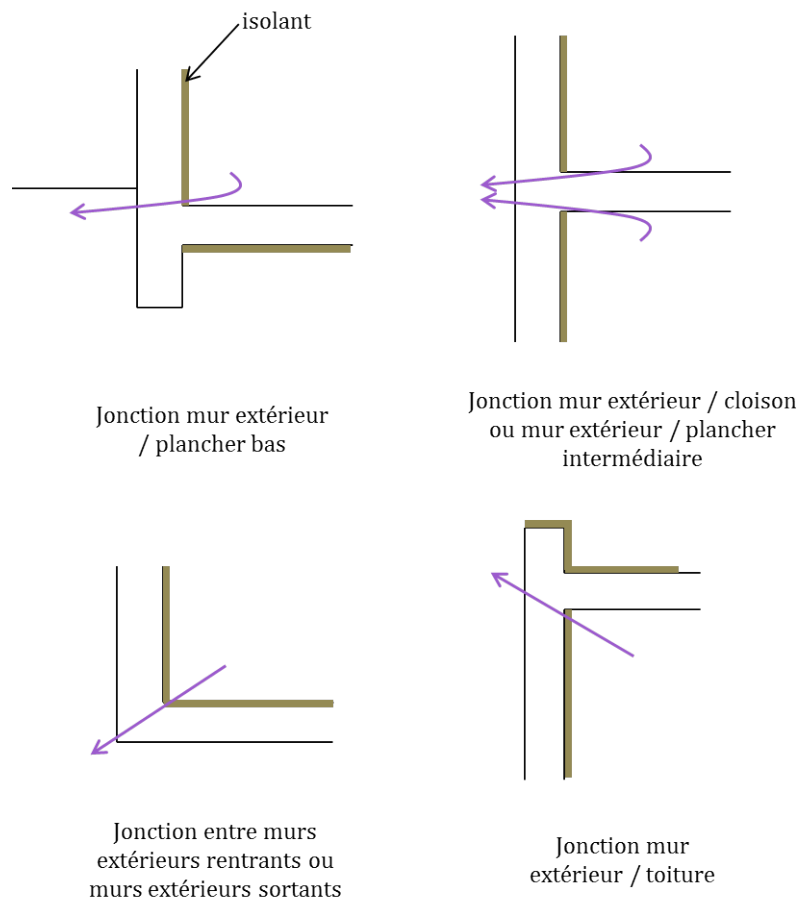


FIGURE 3.11 – Illustration des ponts thermiques pris en compte lors de la modélisation

TABLE 3.3 – Différents ponts thermiques représentés lors de la modélisation du bâtiment test

N	Type de pont thermique	Coefficient thermique linéique ($W/m \cdot ^\circ K$)
1	Jonction mur extérieur / cloison	0.09
2	Jonction entre deux murs extérieurs rentrants	0.05
3	Jonction entre deux murs extérieurs sortants	0.09
4	Jonction mur extérieur / plancher intermédiaire	0.09
5	Jonction mur extérieur / toiture	0.06
6	Jonction terre-plain / plancher bas	0.39

1. **La perméabilité à l'air :** Une infiltration a été créée pour chaque ouverture extérieure existante (portes, fenêtres). Dans chaque zone, un coefficient de fuite d'air a été attribué à chaque surface déperdition. Le bâtiment a été soumis à un essai de perméabilité à l'air. L'essai a été effectué conformément à la norme NF EN 13829 : 2001 [AFNOR \(2001\)](#) sous une pression de 4 Pascals. La perméabilité à l'air du bâtiment mesurée est de $1.3 \text{ m}^3/h$ par m^2 de surface.

Le coefficient d'infiltration correspondant à chaque paroi a été calculé au prorata des surfaces des parois conformément à la surface totale du bâtiment (les zones n'ayant pas de contact avec l'extérieur n'ont donc pas de coefficient d'infiltration). Pour le calcul des fuites d'air, les surfaces totales des parois ont été considérées (avec les ouvertures). Des infiltrations ont aussi été créées sur le toit du bâtiment.

Pour chaque paroi déperditive, un coefficient de fuite d'air C_i sous 4 Pa est calculé pour chaque débit d'infiltration correspondant Q_i suivant l'équation 3.37.

$$C_i = Q_i / (dP)^n \quad (3.37)$$

Où,

C_i est le coefficient de fuite d'air pour chaque paroi de la zone i considérée ($m^3/h.Pa$),
 Q_i est le débit d'infiltration d'air sous 4Pa pour chaque zone i considérée (m^3/h),
 dP est le différentiel de pression appliquée (Pa),
 n est le coefficient de débit d'air pour un écoulement laminaire.

2. **Les systèmes de ventilation :** Sous Contam, les bouches d'insufflation et d'extraction ont aussi été modélisées, fidèlement à la construction du bâtiment, de manière à créer des interactions avec les systèmes de ventilation sous Trnsys et à surtout prendre en compte leur effet sur les infiltrations.
3. **Les ouvrants :** Toutes les portes et fenêtres extérieures ont été supposées fermées, contrairement à toutes les portes communicantes à l'intérieur du bâtiment.

Le fichier *.air* généré par Contam est ensuite chargé sous Trnsys.

3.4.1.6 Les systèmes de ventilation

Pour assurer la ventilation du bâtiment, deux centrales de traitement d'air (CTA) sont installées ainsi qu'un système de ventilation mécanique contrôlée (VMC) simple flux. Elles assurent un renouvellement d'air de $1915 m^3/h$ dans tout le bâtiment avec des débits d'insufflation et d'extraction variables, suivant les différentes zones (Annexes table 6.6). Ces systèmes ont été modélisés sous Trnsys. Une ventilation double flux pour les zones à utilisation de bureaux, la deuxième CTA pour la salle de réunion et la ventilation mécanique pour les zones à utilisation dite sanitaires (Annexes figures 6.2 et 6.3).

Pour modéliser les débits réellement insufflés et extraits dans chaque zone sous Trnsys3D, un premier coefficient α_1 , a été affecté au débit total propre à chaque système de

ventilation, pour représenter le ratio des débits individuels, en fonction des débits de sortie des systèmes de ventilation et un deuxième coefficient α_2 pour représenter les fuites dans le réseau de ventilation suivant chaque zone. Pour estimer α_2 , des mesures de débit d'air ont été effectuées à chaque bouche d'insufflation et d'extraction dans chaque zone i considérée.

Le débit réel de ventilation en insufflation ou en extraction dans une zone i considérée est donnée par l'équation :

$$q_i = \alpha_1 * \alpha_2 * q \quad (3.38)$$

Où,

q_i est le débit d'insufflation ou d'extraction dans la zone i considérée en m^3/h ,

α_1 est le ratio de débit par rapport au débit total,

α_2 est le coefficient de fuite dans le réseau de ventilation,

q est le débit total d'insufflation correspondant à un système de ventilation donné en m^3/h .

3.4.1.7 Les masques solaires

Un masque solaire représente tout objet qui aurait pour effet de réduire les apports solaires apportés au bâtiment. On distingue deux types de masques solaires : les masques solaires lointains qui peuvent être des bâtiments voisins, des arbres ou des reliefs et les masques solaires proches qui sont le plus souvent causés par les rebords des éléments de toiture, des fenêtres ou des portes au-dessus des vitrages.

Pour prendre en compte les masques solaires dans le modèle de simulation, un relevé de hauteurs angulaires a été effectué suivant chaque azimut (figure 3.12) et ceci pour chaque orientation suivant l'équation 3.39. Les hauteurs angulaires ont été mesurées sous Sketchup. Ainsi, pour une orientation donnée, les hauteurs angulaires ont été mesurées à chaque azimut et à un pas de 10° de -180° à 180° .

$$\beta = \text{arctg} \left(\frac{h}{l} \right) \quad (3.39)$$

3.4.1.8 Initialisation de la simulation

La température de chaque zone a été initialisée à 20°C et l'humidité relative à 50 %. Des consignes ont ensuite été assignées à chaque zone de manière à décrire les phénomènes physiques mais aussi les systèmes thermiques dans le bâtiment : le chauffage, la ventilation, les infiltrations ainsi que les apports internes liés à l'occupation et aux différents appareils électriques.

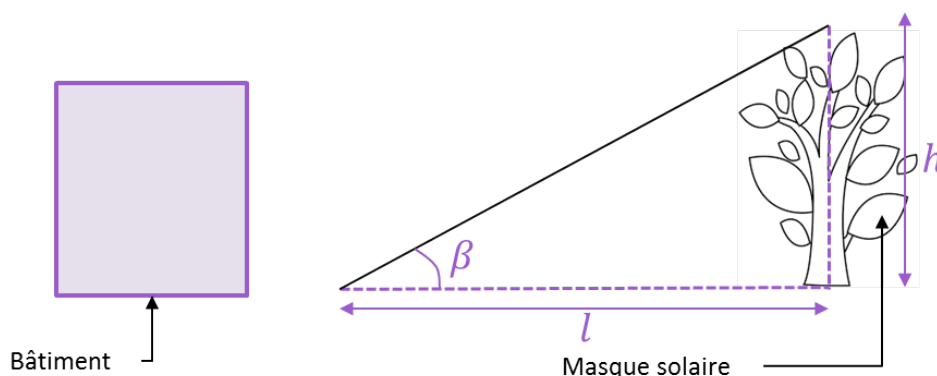


FIGURE 3.12 – Illustration de l'estimation des hauteurs angulaires d'un masque solaire lointain donné suivant un intervalle de 10 degrés, à chaque azimut

3.4.2 Comparaison Calcul / Mesures

Cette étape consiste à soumettre le modèle construit aux conditions réelles d'utilisation durant une période donnée et à comparer les résultats obtenus à la consommation calorifique effectivement constatée durant cette même période. Elle est en quelque sorte une étape de réglage et permet de s'assurer de la cohérence du modèle. Pour ce faire, les données de température ambiante mesurées dans chaque zone ont été mises en entrée du modèle ainsi que les apports internes effectifs mesurés durant cette même période.

L'hypothèse selon laquelle toute l'énergie consommée par les appareils électriques est entièrement restituée dans le bâtiment et respectivement dans les zones concernées à un instant donné a été émise. Ainsi, les consommations électriques mesurées ont été considérées comme apports internes pour la comparaison. En ce qui concerne les autres variables, les hypothèses sont regroupées dans la table 3.4. On constate à la table 3.5 que la consommation estimée par le modèle de simulation est légèrement plus élevée que la consommation réelle ($\approx 2\%$).

TABLE 3.4 – Hypothèses de calibration en entrée du modèle

Variabes d'entrée	Hypothèses de calibration
Température de consigne	Température ambiante
Apports internes	Consommation électrique
Données météorologiques	Météo mesurée sur site
Débits de ventilation	Débits mesurés
Rendement des CTA	Estimé par les mesures de température
Perméabilité à l'air	Essai de perméabilité
Ouverture des fenêtres	Fenêtres fermées
Ouverture des portes donnant sur l'extérieur	Portes fermées
Ouverture des portes donnant vers l'intérieur	Portes ouvertes

TABLE 3.5 – Consommation calorifique annuelle réelle et estimée par le modèle de simulation

Consommation calorifique annuelle (MWh)	
Réelle	36.67
Estimée	37.47

3.4.2.1 Signature énergétique

La signature énergétique est la droite de régression obtenue pour représenter la différence de température journalière entre l'intérieur et l'extérieur du bâtiment en fonction de la consommation énergétique correspondante. Plus la pente de la droite est élevée, plus le bâtiment est sensible aux rigueurs climatiques. On effectue généralement une régression linéaire du premier ordre mais pour obtenir plus de précision et une meilleure interprétation des résultats, une régression du second ordre peut aussi être réalisée.

La signature énergétique des résultats obtenus via la STD est comparée à celle obtenue par les données réelles de température ambiante et de consommation mesurée. La comparaison des deux droites de régression a permis de caler au mieux le modèle STD et de pouvoir l'ajuster en relevant les éventuelles dysfonctionnements.

Les résultats sont représentés sur la figure 3.13, avec en abscisse, les degrés jour (la différence entre la température intérieure moyenne et la température extérieure) et en ordonnée la consommation calorifique journalière comme spécifié par l'équation 3.40 ci-dessous.

$$\sum_{h=1}^{24} Y_{calo} = f \left(\frac{\sum_{h=1}^{24} (T_{int} - T_{ext})}{24} \right) \quad (3.40)$$

Où Y_{calo} représente la consommation calorifique horaire.

La comparaison a été effectuée sur la base de données d'une année de simulations et de mesures (de mars 2015 à février 2016). La consommation calorifique totale se compose de la consommation calorifique des radiateurs et de celles des deux CTA.

On peut remarquer dans la figure 3.13 la présence d'un biais entre les consommations réelles et celles de la consommation obtenue par simulation. Ceci peut être dû à un paramètre statique qui a été mal fixé dans le modèle, que ce soit au niveau des matériaux constitutifs du bâti (conductivité d'un isolant, pont thermique manquant, ...) qu'au niveau de la modélisation de la structure physique (surface, orientation, ...).

On remarque notamment que pour des fortes variations de températures, le modèle a tendance à fournir des consommations de chauffage plus faibles que la réalité contraire-

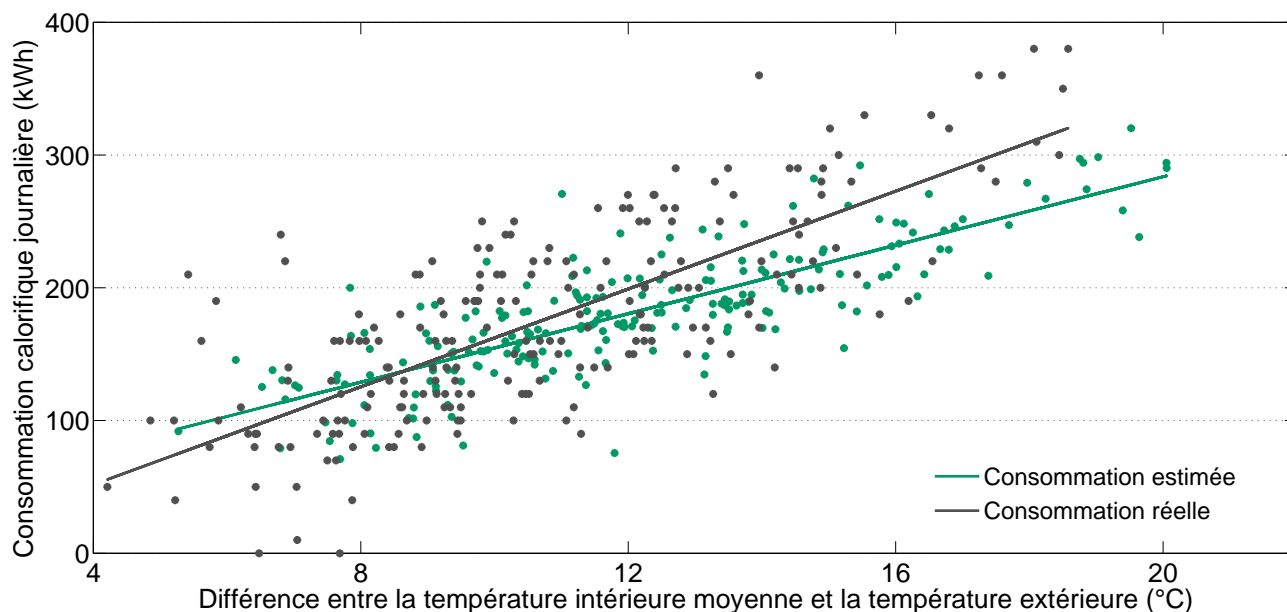


FIGURE 3.13 – Signatures énergétiques obtenues par régression linéaire des consommations journalières par rapport aux degrés jours

ment aux variations de température plus faibles. En comparant de plus près les résultats obtenus (au pas de temps horaire) par la figure 3.14, on constate que la consommation calorifique journalière a tendance à être surestimée pendant la période estivale et sous-estimée pendant la période de chauffe.

3.4.2.2 Indicateurs de performance d'un modèle

Plusieurs indicateurs permettent d'évaluer la performance d'un modèle et sont fournis par les standards internationaux tels que l'ASHRAE et l'IPMVP (ASHRAE, 2013; IPMVP, 2010). Parmi eux, les deux indicateurs les plus fréquemment utilisés sont :

- Le coefficient de variation de l'erreur quadratique moyenne CVRMSE (équation 3.41)
- L'erreur normalisée moyenne NMBE (équation 3.41)

$$CVRMSE(\%) = 100 \cdot \frac{\sqrt{\sum_{h=1}^N (Y_{mes} - Y_{sim})^2 / N}}{\bar{Y}_{mes}} \quad (3.41)$$

$$NMBE(\%) = 100 \cdot \frac{\sum_{h=1}^N (Y_{mes} - Y_{sim})}{\sum_{h=1}^N Y_{mes}} \quad (3.42)$$

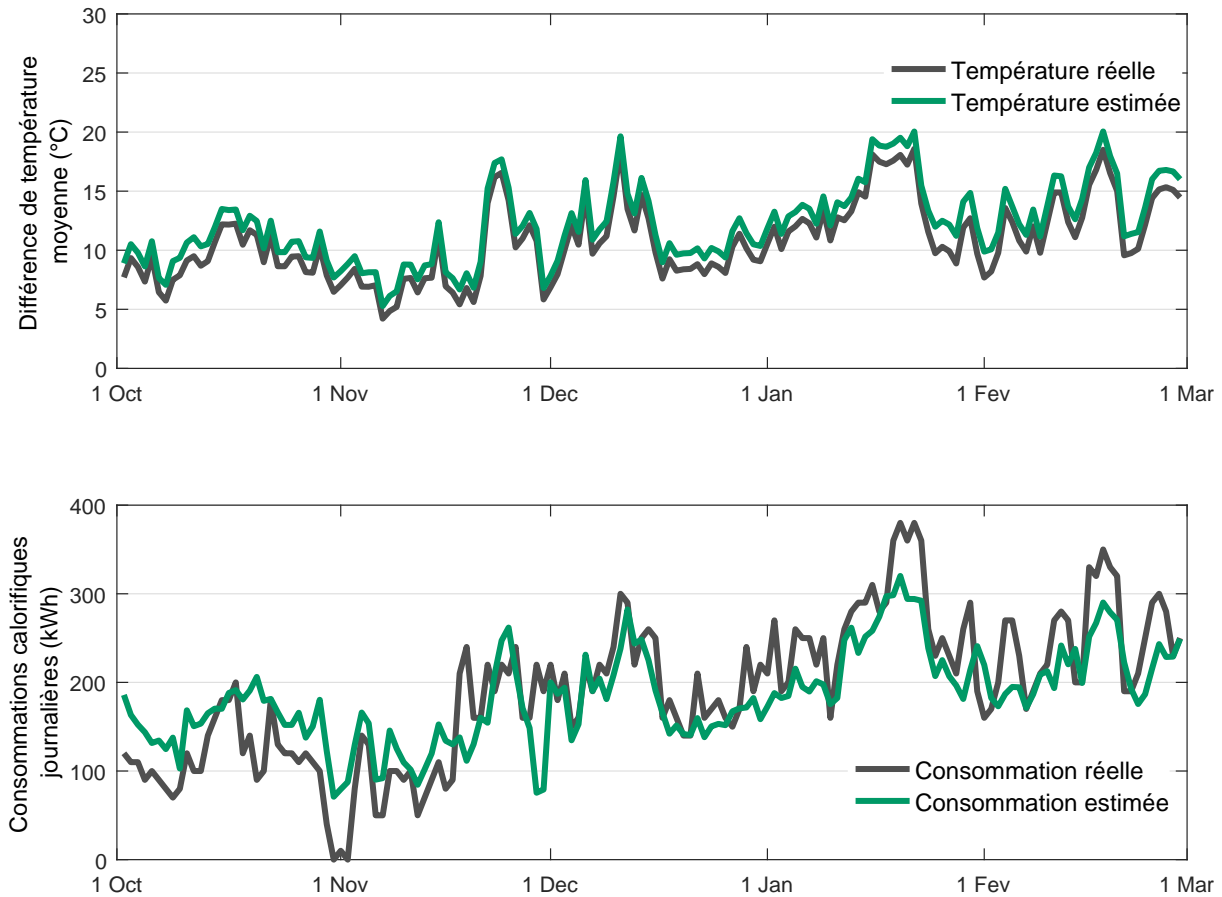


FIGURE 3.14 – Comparaison des températures intérieures mesurées et simulées ainsi que des consommations calorifiques mesurées et simulées journalières du 1 Octobre 2015 au 1 Mars 2016

TABLE 3.6 – Indicateurs de performance pour le calibrage d’un modèle de simulation selon un critère horaire et résultats du calibrage du modèle STD

Standards	NMBE (%)	CVRMSE (%)
ASHRAE 14	10	30
IPMVP	5	20
Modèle STD	-2.82	27.024

Le NMBE, comme indiqué dans l’équation 3.42, permet de juger de la proximité de la consommation calorifique totale estimée par le modèle par rapport à celle réellement perçue par le bâtiment sur une base horaire ou mensuelle. Elle donne ainsi une idée générale du biais qui peut être présent entre les deux. Elle est cependant affectée par l’effet de compensation sur la période considérée. Il est alors nécessaire d’utiliser un second indicateur. En déterminant le CVRMSE (équation 3.41), on peut déterminer la façon dont le modèle s’adapte aux données d’entrée. On obtient ainsi des informations sur les écarts plus ou moins importants entre les consommations réelles et celles estimées. Ces indicateurs fournissent une information sur la qualité des résultats obtenus (des consommations estimées) et non sur l’incertitude des variables en entrée du modèle ni sur le paramétrage du modèle en lui même.

Ces standards fournissent des critères d’acceptation d’un modèle pour juger de sa fiabilité en confrontation avec les valeurs effectives mesurées. Un modèle est donc considéré comme adéquat, si la valeur prise par les grandeurs d’intérêt est en accord avec les valeurs cibles fixées par l’ASHRAE et l’IPMVP. Au vu des résultats obtenus présentés à la table 3.6, le modèle STD construit peut être considéré comme cohérent.

3.5 Analyse de sensibilité et réduction de modèle sur le bâtiment test

3.5.1 Algorithme de simulation pour le couplage physico-probabiliste

Les analyses de sensibilité et de propagation d’incertitudes nécessitant un nombre important de calculs, il a fallu les optimiser pour réduire le temps de calcul. Nous disposons d’un ordinateur multi-cœur possédant 20 cœurs physiques pouvant fonctionner simultanément. Cet ordinateur étant uniquement dédié aux calculs, 20 calculs pouvaient être effectués au même moment. Nous avons pour cela implémenté un code de calculs sous Matlab qui permettait de paralléliser les simulations. Ainsi, pour tous les calculs à suivre la procédure suivante a été adoptée pour effectuer M simulations :

Étape 1 : Caractérisation des variables d’entrée : La première étape consiste à caractériser toutes les variables en entrée du modèle de simulation. Les grandeurs n’ayant pas été retenues et caractérisées au chapitre 2 sont gardées à leurs valeurs

initiales au fil des simulations. Celles dont les incertitudes ont été quantifiées sont modifiées à chaque pas de temps de simulation. On attribut ainsi à chacune d'elle une incertitude ainsi qu'une loi de distribution. Un échantillonnage suivant ces caractéristiques par la méthode de Sobol est d'abord effectué sur l'intervalle de variation de chaque variable, puis la valeur tirée est ajoutée à la valeur nominale préexistante pour générer les M nouvelles valeurs d'entrée des simulations.

Étape 2 : Génération des fichiers d'entrée Trnsys : Une fois les données générées pour chaque variable, M fichiers d'entrée sont créés pour chacune d'elle et répertoriés dans M répertoires différents indexés afin de faciliter le lancement des simulations. Les fichiers d'entrée du modèle Trnsys sont au format .txt ce qui facilite leur manipulation.

Étape 3 : Lancement des simulations : C'est à cette étape qu'intervient le code de parallélisation des calculs. Le répertoire dans lequel se trouve l'exécutable du modèle Trnsys est dupliqué 20 fois. Ce répertoire contient notamment les fichiers .txt des variables non modifiés. Puis à chaque itération du code Matlab, les fichiers d'entrée modifiés dans les répertoires précédents sont mis dans les répertoires Trnsys par foulée de 20. Une fois les 20 simulations terminées, les fichiers de sortie sont récupérés, indexés et enregistrés. Cette étape permet de garder une trace sur chaque simulation et de se retrouver en cas d'erreur dans le code de calcul.

Étape 4 : Dépouillement des résultats : Lorsque les M simulations sont terminées, les fichiers de sortie Trnsys peuvent être récupérés et traités. La sortie d'intérêt étant la consommation calorifique annuelle, la consommation calorifique dans chaque zone du modèle de simulation est sommée à chaque pas de temps (8760 heures) et ensuite dans toutes les zones. Seules les données de consommation totale du bâtiment ont été stockées.

Cette procédure a été adoptée pour toutes les analyses. Elle nous a permis de gagner en temps de calcul mais aussi de garder une trace sur l'ensemble des simulations effectuées. Lorsque l'on est face à un nombre important de simulations, il est souvent difficile, voir impossible de repérer les erreurs de code éventuelles. En structurant chaque code de calcul de la sorte, nous avons pu localiser les défauts (si besoin y est) et intervenir très rapidement lorsque cela était nécessaire. Aussi, bien avant que les M simulations soient terminées, nous pouvions pré-traiter les résultats déjà obtenus et avancer dans nos travaux.

Dans la suite, pour une meilleure identification graphique, les résultats associés au niveau haut (mesures complètes) seront représentés en bleu et ceux associés au niveau bas (aucune mesure) seront représentés en rouge.

TABLE 3.7 – Résultats d’analyse de sensibilité par la méthode de Morris : distance ξ_i estimée suivant l’équation 3.6 pour 40 trajectoires effectuées au niveau haut et au niveau bas d’incertitude.

Notation	Variable	Niveau haut	Niveau bas
X ₁	Température extérieure	7.51	20.25
X ₂	Rayonnement solaire	1.47	1.49
X ₃	Vitesse du vent	0.44	2.29
X ₄	Débits de ventilation de la CTA BUR	1.34	1.86
X ₅	Débits de ventilation de la CTA SDR	0.41	0.72
X ₆	Efficacité de la CTA BUR	1.31	2.86
X ₇	Efficacité de la CTA SDR	0.54	1.25
X ₈	Perméabilité à l’air	0.024	11.09
X ₉	Température intérieure	6.85	28.18
X ₁₀	Occupation	0.95	0.84
X ₁₁	Apports internes	2.78	8.69

3.5.2 Identification des variables influentes

Le but de cette étape est de pouvoir réaliser une analyse de sensibilité sur les variables retenues dans la section 2.2 suivant leur plage de variation en utilisant la méthode de Morris (section 3.1.2). Nous n’avons pas effectué d’analyse de sensibilité globale de type Sobol sur le modèle initial car cela aurait nécessité 10000 à 15000 simulations pour obtenir des résultats satisfaisants. Ce temps de calcul était trop important comparativement à la méthode de Morris. Rappelons que la variable d’intérêt est la consommation annuelle de chauffage et que les plages de variation considérées sont celles assignées à chaque variable dans la table 2.11 à la page 67. Ainsi les incertitudes du niveau bas (respectivement niveau haut) ont été assignées aux variables dans le modèle a priori (respectivement a posteriori). En effectuant une analyse de sensibilité par screening, à partir d’un nombre de trajectoires r , les indices de sensibilité se stabilisent. L’identification du nombre de trajectoires r optimal, permet de garantir la fiabilité des résultats obtenus sans pour autant augmenter le temps de calcul. Chaque variable a été discrétisée en six niveaux ($q = 6$) et le nombre de trajectoires r a été itéré jusqu’à ce que les indices de sensibilité présentés dans les équations 3.3, 3.4 et 3.5 se stabilisent.

Les résultats sont présentés dans la figure 3.15 (et en Annexes 6.6 figures 6.5 et 6.6). L’indice ξ_i de chaque variable X_i sur la consommation calorifique annuelle est présenté suivant r trajectoires considérées au niveau haut (figure 6.5) et au niveau bas (figure 6.6). On constate qu’à partir de 30 trajectoires l’indice de sensibilité commence à se stabiliser pour la plupart des variables suivant les deux niveaux d’étude, sauf pour quelques unes d’entre elles (température extérieure, rayonnements solaires et les débits de ventilation). Leur effet ξ_i sur la consommation calorifique annuelle se stabilise à partir de 40 trajectoires. Huit heures de simulations sont nécessaires pour effectuer 40 trajectoires; ce qui rentrait tout à fait dans notre budget temps. La comparaison des résultats d’analyse de sensibilité a donc été effectuée pour 40 trajectoires (figure 3.15).

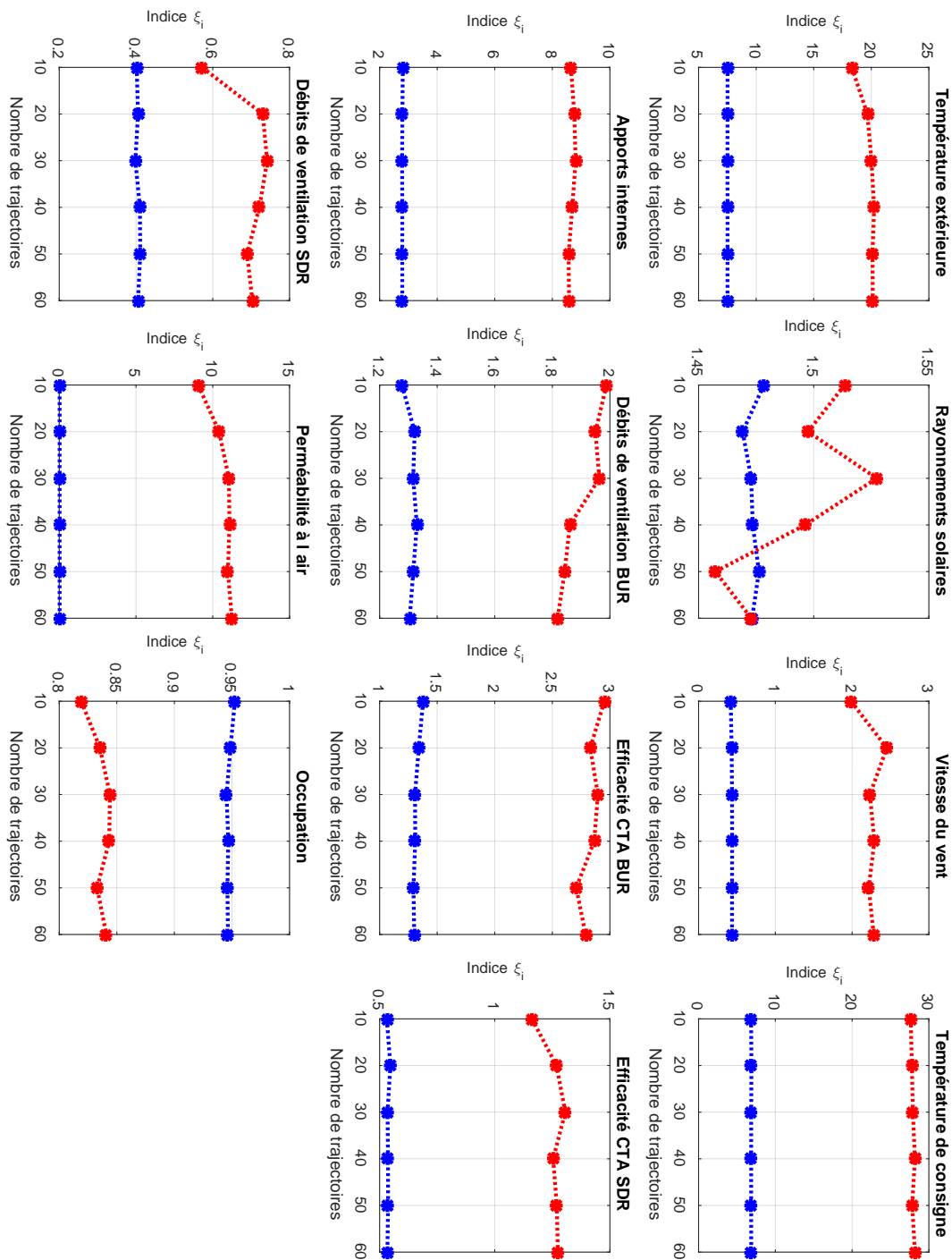


FIGURE 3.15 – Comparaison des indices ξ_i de chaque variable X_i sur la consommation calorifique annuelle (MWh) suivant r trajectoires considérées.

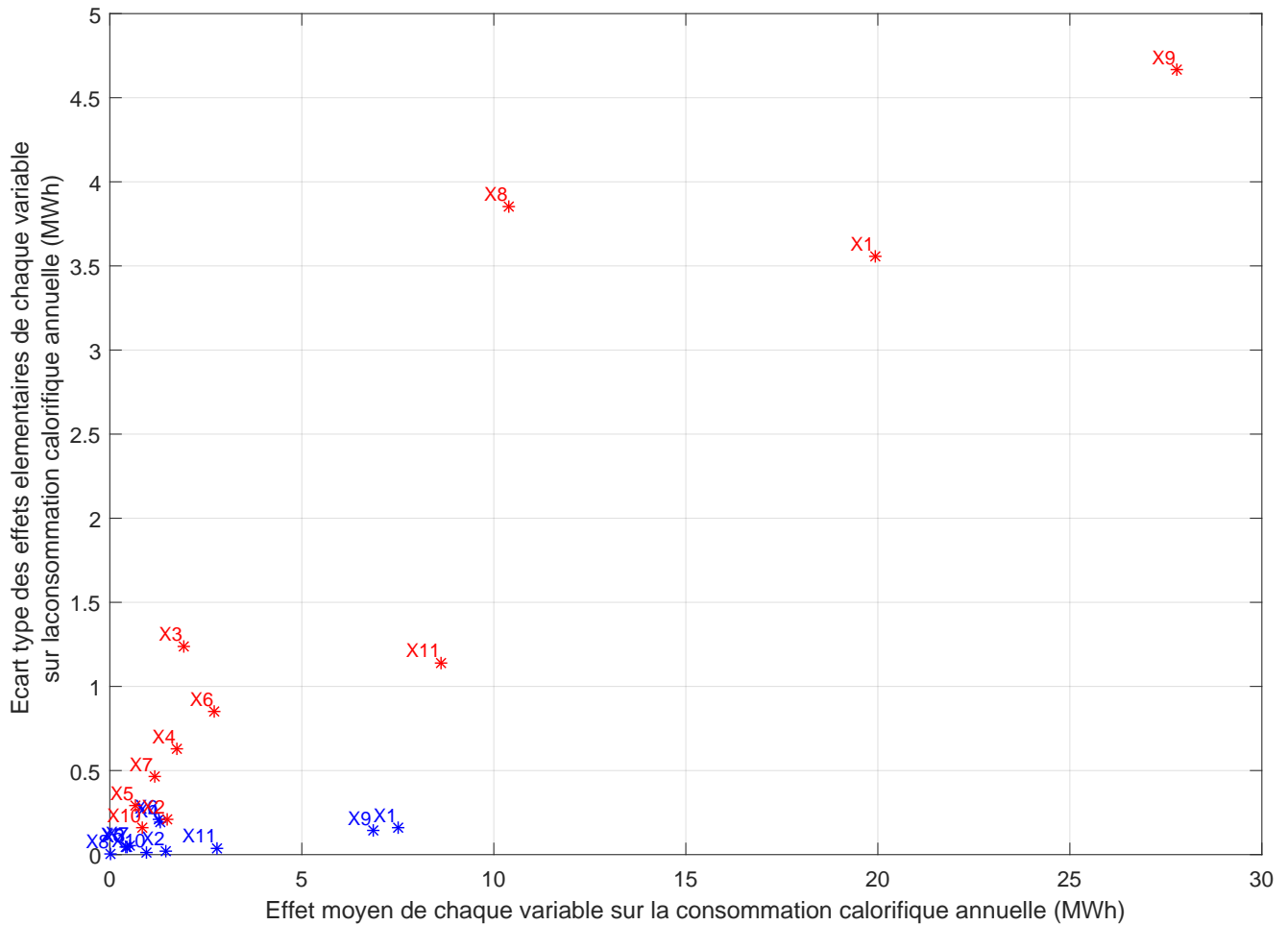


FIGURE 3.16 – Comparaison des effets μ_i et σ_i de chaque variable X_i sur la consommation calorifique annuelle (MWh) suivant 40 trajectoires considérées.

Ainsi, à la figure 3.16, sont représentés les indices de sensibilité sur les deux niveaux d'étude à 40 trajectoires. Ce diagramme permet de renseigner sur le degré d'influence qu'a chaque variable sur la consommation calorifique annuelle ainsi que leur éventuelle implication dans des interactions ou des linéarités (précédemment exposé à la figure 3.1). Les variables ayant une moyenne μ_i importante et une forte dispersion σ_i seront considérées comme influentes et à effets non linéaires et par opposition celles ayant un effet élémentaire μ_i et une dispersion σ_i faibles seront considérées comme ayant un effet négligeable sur la sortie du modèle (section 3.1.2). Un classement des variables est aussi présenté à la figure 3.17. Au niveau haut, les variables $\{X_1, X_9, X_{11}\}$, à savoir la température extérieure, la température intérieure et les apports internes s'avèrent avoir une influence plus importante sur la sortie du modèle que les autres variables. Au niveau bas, les variables $\{X_9, X_1, X_8, X_{11}\}$, à savoir la température intérieure, la température extérieure, la perméabilité à l'air et les apports internes se démarquent aussi des autres.

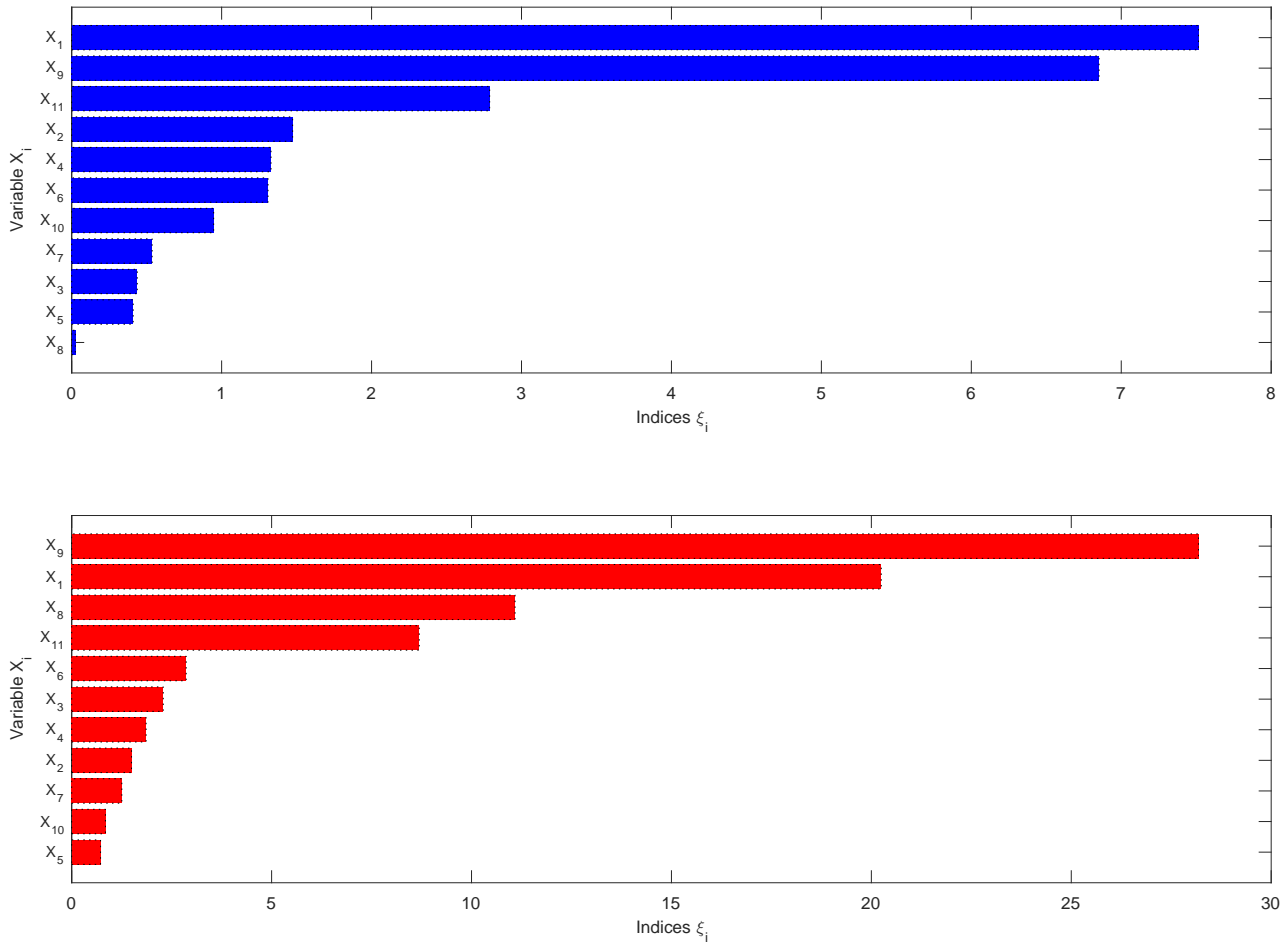


FIGURE 3.17 – Comparaison des indices ξ_i de sensibilité de Morris pour l’analyse de sensibilité au niveau haut (mesures complètes) et au niveau bas (non mesures)

En comparant les résultats du niveau haut et ceux du niveau bas, on s’aperçoit que quel que soit le niveau, la température intérieure, la température extérieure et les apports internes $\{X_9, X_1, X_{11}\}$ se démarquent toujours comme ayant une influence importante et à effet non linéaire sur la consommation calorifique annuelle (figure 3.17). On peut d’ores et déjà considérer qu’elles sont en tête des variables les plus influentes de l’étude car quel que soit le niveau d’incertitudes qui leur est appliqué, elles ressortent du lot.

À ce stade, il ne nous est pas possible de quantifier leur impact sur la consommation calorifique annuelle mais nous verrons plus loin comment l’identification des variables influentes et leur hiérarchisation constitue une réelle avancée dans la démarche d’optimisation des plan de mesure.

Pour conclure, l'analyse de sensibilité par la méthode de Morris a permis d'effectuer une hiérarchisation des variables considérées dans l'étude. Elle nous a fourni une information quantitative sur la sensibilité du modèle aux variables choisies en amont et ceci en seulement huit heures de simulation en dehors du temps de code, de préparation et de post-traitement des données. Ce temps varie bien évidemment en fonction du nombre de variables en jeu et des moyens à disposition pour effectuer les simulations. Et, dans le but d'effectuer un premier tri ou une hiérarchisation des variables d'entrée du modèle l'investissement en temps de calcul s'avère rentable.

3.5.3 Réduction du modèle de simulation a posteriori

Dans les sections précédentes ont été présentées les méthodes d'expansion polynomiale qui permettent de construire un méta-modèle en se basant sur le modèle physico-probabiliste construit. Nous avons vu que la méthode de régression était plus adaptée à l'étude menée et qu'elle permettait non seulement de sélectionner les points constituant le plan d'expériences mais aussi d'en augmenter le nombre pour atteindre un niveau de précision satisfaisant. Un seul niveau d'incertitudes sur les variables aléatoires en entrée du modèle Trnsys a été considéré pour la construction du méta-modèle : le niveau haut. Cette approximation est effectuée sur le modèle a posteriori présenté précédemment à la section 3.4.2 et avec en entrée les variables mesurées. Le but étant de voir si la construction d'un métamodèle sur les données réelles d'exploitation pouvait reproduire le comportement du bâtiment en un temps de simulation relativement faible. Une simulation complète du modèle Trnsys initial s'effectue en une dizaine de minutes.

3.5.3.1 Identification du degré optimal

Ainsi, pour différents degré p d'étude, une approximation polynomiale a été effectuée comme présenté dans la table 3.9. La taille de l'approximation polynomiale, pour une interaction complète entre les variables, est déterminée par le nombre de coefficients de l'expansion polynomiale à déterminer P , et est donnée par l'équation 3.43 et la taille n du plan d'expériences sur lequel le méta-modèle sera construit par l'équation 3.44.

$$P = \frac{(N + p)!}{p!N!} = \binom{N + p}{p} \quad (3.43)$$

$$n = 2 * (P + 1) \quad (3.44)$$

Il est possible de réduire la taille du plan d'expériences en ne considérant que les interactions du second ordre entre les variables (table 3.9). Cela permettra de réduire la taille de l'expansion polynomiale et par la même occasion le temps de calcul. Le chaos polynomial convergerait donc plus rapidement. Vu le nombre important de simulations à

effectuer pour construire le méta-modèle, les calculs sont parallélisés suivant l'algorithme physico-probabiliste. Le seuil d'erreur a été fixé à 25% pour l'erreur relative E et l'erreur en validation croisée (équations 3.34 et 3.36) par analogie au CVRMSE fixé par l'ASH-RAE et l'IPMVP. Au delà de 25 %, on considère que le méta-modèle n'est pas assez précis pour transcrire toutes les caractéristiques du modèle initial. Ce seuil est dépassé au degré $p = 3$ du polynôme du chaos (figure 3.18) pour l'erreur relative mais n'est jamais atteint pour l'erreur en validation croisée. Si l'on ne considère que l'erreur relative qui n'évalue le modèle uniquement sur la base d'apprentissage, une expansion polynomiale de degré 2 est amplement suffisante. L'erreur d'approximation serait de 15 % et on obtiendrait un méta-modèle avec un temps de simulation de 0.21 secondes. Or au degré 2 l'erreur en validation croisée est de 40 % et ne fait que croître en augmentant le degré du polynôme.

TABLE 3.8 – Taille du plan d'expériences en fonction du degré polynomial considéré pour 11 variables en considérant toutes les interactions

Degré polynomial p	1	2	3	4	5	6
Taille du chaos P	12	78	364	1365	4368	12376
Taille du plan d'expériences n	26	158	730	2732	8738	24754

TABLE 3.9 – Taille du plan d'expériences en fonction du degré polynomial considéré pour 11 variables dans le cas d'un chaos polynomial tronqué aux interactions d'ordre 2

Degré polynomial p	1	2	3	4	5	6
Taille du chaos P	12	78	199	375	606	892
Taille du plan d'expérience n	26	158	400	752	1214	1786
Erreur relative E	0.16	0.15	0.29	0.31	1.3	1.91
Erreur E_{LOO}	0.34	0.4	1.27	0.74	27.91	29.98
Temps de simulation (secondes)	0.034	0.21	0.54	1.25	1.55	2.87

3.5.3.2 Comparaison qualitative et quantitative

Nous avons vu à l'étape précédente que, lorsque l'on évaluait le méta-modèle sur les points du plan d'expérience qui ont servi à sa construction, l'erreur d'estimation était très faible mais dès qu'il était évalué sur d'autres points, elle ne cessait de croître. L'objectif de cette étape est de comparer l'estimation de la consommation calorifique annuelle par les deux modèles sur des points autre que ceux de la base d'apprentissage. 200 simulations ont ainsi été réalisées. À la figure 3.19, est représenté le nuage de points des consommations de chauffage construit aux mêmes points des variables d'entrée par les deux modèles. On constate qu'il y a une relation linéaire entre le modèle Trnsys et le méta-modèle mais que les points ne sont pas alignés sur la première bissectrice. Les points sont dispersés et traduisent une non-régularité de la part du modèle.

En effectuant une comparaison supplémentaire des distributions des consommations de chauffage des deux modèles (figure 3.19), on constate que la densité de probabilité du méta-modèle est plus éfilée. Les points qui constituent sa courbe sont moins dispersés par rapport au modèle trnsys. L'irrégularité du modèle est aussi constatée. Les deux distributions sont néanmoins axées sur la même moyenne (table 3.10) et présentent une allure similaire. En comparant les moments en tendance centrale (moyenne et écart type) obtenus sur les deux modèles à la table 3.10, la différence entre les deux écarts type est acceptable.

TABLE 3.10 – Comparaison des caractéristiques du modèle initial et de celles du méta-modèle aux points auxquels le méta-modèle a été construit (MWh) suivant les équations 3.24 et 3.25.

Modèle	Moyenne	Écart type
Modèle Trnsys	32.29	2.8
Métamodèle PC	32.29	2.05

TABLE 3.11 – Indices de Sobol du premier ordre obtenus en réarrangeant les coefficients du méta-modèle sur l'expansion polynomiale de degré 2

Indices de Sobol		Indices de Sobol classés	
S ₁	0.0000093	S ₆	0.0724
S ₂	0.00000105	S ₇	0.0460
S ₃	0.0058	S ₅	0.0268
S ₄	0.0051	S ₁₀	0.0112
S ₅	0.0268	S ₃	0.0058
S ₆	0.0724	S ₄	0.0051
S ₇	0.0460	S ₁₁	0.00124
S ₈	0.000458	S ₈	0.000458
S ₉	0.0000238	S ₉	0.0000238
S ₁₀	0.0112	S ₁	0.0000093
S ₁₁	0.00124	S ₂	0.00000105

On peut aussi comparer les indices de sensibilité de Sobol obtenus sur les polynômes du chaos avec les résultats de l'analyse de sensibilité de screening. Les indices de sensibilité de Sobol décrits précédemment par les équations 3.29 et 3.30 ont été estimés en réarrangeant les coefficients du polynôme. Les indices d'ordre 1 sont présentés à la Table 3.11. Les indices totaux contenant en eux l'effet de chaque variable mais aussi l'effet de leur interaction respective sur la sortie du modèle, seuls les indices d'ordre 1 ont été considérés. Ces indices sont notamment censés fournir des informations complémentaires et venir corroborer celles obtenues par l'analyse de sensibilité conduite par screening. Cependant, on constate que les variables qui ressortent comme ayant un fort impact sur la variabilité de la consommation calorifique annuelle sont l'efficacité des deux CTA, X_6 et X_7 . En représentant les nuages de points des température d'air en insufflation des deux CTA et de celle de la sortie du

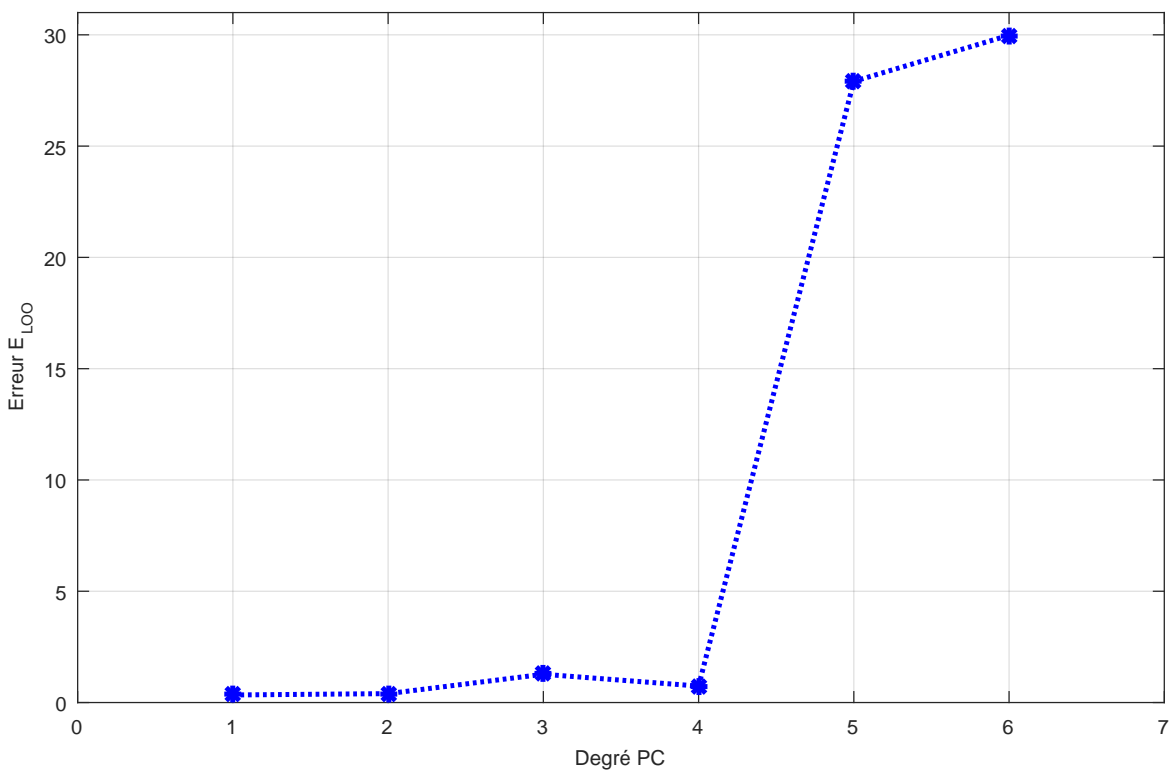
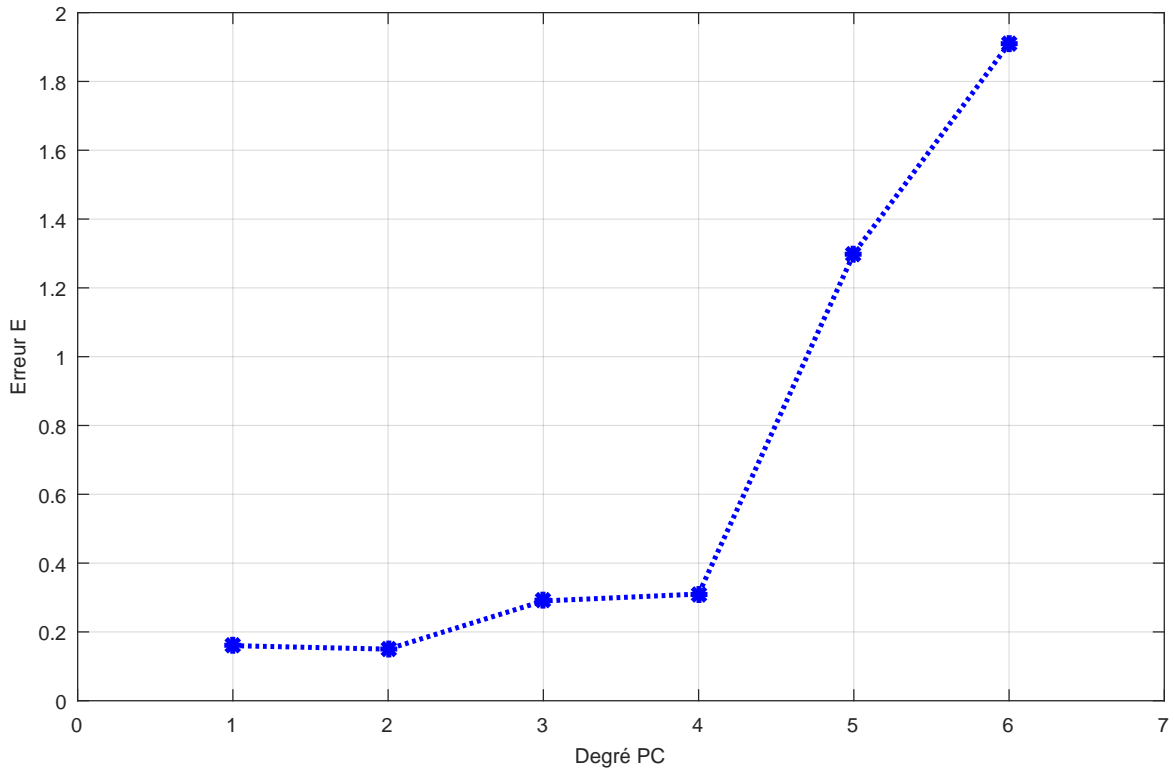


FIGURE 3.18 – Évolution de l'erreur relative E et de l'erreur en validation croisée E_{LOO} en fonction des différents degrés des polynômes suivant les termes d'interactions d'ordre 2

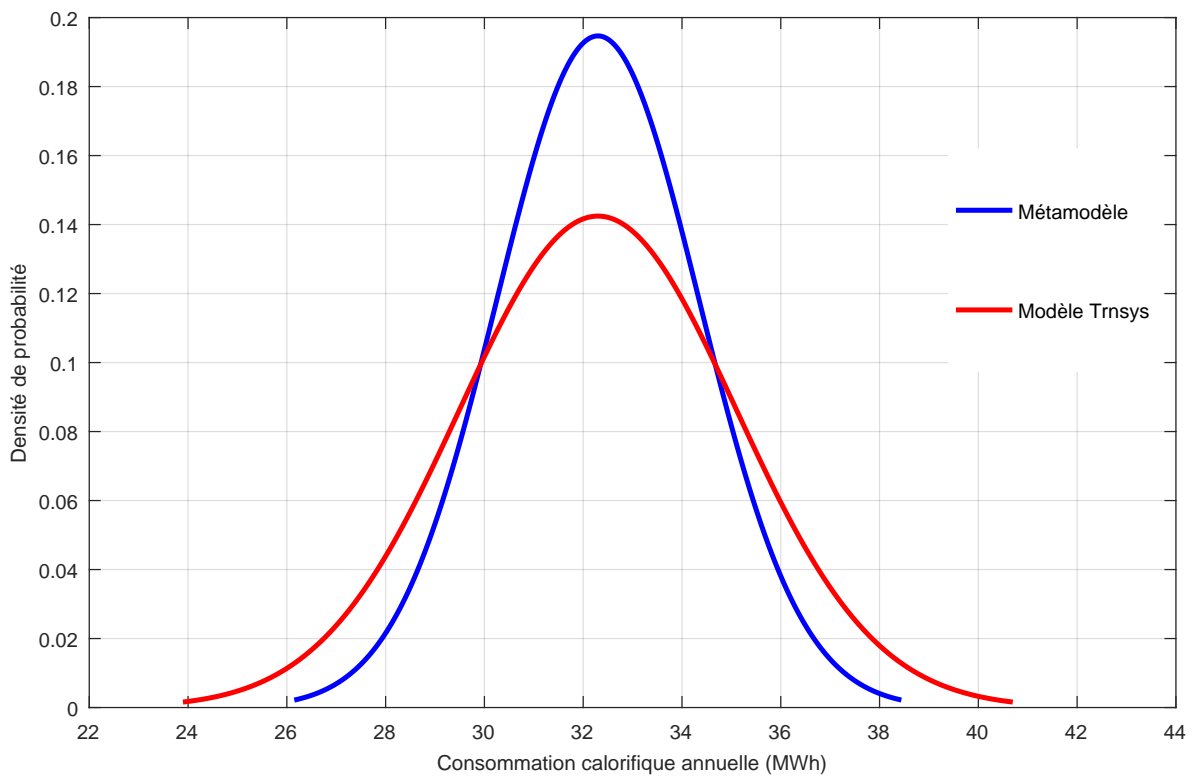
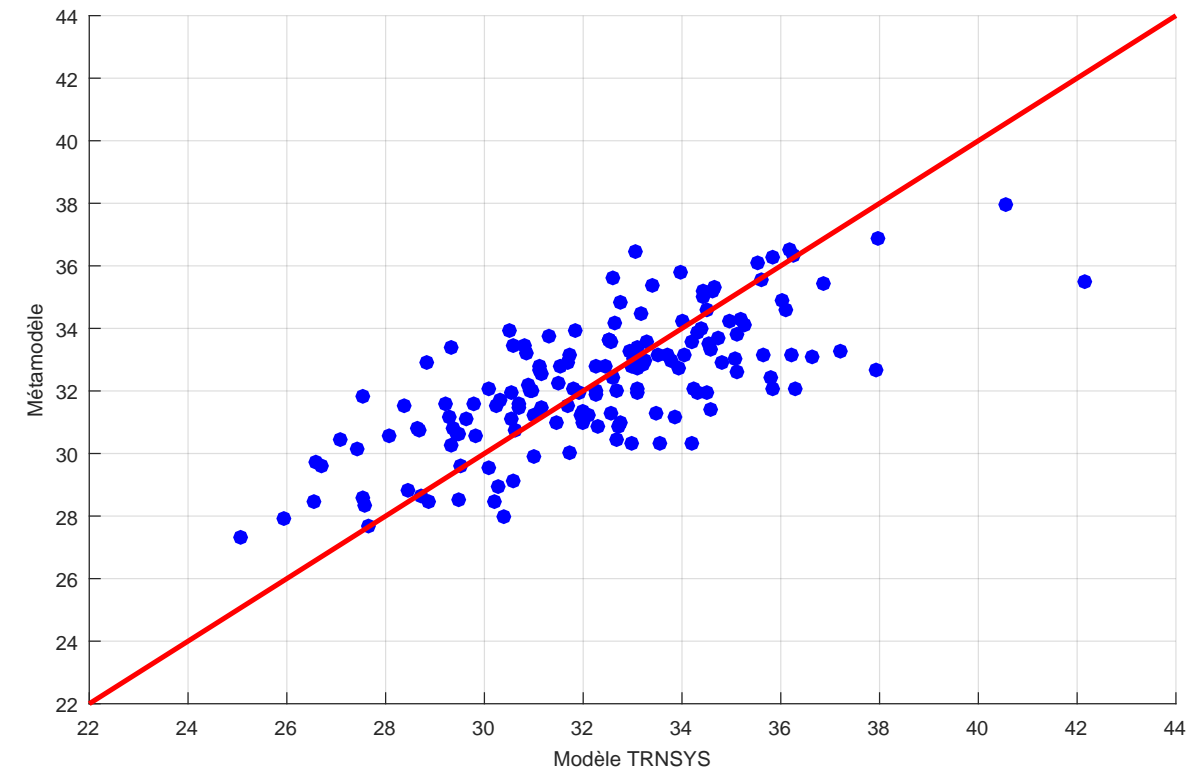


FIGURE 3.19 – Comparaison de la consommation calorifique annuelle (MWh) du bâtiment test du modèle initial Trnsys et du méta-modèle au degré 2

modèle, nous avons pu constater un effet de seuil. La modélisation des deux centrales de traitement d'air sont responsables de la non régularité du modèle.

3.5.3.3 La non régularité du modèle et l'effet de seuil

Un modèle physico-probabiliste régulier traduit toutes les perturbations appliquées aux variables d'entrée sur les sorties de façon linéaire. Tandis qu'un modèle non-régulier a tendance à ne pas forcément traduire ces perturbations sur certaines sorties du modèle : dans notre cas il s'agit de la consommation de chauffage horaire. En effet, lorsque certains systèmes énergétiques (centrales de traitement d'air, pompes à chaleur) sont installés dans un bâtiment et qu'ils sont régulés pour ne fonctionner que sous certaines conditions spécifiques (au-dessus ou en dessous d'un maximum ou un minimum de température par exemple), la consommation énergétique du bâtiment en est affectée et ne fonctionne plus de façon linéaire. Ces mécanismes sont souvent présents pour éviter des phénomènes de surchauffe et se déclenchent lorsque la température du bâtiment est assez élevée pour ne pas nécessiter d'apport calorifique supplémentaire. C'est souvent le cas lorsque les apports internes métaboliques et électriques sont suffisants pour chauffer les locaux ou lorsque certains locaux bénéficient d'apports énergétiques gratuits de la part de locaux voisins.

Comme a pu le souligner Rivalin dans ses travaux ([Rivalin, 2016](#)), cet effet de seuil a un impact sur la régularité du modèle. Et lorsqu'un modèle n'est pas régulier, il est difficile de le réduire et de construire un méta-modèle qui puisse traduire cet effet de seuil. Deux solutions existent. La première consiste à élever le degré du polynôme (un polynôme de degré 9 voir 10) qui permettrait d'approximer au mieux la sortie et diminuer l'erreur relative. L'inconvénient de cette méthode est qu'elle nécessite un nombre important d'appel au modèle pour construire l'approximation polynomiale (un polynôme de degré 10 dans notre cas, aurait nécessité 5174 appels au modèle initial). La seconde solution consiste à construire un méta-modèle composé d'approximation polynomiale spécifique à chaque régime de fonctionnement des systèmes ([Nouy, 2010](#)). Ce méta-modèle serait ainsi mieux adapté au fonctionnement des systèmes et permettrait de palier à l'effet de seuil provoqué par les centrales de traitement d'air. Il serait plus robuste et chaque sous-méta-modèle pourrait être construit à des degrés nettement bas (2 voir 3), pour une évaluation très rapide.

Dans nos travaux, nous avons vu que l'approximation polynomiale construite ne permettait pas d'obtenir les indices de sensibilité de Sobol mais permettait d'effectuer une analyse en tendance centrale (moyenne, écart type) robuste. Le méta-modèle pouvait donc être utilisé à ces fins.

3.6 Conclusion du chapitre

Dans ce chapitre, une méthodologie de construction d'un modèle physico-probabiliste a été présentée dans un premier temps. Le modèle a été construit à partir des caractéristiques physiques et thermiques du bâtiment et une dimension probabiliste lui est ensuite attribuée pour prendre en compte le caractère stochastique des variables dynamiques. Le modèle a ensuite été comparé aux mesures in situ et considéré cohérent suivant les indicateurs proposés par l'ASHRAE. Néanmoins, vu l'exigence du modèle en temps de calcul, il est coûteux en temps d'y appliquer des méthodes d'analyse de type propagation d'incertitudes ou analyse de sensibilité. La méthode de propagation d'incertitudes de Monte Carlo par exemple, exige un nombre important ($\approx 10^3$) de calculs pour l'évaluation des moments statistiques de la variable d'intérêt. Nous avons proposé de réduire le modèle Trnsys en construisant une surface de réponse par les polynômes du chaos.

Par ailleurs, les différentes sources d'incertitudes présentées et quantifiées dans le chapitre 2 sont multiples. Ne disposant pas d'informations préalables quant à leurs poids respectifs sur la consommation calorifique annuelle, elles doivent toutes être prises en considération. Ce qui engendre la présence d'un nombre conséquent de variables incertaines pour la suite. Afin de surmonter cette difficulté, des analyses de sensibilité ont été effectuées. Car, parmi l'ensemble des variables incertaines agissant sur la sortie d'un modèle, seul un nombre limité d'entre elles contribuent effectivement à la variabilité de cette dernière. Une analyse de sensibilité par screening a dans un premier temps été effectuée sur les variables considérées. Elle a permis d'effectuer un premier tri et de dégager trois variables comme étant les plus prépondérantes. L'analyse de sensibilité par screening ne fournissant que des résultats d'ordre qualitatif, une analyse de sensibilité globale par la méthode de Sobol a été effectuée sur le méta-modèle construit. Les résultats n'ont pas permis de confirmer la hiérarchisation des variables précédemment effectuée car le modèle présentait une irrégularité.

Une des solutions par la suite, pour l'optimisation du plan de mesure serait de ne considérer que les incertitudes des variables présentant une forte influence sur la consommation calorifique annuelle. Les autres variables étant fixées à leur valeur nominale. Suite à la réduction du modèle physico-probabiliste construit et à l'identification des variables les plus influentes, il sera question dans le chapitre 4, de développer une méthodologie permettant de faire le choix d'un plan de mesure optimal à mettre en place en fonction du niveau d'incertitude (niveau haut ou niveau bas) associé à chaque variable.

Optimisation du plan de mesure

Introduction

Comme expliqué dans le chapitre 1, les gains d'énergie réalisés suite à la mise en place d'actions de performance énergétique ne peuvent être évalués directement. Les plans de mesure et de vérification permettent de quantifier ces gains en mettant l'accent sur la précision des différentes options de mesure. Ainsi, selon les quatre options proposées par l'IPMVP, un MV permettra de mesurer soit l'intégralité des grandeurs ou soit seulement les grandeurs clés et en effectuant une estimation sur les autres. Le choix de l'option dépend en général de la précision souhaitée. Dans ce chapitre, il sera question d'étudier l'effet effectif de la mesure des grandeurs sélectionnées dans le chapitre 3 sur la consommation calorifique annuelle. Les incertitudes quantifiées pour chacune d'elle (niveau haut et niveau bas) feront alors l'objet d'une propagation d'incertitudes à travers le modèle physico-probabiliste construit. Les plans de mesure et de vérification sont situés dans un premier temps dans le processus des contrats de performance énergétique. Puis, la méthodologie d'analyse d'incertitudes suivant chaque niveau d'études et pour chaque variable incertaine considérée est présentée. La méthodologie est ensuite appliquée sur le cas d'étude du bâtiment du Cerema.

4.1 La place des MV au sein des contrats de performance énergétique

4.1.1 La procédure de contractualisation

Comme évoqué précédemment à la section 1.2.1.2, la directive européenne 2012/27/UE a mis sur pied un moyen légal obligeant les prestataires de services à atteindre les objectifs d'efficacité fixés lors de la livraison des bâtiments. Cette obligation est contractualisée entre le prestataire de services et le bénéficiaire et a pour objectif d'améliorer l'efficacité énergétique à travers des dispositifs d'amélioration consentis afin de parvenir à un niveau d'amélioration de la performance énergétique précisé contractuellement. En dehors de ces actions de performance, l'évaluation de la performance peut aussi avoir lieu dans une démarche de suivi.

Dans le cadre d'actions de performance énergétique, la procédure de contrat de performance énergétique comprend plusieurs étapes comme l'a souligné Duplessis dans ses travaux (Duplessis, 2008). Le prestataire effectue dans un premier temps une analyse préalable du site et estime le potentiel d'amélioration des installations énergétiques sur lesquelles il pense pouvoir s'engager. Ces améliorations, qui constituent des actions de performance énergétiques (APE), peuvent porter aussi bien sur le bâti (performances thermiques), sur les équipements, que sur les conditions d'usage et d'exploitation. Des solutions sont ensuite proposées au client suivant les économies d'énergie envisageables sur le projet ainsi que les modalités de contractualisation comprenant toutes les données financières. Lorsque le client décide de poursuivre la procédure, les deux parties s'engagent en outre sur le suivi des performances et particulièrement sur les méthodes de calculs des économies d'énergies. Ainsi, après la mise en œuvre des travaux d'amélioration prévues, la phase de suivi, de mesure et de vérification de la performance est amorcée. Cette étape est essentielle et permet de vérifier les engagements de garantie. Les résultats de mesure permettront de rémunérer les deux parties suivant les économies d'énergies constatées. Les protocoles de mesure et de vérification de la performance qui ont été contractualisés sont alors mis en place.

De ce fait, parmi les éléments définis dans un contrat de performance énergétique figurent entre autre :

- 1. La procédure de mesure et de vérification détaillée, indiquant les grandeurs à mesurer, le périmètre de mesure ainsi que les indicateurs permettant de vérifier la garantie (les variables considérées ainsi que leurs modalités, les formules d'ajustements, les périodes de mesure, ...)** ;
- 2. La performance attendue et garantie suite aux économies d'énergies.**

En résumé, l'évaluation de la performance énergétique fait partie intégrante des contrats de performance énergétiques. Le prestataire de service et le client s'engagent sur la mé-

thodologie adaptée pouvant chiffrer l'amélioration réelle des actions entreprises ainsi que la performance du bâtiment. C'est dans cette logique que dans la suite, à chaque plan de mesure proposé, nous définirons les grandeurs à mesurer, les spécificités de chaque mesure ainsi que la performance qui pourrait être garantie. Cette consommation sera à chaque fois comparée à la consommation réelle pour que l'opérateur puisse évaluer la portée des mesures qu'il entreprend.

4.1.2 Équilibre entre coût et précision

Les quatre options proposées par l'IPMVP permettent de sélectionner le plan de mesure adapté en fonction de l'ampleur du projet. Elles offrent la possibilité de faire un choix en fonction de l'étendue du projet, du temps de mise en œuvre, des moyens dont dispose l'opérateur mais surtout de l'exigence en précision sur l'évaluation de la performance. Cependant cette précision a un coût qui est proportionnel aux moyens déployés. Le coût d'un plan de mesure dépend de la qualité des moyens mis en œuvre (équipements et instruments de mesure), de la durée de l'évaluation, mais dépend aussi des moyens humains déployés (mains d'œuvre à disposition pour le post traitement des données, la collecte d'informations pour l'estimation des variables à estimer, la construction d'un modèle de simulation, ...). L'opérateur doit dans ce sens disposer de tous les moyens nécessaires pour optimiser ses critères de sélection et trouver un équilibre entre le coût du plan de mesure et la précision des résultats.

Or la précision d'évaluation de la performance est proportionnelle à l'incertitude des protocoles de mesure. L'IPMVP préconisa dans ce sens que le coût d'évaluation de la performance ne dépasse pas 10% des économies d'énergie dans le cadre de la mise en place d'actions de performance énergétique. En tenant compte des éléments présentés ci-dessus, nous proposons une méthodologie d'aide à l'identification d'un plan de mesure adaptée à la dimension d'un projet donné et basée sur la contrainte technico-économique à laquelle tout prestataire de services est confronté.

4.2 Méthodologie d'optimisation des plans de mesure

Dans cette section, sont présentées les différentes étapes qui constituent la méthodologie d'optimisation des plans de mesure. Nous nous plaçons dans le contexte d'un prestataire de services ou d'un opérateur qui souhaiterait définir le plan de mesure idéal à mettre en place pour évaluer la performance de son bâtiment. Pour ce faire, nous préconisons sur la base des chapitres précédents les étapes suivantes :

Étape 1 : Construction du modèle de simulation thermique dynamique :

La première étape consiste à construire un modèle de simulation comme effectué au chapitre 3 à la section 1.4. Il permet à l'opérateur de poser les bases de son étude mais aussi le périmètre des variables qui entrent dans son champ d'action.

L'opérateur se place dans une phase d'étude. Aucune mesure n'a donc été effectuée pour caractériser ces grandeurs. Le modèle construit ici est donc un modèle a priori.

Étape 2 : Caractérisation des variables d'entrées : Les variables choisies à l'étape 1 sont caractérisées. Ces variables peuvent contenir aussi bien les scénarios d'usage, les variables météorologiques que les paramètres intrinsèques à la structure du bâti. L'utilisateur choisit donc les variables en fonction de l'orientation qu'il souhaite donner à son étude. Les valeurs nominales des variables sont définies ainsi que les incertitudes qui leurs sont associées. Ces dernières sont estimées sur la base qu'aucune mesure n'a été effectuée.

Étape 3 : Performance garantie a priori : Une fois le modèle construit et les incertitudes caractérisées, le modèle STD a priori peut être réduit (si l'opérateur le souhaite et si la régularité du modèle le permet) suivant la méthode exposée dans le chapitre 2. Les analyses de sensibilité et de propagation d'incertitudes pourront être effectuées à moindre coût de simulation. La propagation d'incertitudes va permettre de connaître toutes les valeurs possibles que peut prendre la consommation calorifique annuelle et ainsi de pouvoir fixer une consommation garantie a priori (avec un risque de 5%) dans le cadre de non mesures. Cette consommation garantie a priori va ensuite servir de repère dans le processus d'optimisation et permettra de juger si la mesure apporte de la précision ou non à la démarche entreprise. L'analyse de sensibilité quant à elle va permettre de hiérarchiser les variables et orientera l'opérateur sur l'implication de chacune d'elle sur la variabilité de la consommation calorifique annuelle. Elle donne ainsi une première indication des variables sur lesquelles pourront porter les mesures par leur classement (figure 4.1).

Étape 4 : Stratégie d'optimisation : Cette étape fait aussi appel aux connaissances dont peut disposer l'opérateur. Pour chaque variable choisie précédemment dans son étude, il aura à caractériser les incertitudes dans un contexte de mesures de celles-ci. Puis-qu'aucune mesure n'aura été effectuée au préalable, les incertitudes de mesure seront estimées selon ses connaissances et devront représenter le plus fidèlement possible les mesures comme si elles avaient été effectuées. Puis sur la base de la hiérarchisation effectuée précédemment, on proposera un plan de mesure pour chaque niveau d'incertitude défini dans l'équation 4.1. $u_b(X_i)$ est l'incertitude associée à la variable X_i au niveau bas (incertitude à priori) et inversement $u_h(X_i)$ est l'incertitude associée à la variable X_i au niveau haut (incertitude de mesure estimée). On obtient $N + 1$ plans de mesure au total. Le plan \mathcal{MV}_i correspondant à celui effectué à l'étape précédente. Pour chaque \mathcal{MV}_i , on effectue une propagation d'incertitudes à travers le modèle physico-probabiliste ou polynomial. On obtient N incertitudes sur la consommation calorifique annuelle.

Étape 5 : Identification du plan de mesure optimal : Les propagations d'incertitudes effectuées à l'étape précédente et celle effectuée sur le modèle a priori

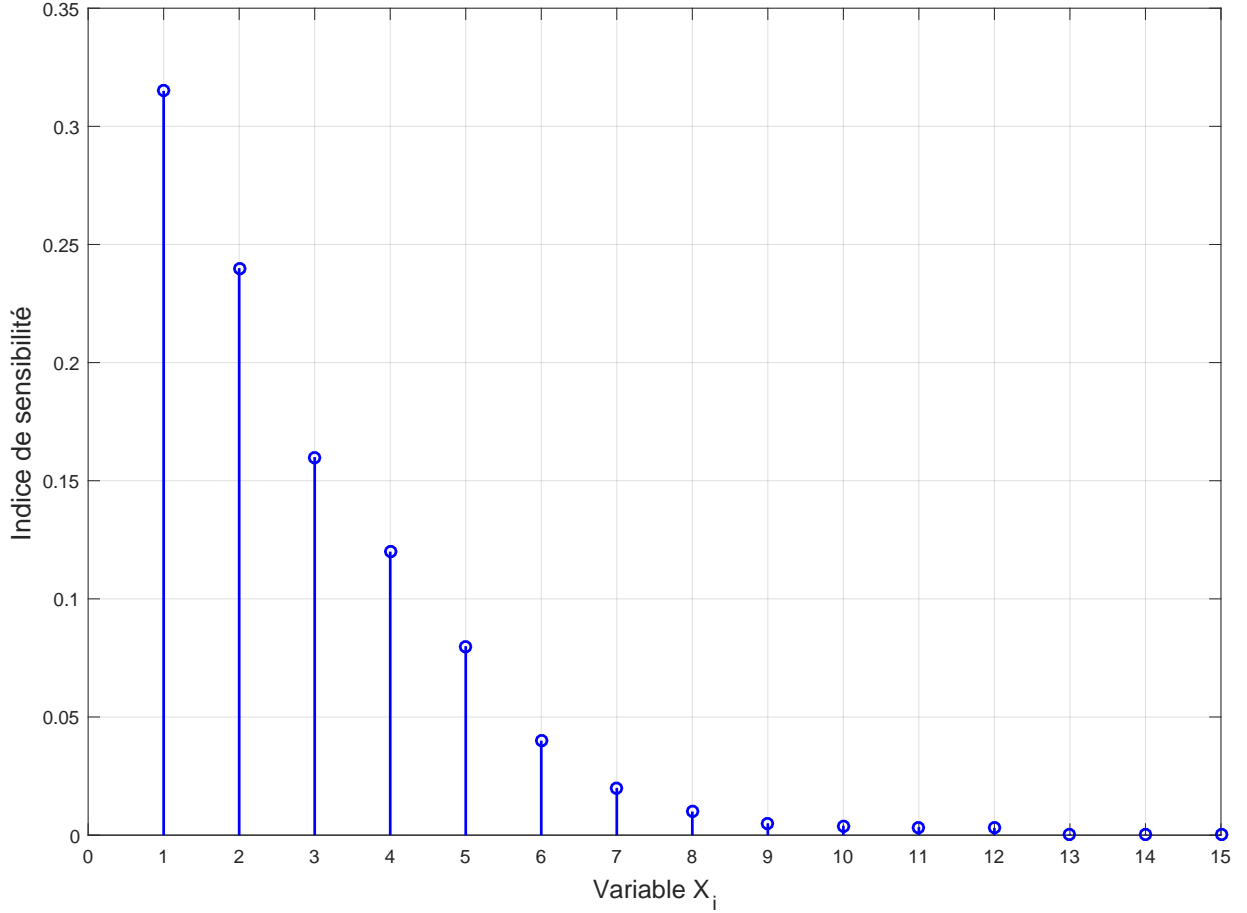


FIGURE 4.1 – Hiérarchisation des variables considérées en se référant à leurs indices de sensibilité

$$\begin{array}{rcl}
 \mathcal{MV}_0 : & u_b(X_0), & u_b(X_i), \dots, u_b(X_N) \\
 \mathcal{MV}_1 : & u_h(X_0), & u_b(X_i), \dots, u_b(X_N) \\
 \mathcal{MV}_i : & u_h(X_0), & u_h(X_i), \dots, u_b(X_N) \\
 \vdots & \vdots & \vdots \quad \ddots \quad \vdots \\
 \mathcal{MV}_{N+1} : & u_h(X_0), & u_h(X_i), \dots, u_h(X_N)
 \end{array} \tag{4.1}$$

permettent d'obtenir $N + 1$ incertitudes associées à la consommation calorifique annuelle pour chaque \mathcal{MV}_i . Pour choisir le plan de mesure optimal, on peut procéder de deux manières. On peut soit choisir le \mathcal{MV} à partir duquel l'incertitude sur la consommation calorifique annuelle commence à se stabiliser (i.e. que la mesure n'a pratiquement plus aucun apport sur la précision de l'estimation), ou soit fixer un seuil de précision. Ce seuil correspondra au niveau de précision fixé par l'opérateur sur sa consommation garantie. Nous choisissons dans nos travaux d'adopter la méthode du Pareto qui consiste à identifier les $\alpha\%$ des variables qui contribueront à

améliorer de $(1 - \alpha)\%$ la précision d'évaluation de la performance (figure 4.2).

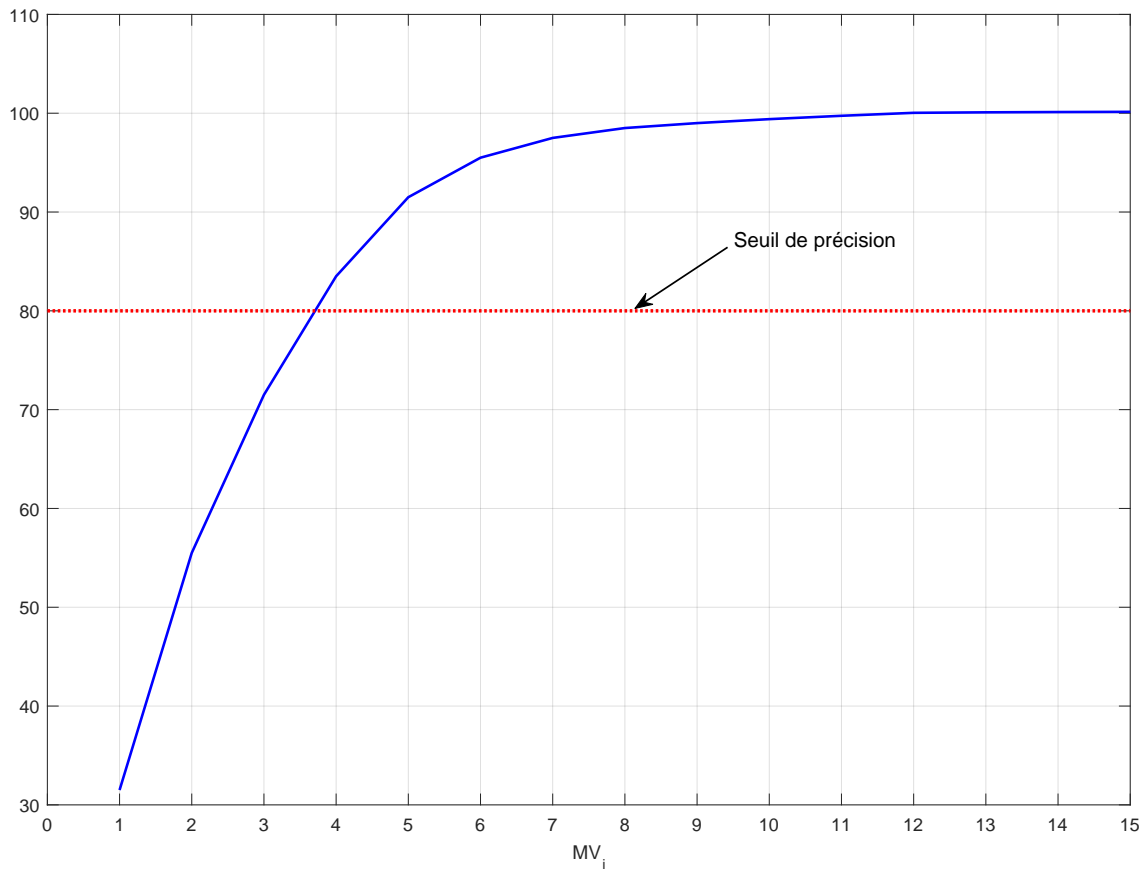


FIGURE 4.2 – Identification du plan de mesure optimal en fonction du seuil de précision fixé

Étape 6 : Intégration des incertitudes des paramètres statiques : Une fois que le plan de mesure optimal a été identifié, les incertitudes des paramètres statiques sont intégrés à la procédure de propagation d'incertitudes afin de prendre en considération les performances énergétiques réelles du bâti liées à l'isolation thermique de l'enveloppe, ainsi que les caractéristiques du site.

Étape 7 : Ajustement de la consommation garantie : Pour finir, la consommation garantie suite à la propagation d'incertitudes effectuée à l'étape 6 est ensuite ajustée avec les mesures effectuées sur les variables identifiées par le plan de mesure optimal.

En résumé, la méthodologie proposée, illustrée à la figure 4.3 pour l'optimisation du plan de mesure est constituée de sept étapes clés.

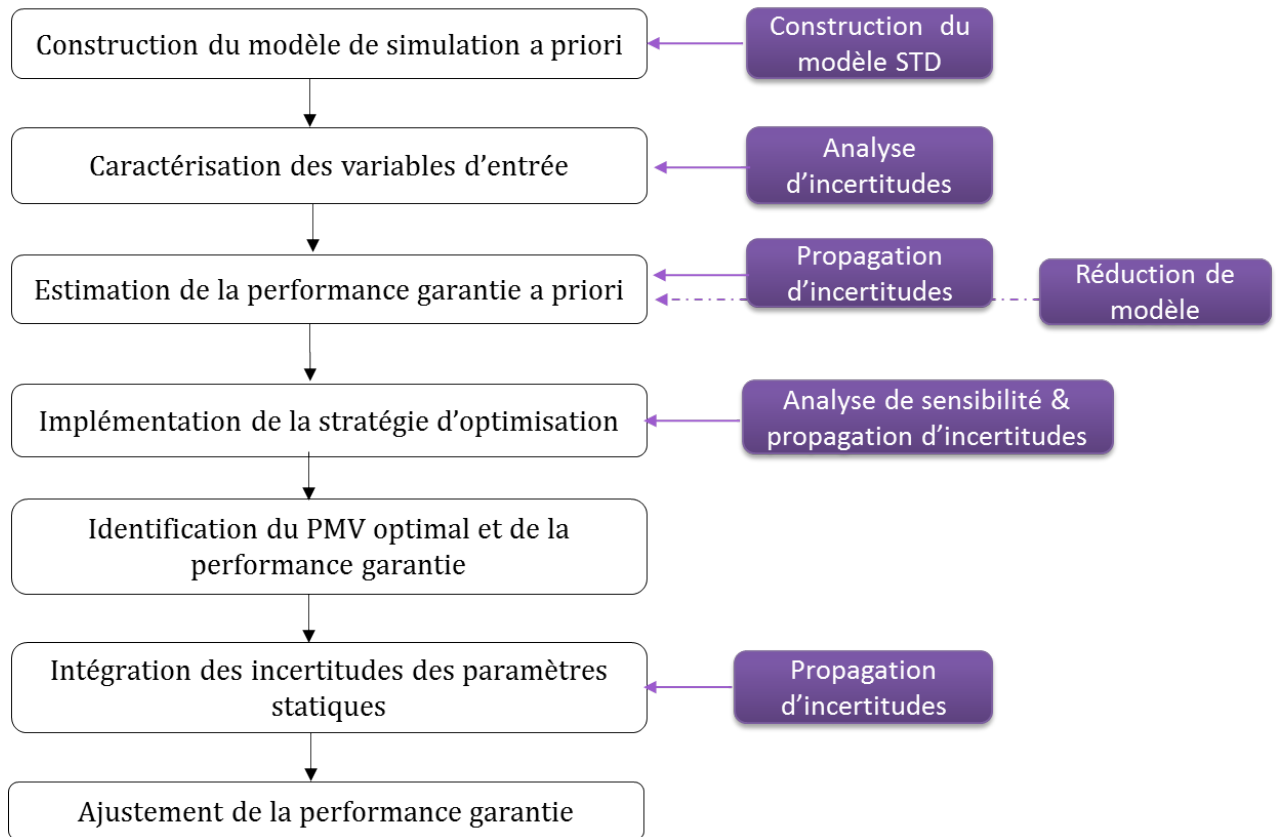


FIGURE 4.3 – Méthodologie d'optimisation d'un plan de mesure ainsi que les outils nécessaires et utilisés à chaque étape

4.3 Cas d'étude

Nous proposons d'appliquer la méthodologie d'optimisation du plan de mesure sur le bâtiment test présenté au chapitre 3. L'idée étant de vérifier si la méthodologie proposée permet d'évaluer la performance énergétique du bâtiment. La performance réelle du bâtiment sur l'année d'étude est de 36.67 MWh. Un modèle a priori basé sur les connaissances antérieures dont nous disposons a été construit. Dans ce modèle, les scénarios d'occupation, les apports internes et la température intérieure appliquée sont ceux provenant des enseignements tirés des opérations PREBAT (PREBAT, 2015) et exclusivement sur les bâtiments tertiaires. Ces données sont présentées en Annexe 6.6. Quant aux données météorologiques, elles sont générées par la plate-forme Météonorm. Les autres paramètres tels que les débits de ventilation et l'efficacité des échangeurs sont issus des données constructeur. On assigne ensuite à chaque variable, une incertitude de non mesure, qui constitue l'incertitude caractérisée au chapitre 2 (table 2.11, page 67).

4.3.1 Utilisation des indices de sensibilité ξ_i de Morris

Nous proposons maintenant d'appliquer la méthodologie d'optimisation des plans de mesure mais cette fois-ci avec les indices de Morris ξ_i obtenus à la section 3.5.2. En effet les indices de Morris sont une alternative aux indices de Sobol. Elles fournissent un classement des variables d'entrée du modèle sans toute fois quantifier leur impact sur la consommation calorifique annuelle mais peuvent servir à optimiser le plan de mesure. Les indices sont classés dans un premier temps à la table 4.1.

TABLE 4.1 – Indices de Morris, ξ_i estimés suivant l'équation 3.6 pour 40 trajectoires effectuées au niveau bas d'incertitude.

Indices de Morris		Indices de Morris classés	
X_1	20.25	X_9	28.18
X_2	1.49	X_1	20.25
X_3	2.29	X_8	11.09
X_4	1.86	X_{11}	8.69
X_5	0.72	X_6	2.86
X_6	2.86	X_3	2.29
X_7	1.25	X_4	1.86
X_8	11.09	X_2	1.49
X_9	28.18	X_7	1.25
X_{10}	0.84	X_{10}	0.84
X_{11}	8.69	X_5	0.72

4.3.2 Méthodologie d'optimisation appliquée au modèle Trnsys a priori

Après avoir classé les variables suivant leur influence sur la consommation calorifique annuelle, la stratégie d'optimisation du plan de mesure peut débuter. À chaque plan de mesure \mathcal{MV}_i , une variable X_i est ajoutée au plan de mesure précédent jusqu'à obtenir un plan de mesure complet. Une propagation d'incertitudes par Monte Carlo est ensuite réalisée à chaque itération. Les variables d'entrée sont échantillonnées suivant la méthode de Sobol puis 500 simulations sont effectuées à chaque \mathcal{MV}_i . Les moments d'ordre 1 et 2 (moyenne et écart type) que peut prendre la consommation calorifique annuelle sont présentés à la table 4.2. Lors-qu'aucune mesure n'a été effectuée pour évaluer la performance, le prestataire estimera que la consommation sera d'environ 26 MWh et garantira une consommation annuelle de 45.46 MWh à 5% de risque. La consommation calorifique réelle du bâtiment est inférieure à la consommation garantie au plan \mathcal{MV}_0 mais la dispersion des valeurs que peuvent prendre cette consommation reste trop importante pour être acceptable.

Le premier constat qu'on peut faire à la table 4.2, c'est que plus on va loin dans la chaîne de mesure, plus l'incertitude sur la consommation de chauffage décroît. Elle a chuté de plus de 6 MWh du plan de mesure \mathcal{MV}_0 où aucune variable n'est mesurée, au plan de

mesure complet \mathcal{MV}_{11} .

TABLE 4.2 – Évolution des moments d'ordre 1 et 2 lors de la propagation d'incertitudes sur le modèle Trnsys a priori

\mathcal{MV}_i	Variables mesurées	Moyenne (MWh)	Écart type (MWh)
\mathcal{MV}_0	-	26.446	9.135
\mathcal{MV}_1	9	26.340	6.175
\mathcal{MV}_2	9-1	26.245	3.928
\mathcal{MV}_3	9-1-8	24.484	3.573
\mathcal{MV}_4	9-1-8-11	24.506	3.132
\mathcal{MV}_5	9-1-8-11-6	24.507	3.086
\mathcal{MV}_6	9-1-8-11-6-3	24.456	3.081
\mathcal{MV}_7	9-1-8-11-6-3-4	24.271	3.044
\mathcal{MV}_8	9-1-8-11-6-3-4-2	24.271	3.043
\mathcal{MV}_9	9-1-8-11-6-3-4-2-7	24.271	3.034
\mathcal{MV}_{10}	9-1-8-11-6-3-4-2-7-10	24.267	2.98
\mathcal{MV}_{11}	9-1-8-11-6-3-4-2-7-10-5	24.136	2.969

Cependant, comme expliqué dans la méthodologie, on aimerait ne pas avoir à mesurer toutes les variables pour atteindre ce niveau d'incertitude. Les incertitudes ont dans un premier temps été maximisé à 5% de risque (Chapitre 2, équation 2.10, page 39). Les résultats sont présentés à la table 4.3. On constate qu'à partir du plan \mathcal{MV}_5 l'incertitude sur la consommation calorifique annuelle se stabilise (figure 4.7 (a)). Le gain en précision est presque nulle sur les plans de mesure suivants. La mesure des variables à partir de ce plan de mesure ne réduit plus l'incertitude sur l'évaluation de la performance (figure 4.4 (b)) mais les coûts des moyens mis à disposition augmenteront au fur et à mesure que l'on évoluera dans la chaîne de mesure.

Afin d'identifier le plan de mesure optimal (étape 5 de la méthodologie), nous avons estimé le pourcentage que représentait l'apport de précision d'un \mathcal{MV}_i à un \mathcal{MV}_{i+1} (table 4.3). En partant sur le principe d'un Pareto (85/15), nous avons pu identifier le plan \mathcal{MV}_2 comme étant le plan optimal. Il permet en mesurant deux variables à savoir la température intérieure et extérieure de réduire d'environ 85% l'incertitude initiale. À la figure 4.7 sont représentés les pourcentages de l'apport de la mesure à l'évaluation de la performance. Nous choisissons donc pour continuer le plan de mesure \mathcal{MV}_2 comme étant le plus adapté pour garantir l'équilibre entre coût et précision. L'opérateur pourrait bien évidemment partir sur un autre critère de sélection et souhaiter réduire l'incertitude de 90%. Son plan de mesure optimal serait alors le plan \mathcal{MV}_3 . Il rajouterait alors à son protocole de mesure, la mesure de la perméabilité à l'air. D'autant plus que les coûts de mesure de la perméabilité d'un bâtiment restent très faibles et que cette dernière est parfois effectuée d'office à la livraison de certains bâtiments.

Les consommations calorifiques annuelles estimées avant et après la mesure de la température intérieure et extérieure sont présentées à la figure 4.7 (a). La consommation réelle est supérieure à la consommation garantie à 5% de risque. L'opérateur s'engage sur une consommation annuelle de 34.36 MWh à cette étape n'ayant pas encore pris en considération les incertitudes sur les paramètres statiques et l'ajustement de la performance calculée.

TABLE 4.3 – Encadrement des écarts type de chaque \mathcal{MV}_i par Khi deux puis normalisation

\mathcal{MV}_i	Écart type encadré (MWh)	Gain en précision (MWh)	Normalisation
\mathcal{MV}_0	9.629	-	0.000
\mathcal{MV}_1	6.509	3.120	48.01
\mathcal{MV}_2	4.140	2.368	84.45
\mathcal{MV}_3	3.766	0.374	90.20
\mathcal{MV}_4	3.301	0.465	97.36
\mathcal{MV}_5	3.253	0.048	98.10
\mathcal{MV}_6	3.247	0.005	98.18
\mathcal{MV}_7	3.208	0.039	98.78
\mathcal{MV}_8	3.207	0.001	98.80
\mathcal{MV}_9	3.198	0.009	98.95
\mathcal{MV}_{10}	3.142	0.056	99.81
\mathcal{MV}_{11}	3.129	0.0013	100.0

4.3.3 Intégration des incertitudes des paramètres statiques

L'étape 6 de la méthodologie d'optimisation consiste à intégrer les incertitudes des paramètres statiques à l'étude. Ils sont représentatifs de l'enveloppe du bâtiment, de la modélisation structurelle qui a été faite, mais aussi des caractéristiques du site où est situé le bâtiment test. Nous proposons donc d'intégrer à l'étude des 11 variables précédentes, 28 paramètres statiques qui représentent :

1. Les propriétés physiques telles que la conductivité et la densité des matériaux constituant l'enveloppe ;
2. Les propriétés structurelles telles que l'épaisseur de ces matériaux ;
3. Les ponts thermiques ;
4. La résistance thermique des vitrages ;
5. Les coefficients de convection des surfaces intérieures et extérieures ;
6. L'albédo du site de l'étude.

Les caractéristiques de chaque variable sont présentées en Annexes à la table 6.5. Les incertitudes ont été estimées sur la base de données provenant de retours d'expérience de

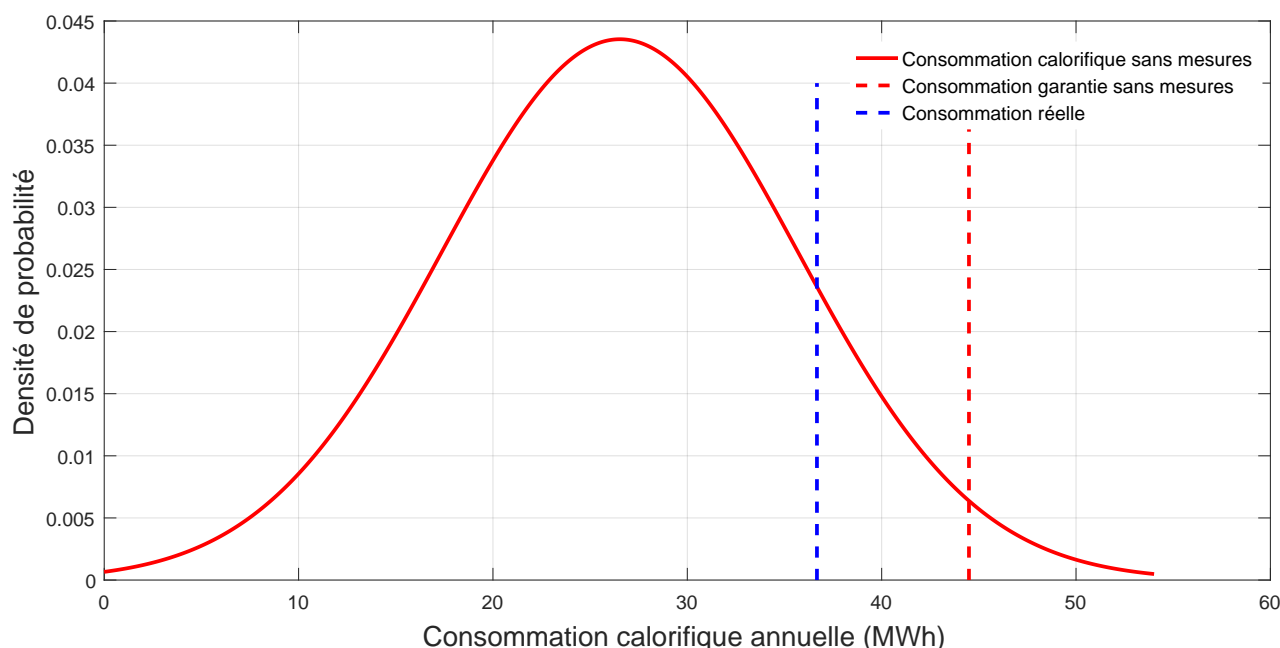


FIGURE 4.4 – Consommation calorifique annuelle estimée suite à la propagation d’incertitudes sur le modèle a priori, au plan de mesure \mathcal{MV}_0 ainsi que la consommation garantie à 5 % de risque

bâtiments tertiaires. Les valeurs nominales sont celles fournies par les données techniques architecturales et qui ont aussi servi lors de la construction du modèle de simulation Trnsys. 500 simulations de Monte Carlo ont été effectuées. Les résultats sont présentés à la figure 4.7 (b). On obtient une incertitude de 5.4 MWh contrairement à l’étape précédente où l’incertitude était de 3.93 MWh. La consommation réelle est maintenant inférieure à la consommation garantie à 5% de risque mais reste supérieure à la performance estimée par l’opérateur. Pour obtenir la performance estimée que fournira l’opérateur à son client, les valeurs nominales des variables mesurées doivent être ajustées à la prochaine étape.

4.3.4 Ajustement de la performance

À cette étape, il s’agit d’adapter l’évaluation de la performance aux conditions réelles d’exploitation pour que les résultats soient en cohérence avec l’année d’étude. Pour ce faire, les variables qui ont été choisies comme variables à mesurer dans le plan de mesure choisi, sont modifiées dans le modèle de simulation. C’est à dire qu’à l’étape précédente leurs valeurs nominales étaient construites sur les connaissances antérieures et des incertitudes de mesure leur étaient assignées. Les données et les incertitudes de mesure in situ leur sont maintenant assignées pour effectuer la propagation d’incertitudes. Les résultats sont présentés à la figure 4.8.

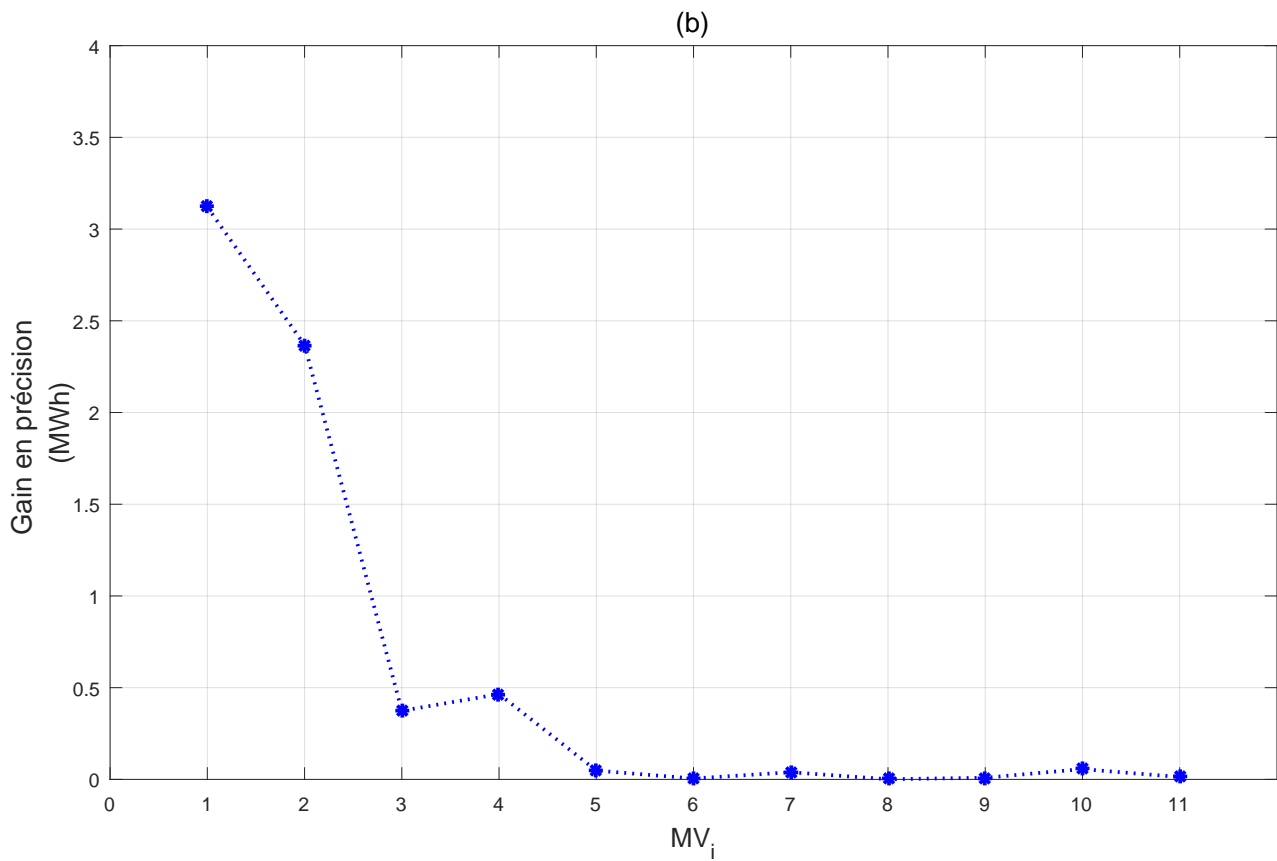
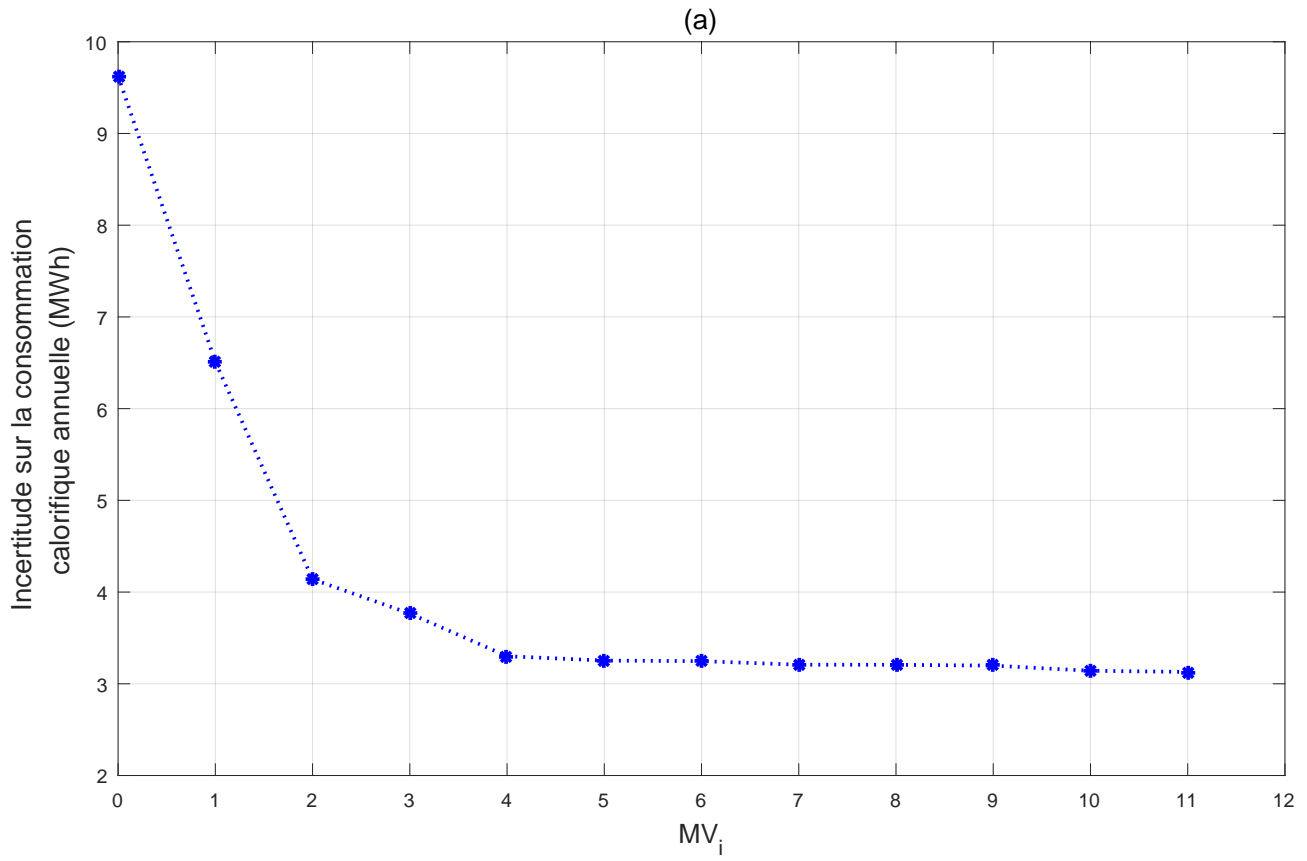


FIGURE 4.5 – (a) Incertitude (MWh) sur la consommation calorifique annuelle à 5 % de risque suivant les valeur de la table 4.3. (b) Évolution du gain en précision (MWh) d'un plan MV_i à un plan MV_{i+1} .

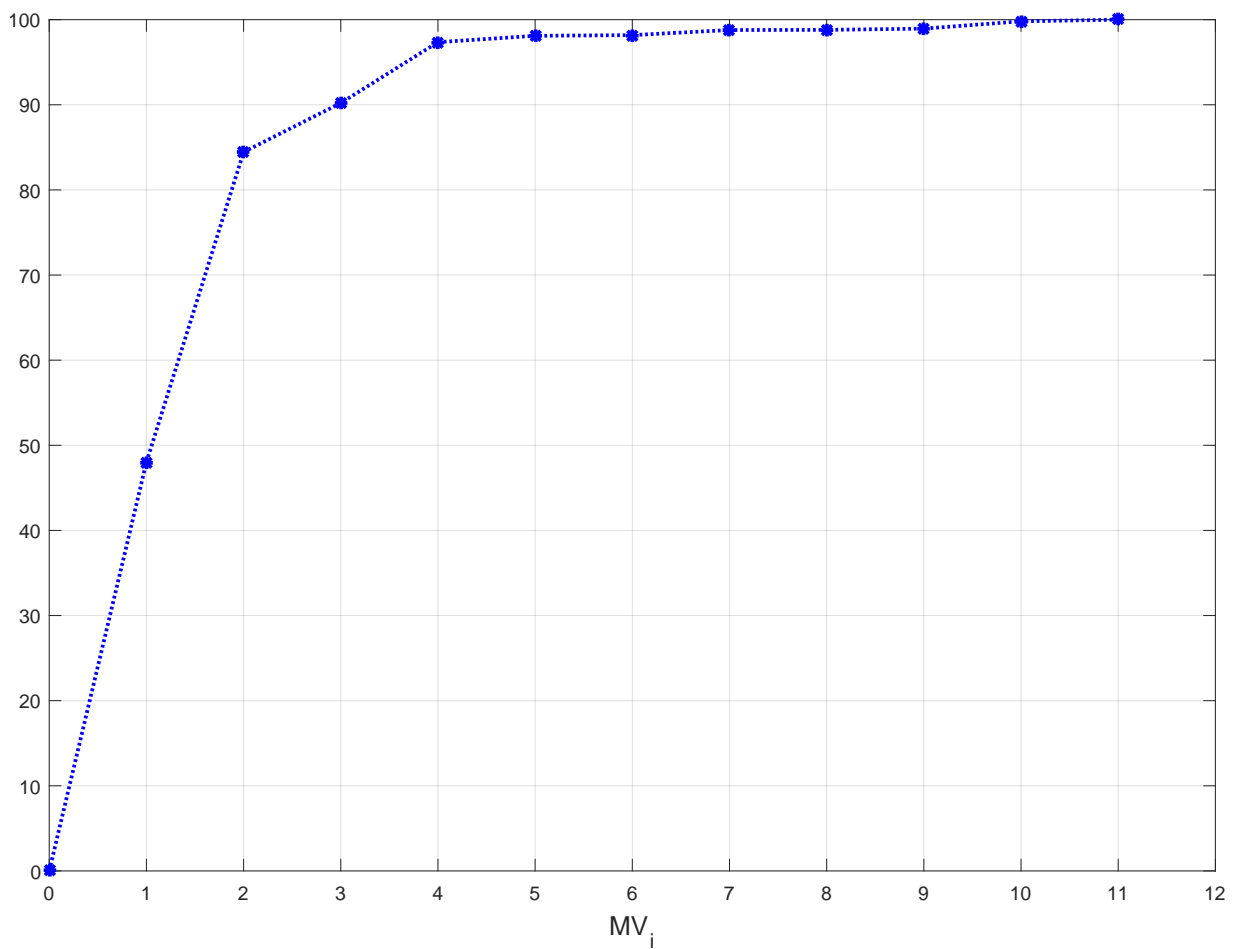


FIGURE 4.6 – Apport de la mesure (%) sur l'incertitude liée à l'évaluation de la performance

TABLE 4.4 – Évolution de la moyenne, de l'écart type et de la consommation garantie au fil de la procédure d'optimisation

	Moyenne (MWh)	Écart type (MWh)	Consommation garantie à 5 % de risque (MWh)
Non mesures	26.529	9.659	45.462
Plan de mesure optimal	26.245	4.140	34.359
Plan de mesure optimal + Paramètres statiques	30.219	5.692	41.376
Plan de mesure optimal + Paramètres statiques + Ajustement	38.768	7.233	52.944

La consommation moyenne ajustée est de 38.77 MWh. Elle est plus élevée que celle évaluée précédemment (table 4.4). Nous avons constaté que les températures intérieures pratiquées dans le bâtiment test étaient en moyenne nettement plus élevées que celles tirées des enseignements PREBAT (PREBAT, 2015), ce qui justifie cette augmentation.

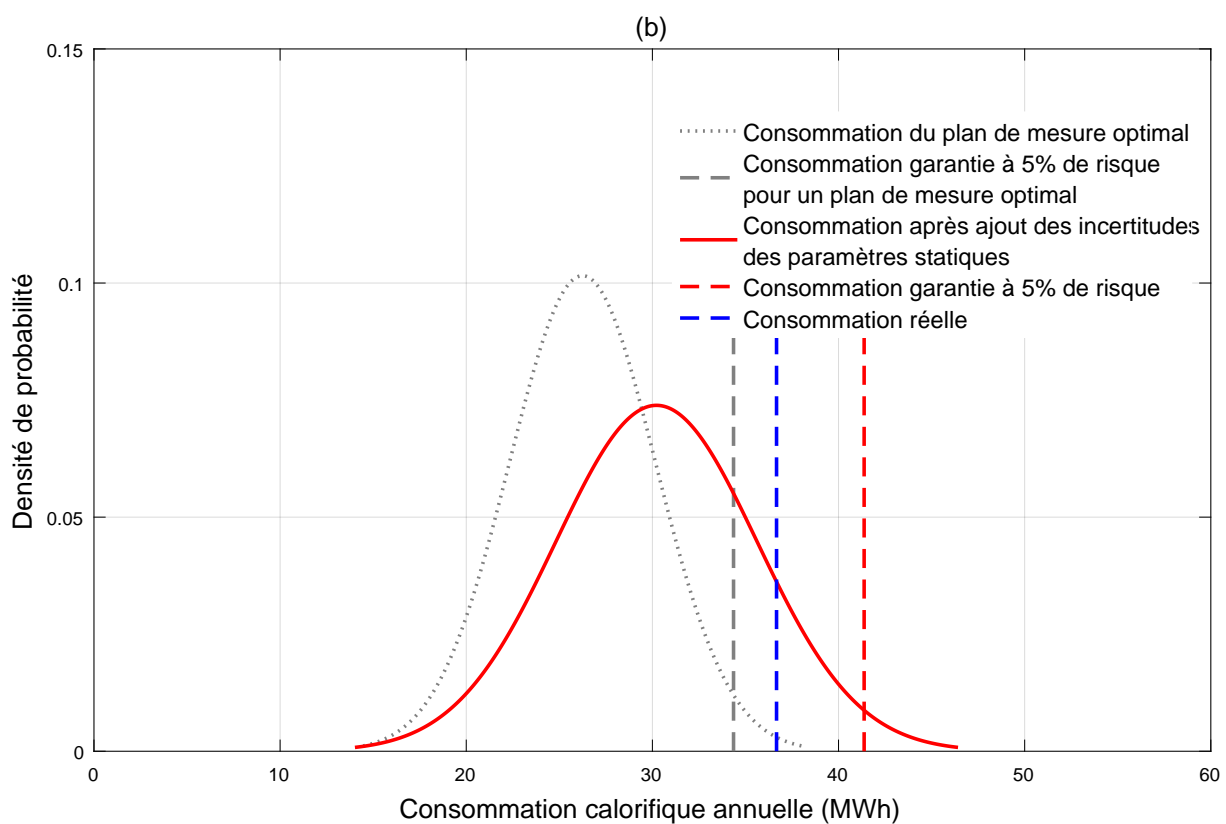
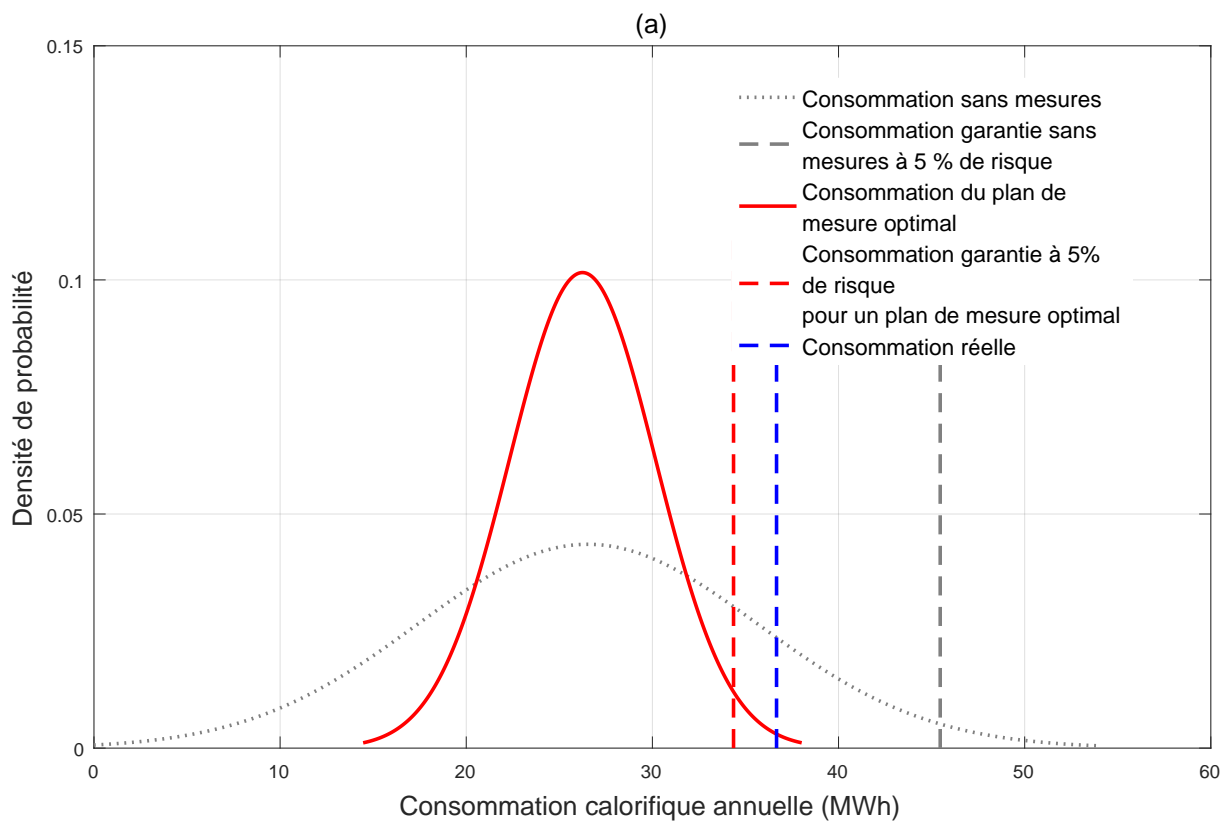


FIGURE 4.7 – (a) Comparaison de la consommation calorifique annuelle estimée suite à la propagation d'incertitudes sur le modèle a priori correspondant au plan de mesure \mathcal{MV}_0 et de celle du plan de mesure optimal \mathcal{MV}_2 ainsi que la consommation garantie à 5 % de risque. (b) Évolution de la consommation calorifique annuelle estimée suite à la propagation d'incertitudes et à l'intégration des paramètres statiques sur le modèle a priori après l'identification du plan de mesure optimal \mathcal{MV}_2 ainsi que la consommation garantie à 5% de risque.

Cette moyenne est aussi supérieure à la consommation réelle de seulement 2 MWh. La mesure de la température intérieure et extérieure permet donc d'évaluer la performance avec une incertitude de 7.23 MWh et un risque de 5% sur les résultats finaux.

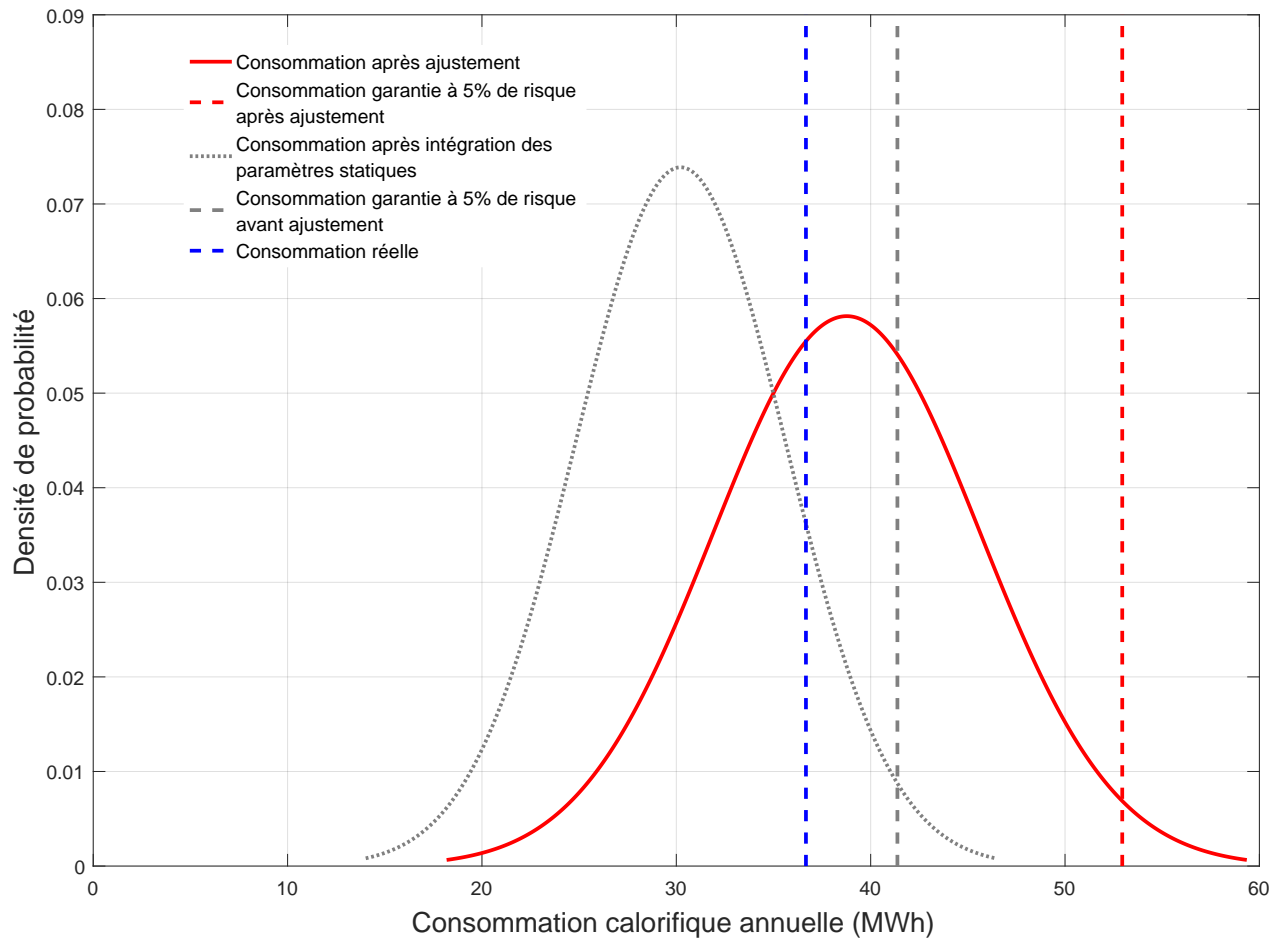


FIGURE 4.8 – Consommation calorifique annuelle estimée suite à l'intégration des paramètres statiques, puis la consommation calorifique annuelle après ajustement ainsi que la consommation garantie à 5 % de risque

4.3.5 La dimension stochastique effective

En somme, les résultats nous permettent de conclure que parmi toutes les variables incertaines considérées dans le calcul probabiliste effectué, seulement deux d'entre elles contribuent réellement à la variabilité de la consommation calorifique annuelle. On peut ainsi distinguer la dimension stochastique effective (ensemble des variables qui contribuent réellement à la variabilité de la réponse du modèle), de la dimension stochastique nominale (ensemble des variables incertaines du modèle). Cette situation est récurrente dans les problèmes d'ingénierie et a notamment été soulevé par Riahi (Riahi, 2013). L'identification de la dimension efficace d'un modèle a pour vocation dans les problèmes à dimension très élevée (nombre de paramètres incertains supérieur à 10^2), de réduire le temps de calcul et

de simplifier les calculs probabilistes en fixant les variables non influentes à leurs valeurs nominales respectives. Cette stratégie est similaire à celle adoptée dans la méthodologie que nous proposons. En identifiant la dimension efficace de la variable d'intérêt de son étude, l'opérateur optimise implicitement son plan de mesure.

4.4 Conclusion du chapitre

L'évaluation et le suivi de la performance énergétique des bâtiments font partie intégrante des contrats de performance énergétique. Elles apportent une garantie au client sur les actions d'amélioration menées et permettent de mesurer leur réelle contribution. En fonction de l'envergure du projet, l'installation d'une instrumentation complète peut se révéler coûteuse et le temps d'investigation très long lorsqu'il s'agit d'évaluer la consommation calorifique annuelle. Or, la réduction des moyens mis en œuvre entraîne une augmentation naturelle des incertitudes. Nous avons proposé dans ce chapitre une méthodologie permettant d'opter pour un plan de mesure adapté. Elle repose en grande partie sur l'estimation des incertitudes des variables potentiellement mesurables et sur la quantification de leur impact éventuel sur la performance finale. Elle favorise le recours aux outils basiques d'évaluation tels que la construction d'un modèle de simulation ainsi que l'utilisation des méthodes d'analyse de sensibilité et de propagation d'incertitudes. Ces outils restent à la portée de prestataires moyennement qualifiés et peuvent être a priori pris en main très rapidement. Communément utilisés dans la communauté du bâtiment, elles sont de ce fait tout aussi adaptés dans une démarche de suivi et d'évaluation.

Conclusion et Perspectives

L'évaluation de la performance énergétique des bâtiments requière la mise en place de plans de mesure et de vérification. Pour ce faire, les opérateurs se basent généralement sur les quatre options déclinées par l'IPMVP. Ces options vont de la mesure de toutes les grandeurs clés, sur des périodes courtes ou longues, à leurs estimations sans toute fois connaître leur impact sur les résultats d'évaluation de la performance. L'objectif de cette thèse a été de proposer une méthodologie d'optimisation d'un plan de mesure et de vérification en phase d'exploitation du bâtiment. Cette méthodologie intègre en elle, les incertitudes des variables d'entrée et permet à l'opérateur, d'identifier le plan de mesure adapté en choisissant en amont le niveau d'incertitudes visé sur la variable d'intérêt de son choix ainsi que la précision qu'il souhaite donner à ses résultats de mesure.

L'accent a dans un premier temps été porté sur la quantification des incertitudes des variables d'entrée. Nous avons proposé l'évaluation de ces incertitudes suivant deux cas de figure : le premier cas où des mesures complètes sont effectuées pour estimer la grandeur, et le deuxième cas où la grandeur est estimée à partir de valeurs hypothétiques ou au mieux, d'informations provenant de retours d'expérience sur des bâtiments similaires. La considération de ces deux cas de figure présente l'avantage de pouvoir étudier une multitude de possibilité de plans de mesure. Ainsi, pour prendre en compte la diversité des grandeurs qui sont susceptibles d'avoir un effet sur la performance énergétique, le panel des variables qui a été étudié, est constitué de variables dynamiques et de variables statiques, qui ont la possibilité d'être ajusté avant ou après la mise en place d'actions de performance énergétiques mais aussi lors de l'évaluation de la performance du bâtiment dans une démarche de suivi et d'évaluation. Ces variables sont aussi représentatives des sollicitations dont fait l'objet le bâtiment à savoir les sollicitations météorologiques, les sollicitations liées à la distribution aéraulique et celles liées à l'usage.

La deuxième phase de la thèse a consisté à mettre sur pied un modèle de simulation qui intègre en lui la physique du bâtiment et qui permette d'étudier le caractère stochastique des variables d'entrée. Ce modèle a été construit sous Trnsys et nécessitait une dizaine de minutes pour une simulation. Vu le nombre de simulations à effectuer pour les analyses de sensibilité et de propagation d'incertitudes, nous avons proposé de réduire le modèle par une expansion polynomiale, moins coûteuse en temps de calcul et qui permettrait d'obtenir les indices de sensibilité de Sobol sans aucune simulation supplémentaire. La surface de réponse a été obtenue avec une erreur relative de 15% par rapport au modèle initial, qui était trop élevée pour être acceptable. Nous avons constaté que l'expansion polynomiale permettait d'effectuer une analyse en tendance centrale mais pas d'obtenir des indices de Sobol avec cette erreur. Une analyse de sensibilité par screening a aussi été effectuée sur le modèle Trnsys.

Ces indices ont servi de base à la méthodologie d'optimisation du plan de meure. En effectuant un classement des variables par ordre d'importance, nous avons pu identifier celles qui contribuent réellement à la variabilité de la consommation calorifique annuelle. En mesurant seulement 20% des variables, l'incertitude sur l'évaluation de la performance a été réduite de plus de 80%, l'apport de la mesure des variables qui suivent n'étant plus conséquent. En résumé, l'approche que nous proposons est simple et intuitive pour un prestataire qui souhaite évaluer la performance d'un bâtiment donné. Elle est basée sur des outils couramment utilisés dans le domaine du bâtiment à savoir : un outil de simulation thermique dynamique, une méthode d'analyse de sensibilité et une méthode de propagation d'incertitudes.

Ces travaux de thèse s'ouvrent sur de nombreuses perspectives. Dans un premier temps, il serait intéressant de pouvoir caractériser les lois de distributions affectées à chaque variable en fonction de l'instrumentation qui a été effectuée. Des lois normales tronquées aux bornes des intervalles de variation ont été assignées à toutes les variables (sauf la perméabilité à l'air) faute de temps. Des tests d'adéquation de loi de type Kolmogorov Smirnov ou droite de Henry, peuvent être effectués. Des lois de distribution appropriées seront ainsi utilisées et permettraient d'améliorer la précision des incertitudes en sortie de modèle.

Puis, un panel plus élargi des variables d'entrée peut être considéré. On pourrait ajouter aux sollicitations externes, des variables météorologiques supplémentaires telles que l'humidité relative ou la pression. Dans le modèle physico-probabiliste construit, elles ont été considérées mais ont été fixées à leurs valeurs déterministes dans les études menées. Ce qui revient entre autre à supposer qu'elles ont un effet négligeable sur la sortie du modèle,

ce qui dans notre sens peut être vérifié.

Par ailleurs, certaines variables peuvent être agrégées c'est à dire regroupées entre elles afin d'étudier l'impact du groupe de variables sur la sortie du modèle. Les variables météorologiques peuvent par exemple être regroupées entre elles. Puisque, lorsque l'on mesure l'une d'entre elles, on utilise une station météo qui mesure par la même occasion toutes les autres variables météorologiques. On peut donc envisager de ne pas dissocier leur mesure. D'un autre côté, la granularité de certaines variables peut être plus approfondie. La variable occupation peut aussi être décomposé en un nombre d'occupant et en des scénarios d'occupation pour être traité séparément. Dans notre étude nous avons choisi de nous concentrer sur le nombre d'occupant par zone du bâtiment, les scénarios d'occupation étant déterministes.

Pour finir, nous aurions souhaité appliquer cette méthodologie à d'autres cas d'étude pour vérifier son applicabilité dans le cadre de l'évaluation de la performance ou dans le cadre de la mise en place d'actions de performance énergétique. La performance serait ainsi évaluée avant et après travaux pour pouvoir caractériser l'impact de la mesure des variables effectives.

Les travaux de cette thèse ont donné lieu à :

Un article de revue :

Titikpina, F., Caucheteux, A., Charki, A. and Bigaud, D., 2015. *Uncertainty assessment in building energy performance with a simplified model*. International Journal of Metrology and Quality Engineering, 6(3), p.308. (*in Press*).

Cinq communications dans des conférences internationales et nationales :

Titikpina, F., Caucheteux, A., Charki, A. and Bigaud, D., *Measurement and Verification plan optimization for the evaluation of building energy performance*. 15th Conference of International Building Performance Simulation Association, 2017. (*Under review*).

Titikpina, F., Caucheteux, A., Charki, A. and Bigaud, D., *Évaluation de l'incertitude en performance énergétique des bâtiments*. CAFMET 2016, Comité Africain de Métrologie, Dakar, Sénégal, Mars 2016.

Titikpina, F., Caucheteux, A., Charki, A. and Bigaud, D., *Construction of temporal data-based files for building energy simulations : the missing data issue*. Proceedings of BS2015 : 14th Conference of International Building Performance Simulation Association, Hyderabad, India, Dec. 7-9, 2015.

Titikpina, F., Caucheteux, A., Charki, A. and Bigaud, D., *Impact de la mesure de la météo sur l'évaluation de la performance énergétique des bâtiments*. Rencontres Francophones sur la qualité et la mesure, Angers, France, Mai 2015.

Titikpina, F., Caucheteux, A., Charki, A. and Bigaud, D., *Uncertainty Assessment in Building Energy Performance*. World Sustainable Energy Days, Wels, Austria, Février 2015.



6

Annexes

Appendix A

6.1 Données météorologiques utilisées pour l'estimation des incertitudes au niveau bas (aucune mesure)

TABLE 6.1 – Moyenne de température extérieure en °C de la station Beaucouzé à Angers de 2005 à 2015. Les données sont issues de la base de données Météo France.

Année\Mois	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Décembre
2015	6	5,1	8,7	12,9	14,3	18,3	20,6	20,3	15,2	12,4	12,3	10,2
2014	8	NaN	9,6	12,2	13,9	18	20,4	18,1	18,2	15,1	11,1	6,5
2013	5,4	4,5	6,8	10	12,2	16,6	21,9	19,8	17,5	15,2	8,2	6,6
2012	7	2,5	10,3	9,7	15	17,4	18,9	20,4	16,1	13,5	8,7	8
2011	5,6	7,9	8,6	14,5	15,8	17,7	18,1	NaN	NaN	14,1	12	8,5
2010	2,5	5,2	8,1	11,9	13,6	18,7	21,5	19,3	16,4	12,4	7,8	2,1
2009	3,1	5,3	8,3	11,5	15,1	18,1	19,6	20,7	17,4	13,4	10,9	5,1
2008	7,5	7,6	8	10,2	16,1	17,2	19,1	18,8	14,6	11,5	8,8	4,5
2007	8	9,3	8,2	14,7	15,1	17,9	18,2	17,9	14,9	12,1	7,5	5
2006	4	4	7,9	11	14,6	19,7	23,9	18,5	19,3	15,4	9,9	5,9
2005	6,6	4	8,3	11,1	14,5	19,9	20,5	19,2	17,7	16,2	7,5	4

TABLE 6.2 – Moyenne des rayonnements solaires en W/m^2 de la station Beaucouzé à Angers de 2005 à 2015. Les données sont issues de la base de données Météo France.

Année\Mois	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Décembre
2015	59,1	112,3	143,3	209,8	200,5	291,5	233,6	202,8	215,2	128,6	78,9	84,2
2014	54,3	103,7	212,2	176,6	208	261,6	216	177,3	258,6	139,5	78,7	53,1
2013	42,1	80,1	100,7	177,3	150	192,2	242,8	264,7	182,1	103,5	68,3	112,2
2012	75,1	147,9	192,9	155,5	247,9	151,8	201,9	248,4	182,6	88,9	86,5	82,3
2011	73,3	63,2	119,1	276,7	280	205,6	179,8	153,7	51	135,1	90,9	60,3
2010	85,8	104,8	160,4	270,3	202,9	212,4	271,8	184,7	212,9	140,2	61,5	63,6
2009	99,9	110,9	195,5	155,7	200,5	252	207,6	253,8	210,4	141,3	71,6	75,2
2008	71	143,1	103,4	160,7	181,6	221,3	229,7	179,3	181,5	136,9	48,8	107,5
2007	43,8	82,9	147,3	278	136,2	163,2	188,7	177,5	216,4	137,4	92,4	119,6
2006	69,7	47,9	124,6	180	168,6	275,1	284,4	183,8	177,1	109,3	11,3	85,9
2005	71,4	83	135,7	155	216,7	248,3	277	269,1	209,5	137,2	124,8	87,7

TABLE 6.3 – Vitesse de vent en m/s de la station Beaucouzé à Angers de 2005 à 2015.
Les données sont issues de la base de données Météo France.

Année	Vitesse du vent
2015	3,706
2014	3,814
2013	3,611
2012	3,386
2011	3,447
2010	3,214
2009	3,364
2008	3,500
2007	3,461
2006	3,158
2005	3,458

6.2 Table de Khi deux

TABLE 6.4 – Table donnant les valeurs de khi-deux pour n degrés de liberté suivant un niveau de risque associé p .

n/p	0.999	0.995	0.99	0.98	0.95	0.9	0.1	0.05	0.02	0.001
1	1.571e-06	3.927e-05	0.0001571	0.0006285	0.003932	0.01579	2.706	3.841	5.412	10.83
2	0.002001	0.01003	0.0201	0.04041	0.1026	0.2107	4.605	5.991	7.824	13.82
3	0.0243	0.07172	0.148	0.1848	0.3518	0.5844	6.251	7.815	9.837	16.27
4	0.0908	0.207	0.2971	0.4294	0.7107	1.064	7.779	9.488	11.67	18.47
5	0.2102	0.4117	0.5543	0.7519	1.145	1.61	9.236	11.07	13.39	20.52
6	0.3811	0.6757	0.8721	1.134	1.635	2.204	10.64	12.59	15.03	22.46
7	0.5985	0.9893	1.239	1.564	2.167	2.833	12.02	14.07	16.62	24.32
8	0.8571	1.344	1.646	2.032	2.733	3.49	13.36	15.51	18.17	26.12
9	1.152	1.735	2.088	2.532	3.325	4.168	14.68	16.92	19.68	27.88
10	1.479	2.156	2.558	3.059	3.94	4.865	15.99	18.31	21.16	29.59
11	1.834	2.603	3.053	3.609	4.575	5.578	17.28	19.68	22.62	31.26
12	2.214	3.074	3.571	4.178	5.226	6.304	18.55	21.03	24.05	32.91
13	2.617	3.565	4.107	4.765	5.892	7.042	19.81	22.36	25.47	34.53
14	3.041	4.075	4.66	5.368	6.571	7.79	21.06	23.68	26.87	36.12
15	3.483	4.601	5.229	5.985	7.261	8.547	22.31	25	28.26	37.7
16	3.942	5.142	5.812	6.614	7.962	9.312	23.54	26.3	29.63	39.25
17	4.416	5.697	6.408	7.255	8.672	10.09	24.77	27.59	31	40.79
18	4.905	6.265	7.015	7.906	9.39	10.86	25.99	28.87	32.35	42.31
19	5.407	6.844	7.633	8.567	10.12	11.65	27.2	30.14	33.69	43.82
20	5.921	7.434	8.26	9.237	10.85	12.44	28.41	31.41	35.02	45.31
21	6.447	8.034	8.897	9.915	11.59	13.24	29.62	32.67	36.34	46.8
22	6.983	8.643	9.542	10.6	12.34	14.04	30.81	33.92	37.66	48.27
23	7.529	9.26	10.2	11.29	13.09	14.85	32.01	35.17	38.97	49.73
24	8.085	9.886	10.86	11.99	13.85	15.66	33.2	36.42	40.27	51.18
25	8.649	10.52	11.52	12.7	14.61	16.47	34.38	37.65	41.57	52.62
26	9.222	11.16	12.2	13.41	15.38	17.29	35.56	38.89	42.86	54.05
27	9.803	11.81	12.88	14.13	16.15	18.11	36.74	40.11	44.14	55.48
28	10.39	12.46	13.56	14.85	16.93	18.94	37.92	41.34	45.42	56.89
29	10.99	13.12	14.26	15.57	17.71	19.77	39.09	42.56	46.69	58.3
30	11.59	13.79	14.95	16.31	18.49	20.6	40.26	43.77	47.96	59.7

6.3 Incertitudes des paramètres statiques

TABLE 6.5 – Paramètres statiques considérés ainsi que leurs incertitudes

	Paramètre statique	Minimum	Maximum	Ecart type	Distribution
1	Conductivité du polystyrène expansé	0.8	1.2	0.1	Normale
2	Densité du polystyrène expansé	0.8	1.2	0.1	Normale
3	Conductivité de la laine de bois	0.8	1.2	0.1	Normale
4	Densité de la laine de bois	0.8	1.2	0.1	Normale
5	Conductivité de la laine de roche	0.8	1.2	0.1	Normale
6	Densité de la laine de roche	0.8	1.2	0.1	Normale
7	Conductivité du béton lourd	0.8	1.2	0.1	Normale
8	Densité du béton lourd	0.8	1.2	0.1	Normale
9	Conductivité du Sol	0.8	1.2	0.1	Normale
10	Densité du sol	0.8	1.2	0.1	Normale
11	Résistance Couche virtuelle du sol	0.8	1.2	0.1	Normale
12	Résistance des ponts thermiques	0.8	1.2	0.1	Normale
13	Longeur des ponts thermiques	0.95	1.05	0.1	Uniforme
14	Absorption solaire des parois (front)	0.8	1.2	0.1	Normale
15	Absorption solaire des parois (back)	0.8	1.2	0.1	Normale
16	Epaisseur du béton GROUND_FLOOR	0.95	1.05	0.1	Uniforme
17	Epaisseur du polystyrène GROUND_FLOOR	0.95	1.05	0.1	Uniforme
18	Epaisseur sol GROUND_FLOOR	0.95	1.05	0.1	Uniforme
19	Epaisseur laine de bois EXT_WALL	0.95	1.05	0.1	Uniforme
20	Epaisseur polystyrène EXT_ROOF	0.95	1.05	0.1	Uniforme
21	Epaisseur polystyrène EXT_FLOOR	0.95	1.05	0.1	Uniforme
22	Epaisseur béton ADJ_CEILING	0.95	1.05	0.1	Uniforme
23	Epaisseur laine de roche ADJ_WALL	0.9	1.1	0.1	Uniforme
24	Coefficient de transfert convectif	0.8	1.2	0.1	Normale
25	Facteur solaire des parois	-0.2	0.2	0.1	Normale
26	Résistance thermique vitrage	0.8	1.2	0.1	Normale
27	Albedo	0.5	2	0.5	Normale
28	Hauteur des masques solaires	0.8	1.2	0.1	Normale

6.4 Modélisation du bâtiment test

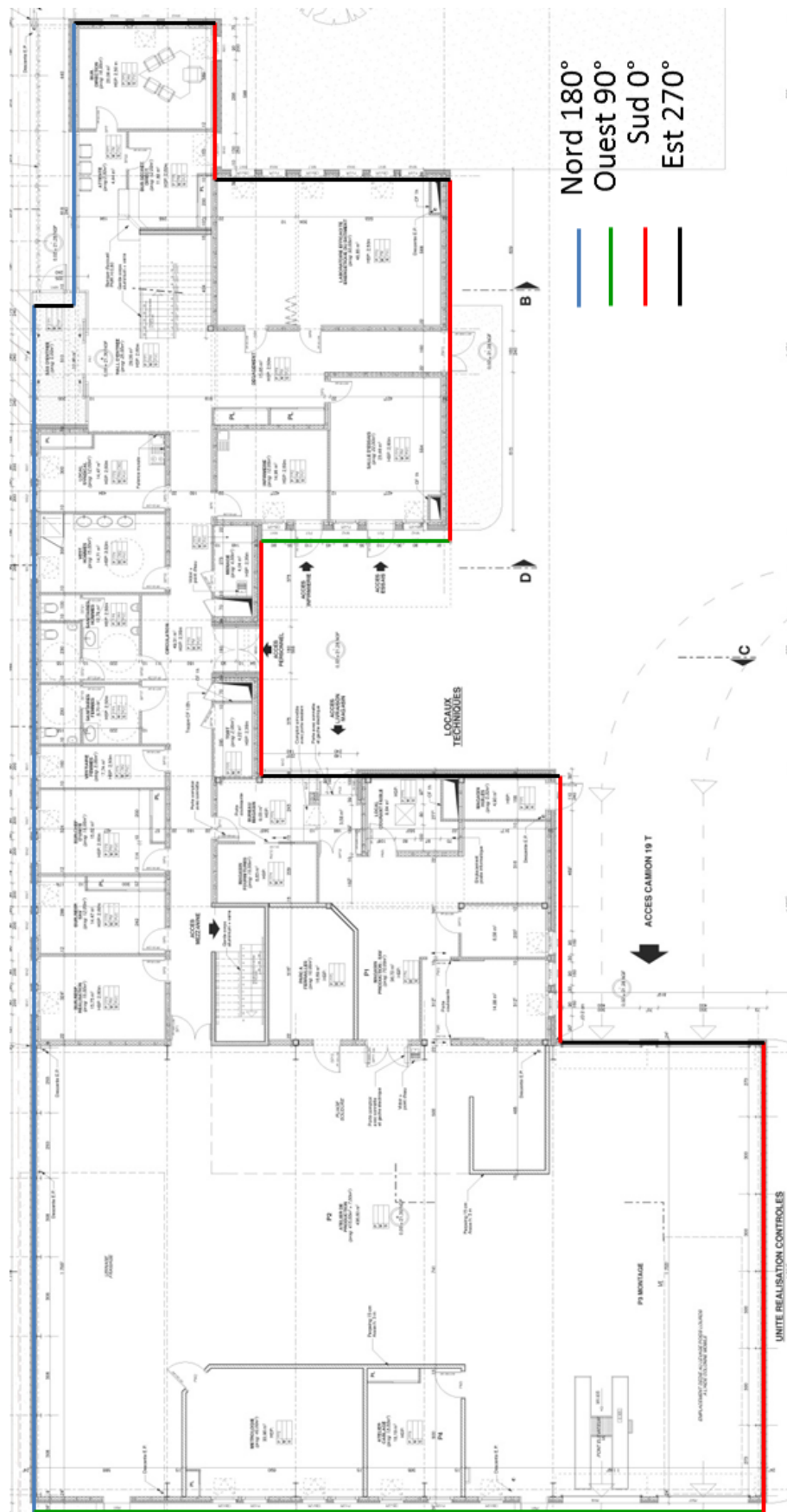


FIGURE 6.1 – Modélisation des orientatons des facades et des masques du bâtiment test



FIGURE 6.2 – Zonage du rez de chaussée du bâtiment test



FIGURE 6.3 – Zonage du premier étage du bâtiment test

6.5 Instrumentation du bâtiment test

6.5.1 Instrumentations in situ

Toute la partie du bâtiment CECP à utilisation bureautique a été instrumentée. Les capteurs ont été placés dans tous les locaux selon les grandeurs à mesurer. Ci-dessous la liste des mesures effectuées ainsi que les plans détaillés des mesures en annexes.

- Température de surface
- Température ambiante
- Mesure de CO₂
- Éclairage
- Ouverture des fenêtres et des stores
- Entrée et sortie des occupants
- Température de soufflage
- Différentiel de pression
- Énergie électrique
- Énergie calorifique

La température d'air ambiante a été mesurée dans tous les locaux et des mesures de températures de surface ont été effectuées dans certains locaux en complément. La température d'air en sortie de bouche des deux centrales de traitements d'air ont aussi été mesurés. Pour estimer l'occupation, des compteurs de présence et de mesure de CO₂ ont été placés ainsi que des compteurs d'énergie électrique et de mesures de taux d'ouverture de fenêtres et de stores.

6.5.2 Instrumentation de la CTA Bureaux

La centrale de traitement d'air de la partie Bureaux fonctionne suivant trois régimes pendant les deux saisons Hiver/Été. A chaque régime est associé un mode de fonctionnement suivant une plage horaire et suivant l'occupation des locaux. La température de confort de chauffage est réglée à 20°C sur la CTA.

TABLE 6.6 – Modélisation de chaque zone du bâtiment test et consignes de température du modèle a priori

Zone	Description	Surface [m ²]	Volumel[m ³]	Nb Occupant	Consigne de température [°C]		Ventilation[m ³ /h]	
					Min.	Max.	Soufflé	Repris
Z ₁	BUR009	16,88	54,6912	1	17,4	20,4	30	30
Z ₂	BUR010	15,4	49,896	1	17,4	20,4	30	30
Z ₃	BUR011	19,55	63,342	1	17,4	20,4	30	30
Z ₄	VEST_FEM	8,41	27,2484	0	17,4	20,4	30	30
Z ₅	SANIT_FEM_RDC	9,36	30,3264	0	17,4	20,4	-	60
Z ₆	SAS	10,91	35,3484	0	17,4	20,4	-	-
Z ₇	BUR001	20,19	65,4156	1	17,4	20,4	30	30
Z ₈	SAL REUN	50,78	164,5272	2	17,4	20,4	150	150
Z ₉	SAL ESSAIS	24,46	79,2504	1	17,4	20,4	60	60
Z ₁₀	BUR006	16,07	52,0668	0	17,4	20,4	30	30
Z ₁₁	MENAG	6,04	19,5696	0	17,4	20,4	-	45
Z ₁₂	TGBT	6,04	19,5696	0	17,4	20,4	-	30
Z ₁₃	BUR MAGASIN	10,4	33,696	1	17,4	20,4	30	30
Z ₁₄	MAG FOURN	9,68	31,3632	1	17,4	20,4	-	-
Z ₁₅	MAG PROD SAV	101,36	328,4064	1	17,4	20,4	-	-
Z ₁₆	ATEL PROD	497,02	3454,289	1	17,4	20,4	-	-
Z ₁₇	BUR102	14,16	52,5336	1	17,4	20,4	30	30
Z ₁₈	BUR103	22,27	82,6217	1	17,4	20,4	60	60
Z ₁₉	BUR104	22,28	82,6588	1	17,4	20,4	60	60
Z ₂₀	BUR107	16,06	59,5826	1	17,4	20,4	30	30
Z ₂₁	BUR108	20,11	74,6081	1	17,4	20,4	130	65
Z ₂₂	BUR109	55,012	204,09452	4	17,4	20,4	150	150
Z ₂₃	BUR110	28,98	107,5158	1	17,4	20,4	30	30
Z ₂₄	BUR111	16,21	60,1391	1	17,4	20,4	30	30
Z ₂₅	ARCHIVES	49,48	183,5708	0	17,4	20,4	180	180
Z ₂₆	BUR112	15,82	58,6922	1	17,4	20,4	30	30

Z ₂₇	BUR113	19,49	72,3079	1	17,4	20,4	30	30
Z ₂₈	SANIT HOM R1	6,04	22,4084	1	17,4	20,4	-	60
Z ₂₉	SANIT FEM R1	6,04	22,4084	0	17,4	20,4	-	60
Z ₃₀	REPRO	22,67	84,1057	1	17,4	20,4	-	60
Z ₃₁	SOUS STATION	31,83	118,0893	0	17,4	20,4	-	-
Z ₃₂	VIDE MAGASIN	15,38	57,0598	0	17,4	20,4	-	-
Z ₃₃	BUR101	24,3	90,153	2	17,4	20,4	60	60
Z ₃₄	CAFET	16,06	52,0344	2	17,4	20,4	120	120
Z ₃₅	CIRC RDC	140,09	453,8916	1	17,4	20,4	240	95
Z ₃₆	SANIT HOM RDC	13,63	44,1612	1	17,4	20,4	-	90
Z ₃₇	VEST HOM	15,57	50,4468	1	17,4	20,4	60	60
Z ₃₈	BUR 105	12,59	46,7089	1	17,4	20,4	30	30
Z ₃₉	BUR106	15,57	57,7647	1	17,4	20,4	30	30
Z ₄₀	CIRC R1	101,78	377,6038	1	17,4	20,4	195	-
Total		1523,97	7024,16732	36	-	-	1915	1915

TABLE 6.7 – Régimes de fonctionnement CTA Bureaux

Fonctionnement	Régime 1	Régime 2	Régime 3	Arrêt
Air insufflé	40%	65%	90%	0
Air extrait	40%	65%	90%	0

TABLE 6.8 – Plages horaires des régimes de fonctionnement de la CTA Bureaux

Plages horaires	Heures	Hiver	Été
Du lundi au vendredi	6H - 9H	Régime 3	Régime 3
	9H - 18H	Régime 3	Régime 2
	18H - 00H	Arrêt	Arrêt
	00H - 6H	Arrêt	Régime 3
Week-ends et jours fériés	7H - 18H	Régime 2	Arrêt
	18H - 7H	Arrêt	Régime 3



FIGURE 6.4 – Instrumentation de la CTA Bureaux

6.6 Consolidation et réduction des modèles

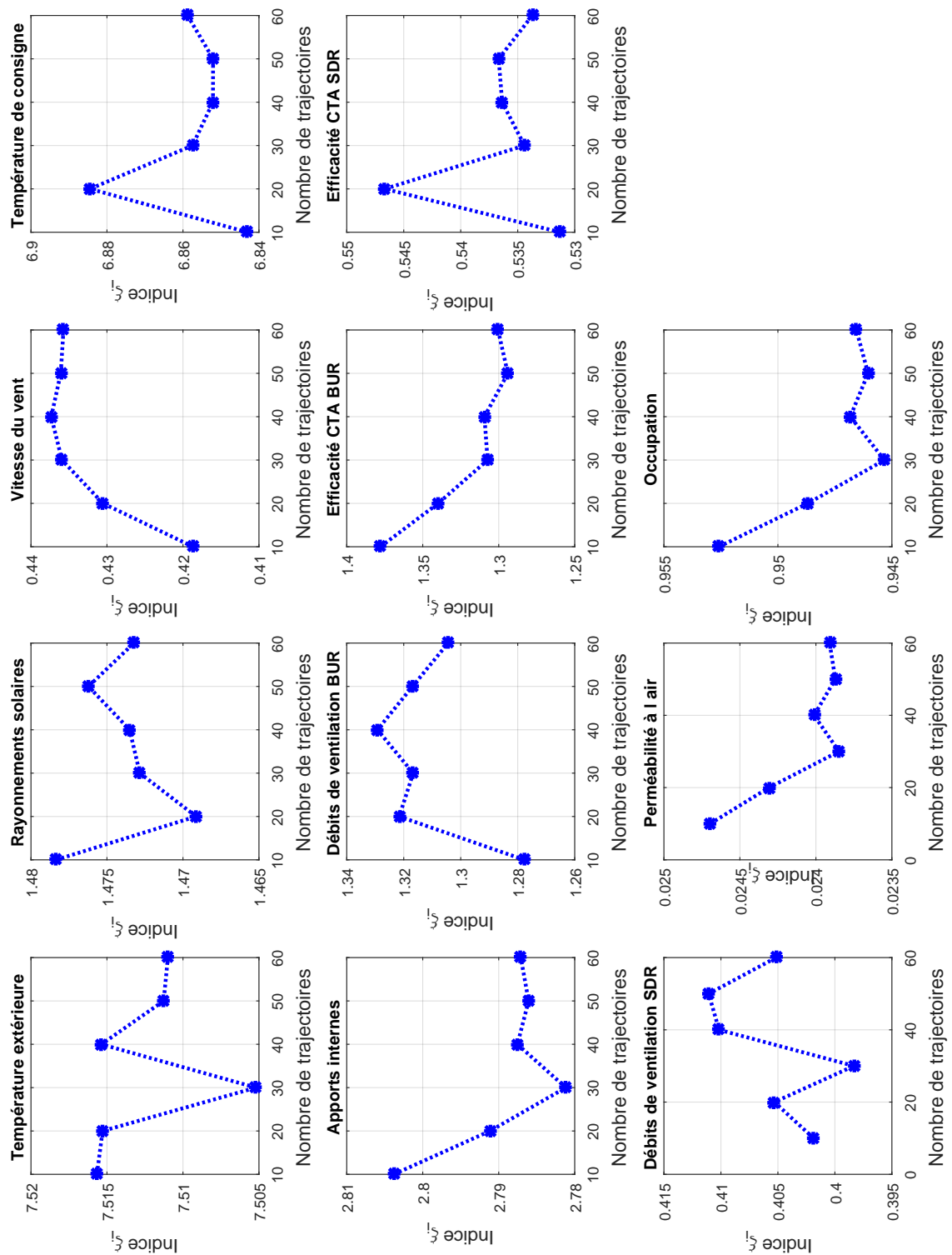


FIGURE 6.5 – Indice ξ_i^* de chaque variable X_i sur la consommation calorifique annuelle (MWh) suivant r trajectoires considérées dans le cas où des mesures ont été effectuées (niveau haut).

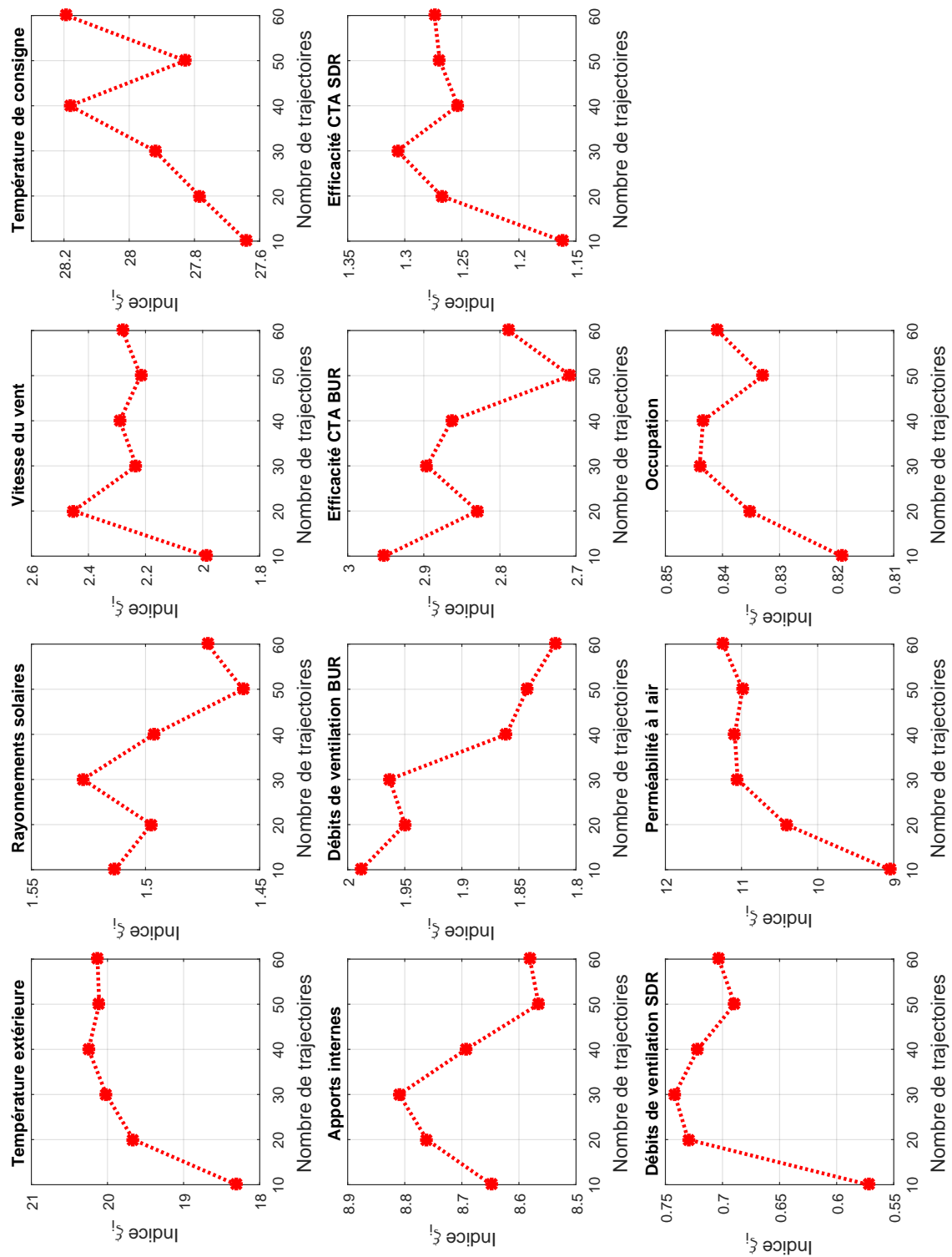


FIGURE 6.6 – Indice ξ_i de chaque variable X_i sur la consommation calorifique annuelle (MWh) suivant r trajectoires considérées dans le cas où aucune mesure n'a été effectuée (niveau bas).

Acronymes

- ANN** Artificial Neural Network. 21, 79
- APE** Action de Performance Énergétique. 22, 23, 26, 40, 120
- AQC** Agence Qualité Construction. 17
- ASHRAE** American Society of Heating, Refrigerating and Air-Conditioning Engineers. 21, 24, 25, 103, 105
- BBC** Bâtiment Basse Consommation. 14
- CECP** Centre d'Études et de Conception de Prototypes du Cerema. 59, 94
- Cerema** Centre d'études et d'expertises sur les risques, l'environnement, la mobilité et l'aménagement. 94
- COP** United nations conference on climate change. 14
- CPE** Contrat de Performance Énergétique. 23
- CSTB** Centre Scientifique et Technique du Bâtiment. 15
- CTA** Centrale de Traitement d'Air. 51, 55–58, 99, 102
- CVC** Chauffage Ventilation Climatisation. 20, 88
- CVRMSE** Coefficient of Variation of Root Mean Square Error. 103, 105
- GIEC** Groupe d'experts Intergouvernemental sur l'Évolution du Climat. 14
- GP** Processus Gaussien. 79
- GPE** Garantie de Performance Énergétique. 17
- GPEI** Garantie de Performance Énergétique Intrinsèque. 18
- GRE** Garantie de Résultats Énergétiques. 18
- GUM** Guide of Uncertainty in Measurement. 35, 56, 58, 64
- IBPSA** International Building Performance Simulation Association. 19

- IEA** Agence Internationale de l'Énergie. 13
- IPMVP** International Performance Measurement and Verification Protocol. 22, 23, 25, 27, 31, 103, 105, 119
- MV** Mesure et Vérification. 23, 31, 119
- NF** Norme Française. 98
- NMBE** Normalized Mean Bias Error. 103, 105
- OAT** One At a Time. 72–75
- OMEGA** Outils Méthodologiques pour la Garantie de performance énergétique. 29
- POPE** Programme fixant les Orientations de la Politique Énergétique. 14
- PREBAT** Plate-forme de recherche et d'expérimentation sur l'énergie dans le bâtiment.
59
- RT** Réglementation Thermique. 58
- SIMUREX** Simulations et Retours d'expériences. 19
- STD** Simulation Thermique Dynamique. 45, 47, 60, 94, 96, 102, 105
- SVD** Décomposition en Valeurs Singulières. 84
- VMC** Ventilation Mécanique Contrôlée. 51, 99

Liste des tableaux

1.1	Les quatre options de plan de mesure et de vérification déclinées par l'IPMVP	24
2.1	Exemple de facteur d'élargissement et probabilités associées pour une variable gaussienne	40
2.2	Quantification des incertitudes des grandeurs d'entrée suivant les différents niveaux d'étude	44
2.3	Illustration des périodes de mesures effectuées aux différents emplacements des stations météo	47
2.4	Différence, pour chaque grandeur (température extérieure, rayonnements solaires et vitesse du vent) entre l'emplacement de la station mobile et la station de référence	48
2.5	Évaluation de l'incertitude liée à la mesure de la température extérieure, des rayonnements solaires et de la vitesse du vent à 95% de confiance	48
2.6	Moyenne de chaque variable météorologique (température extérieure, rayonnements solaires et vitesse du vent) pour chaque année comprise entre les 2005 à 2015 et évaluation de l'incertitude liée aux données de température extérieure de 2005 à 2015 à 95% de confiance et pour 10 degrés de libertés .	51
2.7	Débits nominales et débits mesurés aux bouches de ventilation du bâtiment test (m^3/h)	53
2.8	Incertaince liée à la mesure de l'efficacité de l'échangeur	56
2.9	Mesure de la perméabilité à l'air du bâtiment CECP	59
2.10	Exemple d'estimateurs pour l'approximation du modèle d'occupation	63
2.11	Estimation des incertitudes de mesures et de non mesure liés à chaque grandeur considéré	67
2.12	Annotation assignée à chaque variable d'entrée	69
3.1	Type de polynômes en fonction des distributions des variables d'entrée	81
3.2	Conductivité des parois du bâtiment CECP	97
3.3	Différents ponts thermiques représentés lors de la modélisation du bâtiment test	98
3.4	Hypothèses de calibration en entrée du modèle	101

3.5	Consommation calorifique annuelle réelle et estimée par le modèle de simulation	102
3.6	Indicateurs de performance pour le calibrage d'un modèle de simulation selon un critère horaire et résultats du calibrage du modèle STD	105
3.7	Résultats d'analyse de sensibilité par la méthode de Morris : distance ξ_i estimée suivant l'équation 3.6 pour 40 trajectoires effectuées au niveau haut et au niveau bas d'incertitude.	107
3.8	Taille du plan d'expériences en fonction du degré polynomial considéré pour 11 variables en considérant toutes les interactions	112
3.9	Taille du plan d'expériences en fonction du degré polynomial considéré pour 11 variables dans le cas d'un chaos polynomial tronqué aux interactions d'ordre 2	112
3.10	Comparaison des caractéristiques du modèle initial et de celles du méta-modèle aux points auxquels le méta-modèle a été construit (MWh) suivant les équations 3.24 et 3.25.	113
3.11	Indices de Sobol du premier ordre obtenus en réarrangeant les coefficients du méta-modèle sur l'expansion polynomiale de degré 2	113
4.1	Indices de Morris, ξ_i estimés suivant l'équation 3.6 pour 40 trajectoires effectuées au niveau bas d'incertitude.	126
4.2	Évolution des moments d'ordre 1 et 2 lors de la propagation d'incertitudes sur le modèle Trnsys a priori	127
4.3	Encadrement des écarts type de chaque $\mathcal{M}\mathcal{V}_i$ par Khi deux puis normalisation	128
4.4	Évolution de la moyenne, de l'écart type et de la consommation garantie au fil de la procédure d'optimisation	131
6.1	Moyenne de température extérieure en $^{\circ}\text{C}$ de la station Beaucouzé à Angers de 2005 à 2015. Les données sont issues de la base de données Météo France.	142
6.2	Moyenne des rayonnements solaires en W/m^2 de la station Beaucouzé à Angers de 2005 à 2015. Les données sont issues de la base de données Météo France.	143
6.3	Vitesse de vent en m/s de la station Beaucouzé à Angers de 2005 à 2015. Les données sont issues de la base de données Météo France.	144
6.4	Table donnant les valeurs de khi-deux pour n degrés de liberté suivant un niveau de risque associé p	145
6.5	Paramètres statiques considérés ainsi que leurs incertitudes	146
6.6	Modélisation de chaque zone du bâtiment test et consignes de température du modèle a priori	152
6.7	Régimes de fonctionnement CTA Bureaux	154
6.8	Plages horaires des régimes de fonctionnement de la CTA Bureaux	154

Table des figures

1.1	Exemple de performance annuelle garantie par un contrat de performance énergétique	18
1.2	Positionnement de la thèse par rapport aux différents champs d'actions d'amélioration de la performance énergétique dans le bâtiment.	22
1.3	Suivi de la performance énergétique du bâtiment basé sur la méthodologie proposée par l'IPMVP (IPMVP, 2010).	24
1.4	Représentation physico-probabiliste d'un modèle de simulation. Chaque grandeur d'entrée x_i est représentée par son incertitude $u(x_i)$ et sa distribution $g(x_i)$, ce qui permet d'obtenir une incertitude $u(y)$ et une distribution $g(y)$ sur la sortie du modèle.	27
1.5	Exemple de dispersion de deux distributions de performance énergétique garantie par un contrat de performance	29
2.1	Différentes configurations possibles des erreurs aléatoires et systématiques. La ligne en pointillée représente la valeur vraie de la grandeur physique et l'histogramme est construit sur la base des mesures effectuées.	36
2.2	Illustration de l'intervalle élargi avec un niveau de risque associé α	40
2.3	Bâtiment CECF	42
2.4	Illustration des deux niveaux d'étude considérés pour une grandeur d'entrée donnée X_i	43
2.5	Différents points de déplacement des stations météorologiques installées sur site. La station mobile a été déplacée aux différents emplacements notifiés en bleu.	47
2.6	Variation de la température extérieure, des rayonnements solaires et de la vitesse du vent de 2005 à 2015. La ligne en rouge représente la médiane des données. Sont représentées de part et d'autre de cette ligne et dans le rectangle, 25 % des données et les deux autres quartiles sont représentés par des pointillés.	50
2.7	Estimation de l'incertitude de mesure des débits de ventilation	54
2.8	Schéma représentatif du principe de fonctionnement d'une des centrales de traitement d'air double flux présente dans le bâtiment test	55

2.9	Distribution des valeurs de l'efficacité de l'échangeur obtenu par des mesures effectuées pendant les deux régimes de fonctionnement	57
3.1	Représentation des résultats de l'analyse de sensibilité par la méthode de Morris	76
3.2	Illustration d'une propagation d'incertitudes à travers un modèle de simulation	77
3.3	Approximation par la méthode de collocation de la fonction $\sin(x)$. La fonction est bien approximée aux points de collocation mais le méta-modèle construit révèle une mauvaise approximation aux autres endroits de la courbe.	83
3.4	10 points échantillonnés par Hypercube Latin en dimension 2. On remarque sur la figure de gauche, que les points échantillonnés recouvrent bien l'espace des deux variables. À l'opposé sur la figure de droite, les points échantillonnés sont alignés et laissent des régions de l'espace vides donc non représentés.	86
3.5	Échantillonnage de 500 points avec les suites de Halton réalisé en dimension 20 sur les sous espaces de dimensions 1 et 2 à gauche et de dimension 10 et 19 à droite. On remarque que pour des dimensions élevées, des sous espaces restent vides et non représentés.	87
3.6	Échantillonnage de 1000 points avec les suites de Sobol réalisé en dimension 20 sur les sous espaces de dimensions 1 et 2 à gauche et de dimensions 19 et 20 à droite.	88
3.7	Algorithme proposé de construction d'un méta-modèle par polynômes du chaos	94
3.8	Décomposition schématique d'un modèle de simulation thermique dynamique	95
3.9	Illustration de la modélisation thermique d'une zone à trois surfaces	96
3.10	Modélisation du bâtiment du CECP sous Sketchup 3D ainsi que des masques lointains.	96
3.11	Illustration des ponts thermiques pris en compte lors de la modélisation . .	98
3.12	Illustration de l'estimation des hauteurs angulaires d'un masque solaire lointain donné suivant un intervalle de 10 degrés, à chaque azimut	101
3.13	Signatures énergétiques obtenues par régression linéaire des consommations journalières par rapport aux degrés jours	103
3.14	Comparaison des températures intérieures mesurées et simulées ainsi que des consommations calorifiques mesurées et simulées journalières du 1 Octobre 2015 au 1 Mars 2016	104
3.15	Comparaison des indices ξ_i de chaque variable X_i sur la consommation calorifique annuelle (MWh) suivant r trajectoires considérées.	108
3.16	Comparaison des effets μ_i et σ_i de chaque variable X_i sur la consommation calorifique annuelle (MWh) suivant 40 trajectoires considérées.	109
3.17	Comparaison des indices ξ_i de sensibilité de Morris pour l'analyse de sensibilité au niveau haut (mesures complètes) et au niveau bas (non mesures) .	110

3.18	Évolution de l'erreur relative E et de l'erreur en validation croisée E_{LOO} en fonction des différents degrés des polynômes suivant les termes d'interactions d'ordre 2	114
3.19	Comparaison de la consommation calorifique annuelle (MWh) du bâtiment test du modèle initial Trnsys et du méta-modèle au degré 2	115
4.1	Hierarchisation des variables considérées en se référant à leurs indices de sensibilité	123
4.2	Identification du plan de mesure optimal en fonction du seuil de précision fixé	124
4.3	Méthodologie d'optimisation d'un plan de mesure ainsi que les outils nécessaires et utilisés à chaque étape	125
4.4	Consommation calorifique annuelle estimée suite à la propagation d'incertitudes sur le modèle a priori, au plan de mesure \mathcal{MV}_0 ainsi que la consommation garantie à 5 % de risque	129
4.5	(a) Incertitude (MWh) sur la consommation calorifique annuelle à 5 % de risque suivant les valeur de la table 4.3. (b) Évolution du gain en précision (MWh) d'un plan \mathcal{MV}_i à un plan \mathcal{MV}_{i+1}	130
4.6	Apport de la mesure (%) sur l'incertitude liée à l'évaluation de la performance	131
4.7	(a) Comparaison de la consommation calorifique annuelle estimée suite à la propagation d'incertitudes sur le modèle a priori correspondant au plan de mesure \mathcal{MV}_0 et de celle du plan de mesure optimal \mathcal{MV}_2 ainsi que la consommation garantie à 5 % de risque. (b) Évolution de la consommation calorifique annuelle estimée suite à la propagation d'incertitudes et à l'intégration des paramètres statiques sur le modèle a priori après l'identification du plan de mesure optimal \mathcal{MV}_2 ainsi que la consommation garantie à 5% de risque.	132
4.8	Consommation calorifique annuelle estimée suite à l'intégration des paramètres statiques, puis la consommation calorifique annuelle après ajustement ainsi que la consommation garantie à 5 % de risque	133
6.1	Modélisation des orientatons des facades et des masques du batiment test .	148
6.2	Zonage du rez de chaussée du batiment test	149
6.3	Zonage du premier étage du batiment test	150
6.4	Instrumentation de la CTA Bureaux	154
6.5	Indice ξ_i^* de chaque variable X_i sur la consommation calorifique annuelle (MWh) suivant r trajectoires considérées dans le cas où des mesures ont été effectuées (niveau haut).	155
6.6	Indice ξ_i de chaque variable X_i sur la consommation calorifique annuelle (MWh) suivant r trajectoires considérées dans le cas où aucune mesure n'a été effectuée (niveau bas).	156

Bibliographie

- ADEME (2013), Chiffres clés du bâtiment, Technical report, Agence de l'environnement et de la maîtrise de l'énergie. [13](#), [14](#)
- AFNOR (2001), 'Norme NF EN 13829 2001-02 Norme européenne, norme française, Performance énergétique des bâtiments, détermination de la perméabilité à l'air des bâtiments'. [98](#)
- Aksoezen, M., Daniel, M., Hassler, U. and Kohler, N. (2015), 'Building age as an indicator for energy consumption', *Energy and Buildings* **87**, 74–86. [20](#)
- Amayri, M., Arora, A., Ploix, S., Bandhyopadhyay, S., Ngo, Q.-D. and Badarla, V. R. (2016), 'Estimating occupancy in heterogeneous sensor environment', *Energy and Buildings* . [20](#), [62](#)
- Ansanay-Alex, G., Abdelouadoud, Y. and Schetelat, P. (2016), Statistical and stochastic modelling of french households and their energy consuming activities, in '12th REHVA World Congress-CLIMA 2016'. [62](#)
- Anstett-Collin, F., Goffart, J., Mara, T. and Denis-Vidal, L. (2015), 'Sensitivity analysis of complex models : coping with dynamic and static inputs', *Reliability Engineering & System Safety* **134**, 268–275. [19](#), [72](#)
- AQC (2011), Les retours du terrain, les premiers enseignements, numéro spécial batimat 2011, Agence pour la qualité de la construction. [17](#)
- ASHRAE, S. C. (2013), *Ashrae handbook : fundamentals*. [21](#), [24](#), [25](#), [103](#)
- Azar, E. and Menassa, C. C. (2012), 'A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings', *Energy and Buildings* **55**, 841–853. [20](#)
- Benedetti, M., Cesarotti, V., Introna, V. and Serranti, J. (2016), 'Energy consumption control automation using artificial neural networks and adaptive algorithms : Proposal of a new methodology and case study', *Applied Energy* **165**, 60–71. [21](#)
- Berthou, T. (2013), Développement de modèles de bâtiment pour la prévision de charge de climatisation et l'élaboration de stratégies d'optimisation énergétique et d'effacement, PhD thesis, Ecole Nationale Supérieure des Mines de Paris. [21](#)

- BIPM, I., IFCC, I. and IUPAP, O. (2008), Evaluation of measurement data—supplement 1 to the gum : propagation of distributions using a monte carlo method, Technical report, Technical report, International Organization for Standardization (ISO), Geneva. [77](#)
- Björck, A. (1996), *Numerical methods for least squares problems*, Siam. [84](#)
- Blatman, G. (2009), Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis, PhD thesis, Clermont-Ferrand 2. [79](#), [92](#), [93](#)
- Blatman, G., Sudret, B. and Berveiller, M. (2007), ‘Quasi random numbers in stochastic finite element analysis’, *Mécanique et Industries* **8**(3), 289–298. [89](#)
- Bordass, B. et al. (2004), Energy performance of non-domestic buildings : closing the credibility gap, *in* ‘in Proceedings of the 2004 Improving Energy Efficiency of Commercial Buildings Conference’, Citeseer.
- BOUGRAIN (2012), Les résultats économiques et environnementaux des bâtiments performants à travers une revue de la littérature, Technical report, Projet interne CSTB RETEX. 67p. [17](#)
- Burkhart, M. C., Heo, Y. and Zavala, V. M. (2014), ‘Measurement and verification of building systems under uncertain data : A gaussian process modeling approach’, *Energy and Buildings* **75**, 189–198. [21](#), [79](#)
- Burman, E., Mumovic, D. and Kimpian, J. (2014), ‘Towards measurement and verification of energy performance under the framework of the european directive for energy performance of buildings’, *Energy* **77**, 153–163. [21](#)
- Calì, D., Matthes, P., Huchtemann, K., Streblow, R. and Müller, D. (2015), ‘Co 2 based occupancy detection algorithm : Experimental analysis and validation for office and residential buildings’, *Building and Environment* **86**, 39–49. [62](#)
- Campolongo, F., Cariboni, J. and Saltelli, A. (2007), ‘An effective screening design for sensitivity analysis of large models’, *Environmental modelling & software* **22**(10), 1509–1518. [73](#), [74](#)
- Caucheteux, A., Amayri, M. and Ploix, S. (2016), Mesure de l’occupation pour l’évaluation de la performance énergétique des batiments : plan de mesure et incertitude, *in* ‘IBPSA FRANCE’. [62](#)
- Caucheteux, A., Sabar, A. E. and Boucher, V. (2013), ‘Occupancy measurement in building : A litterature review, application on an energy efficiency research demonstrated building’, *International Journal of Metrology and Quality Engineering* **4**(2), 135–144. [17](#), [20](#), [62](#)
- Chang, W.-k. and Hong, T. (2013), Statistical analysis and modeling of occupancy patterns in open-plan offices using measured lighting-switch data, *in* ‘Building Simulation’, Vol. 6, Springer, pp. 23–32. [20](#)

- Chen, V. C., Tsui, K.-L., Barton, R. R. and Meckesheimer, M. (2006), 'A review on design, modeling and applications of computer experiments', *IIE transactions* **38**(4), 273–291. [79](#)
- Chiffres clés de l'énergie, P. (2016), Commissariat général du développement durable, Technical report, Service de l'observation et des statistiques. [13](#)
- Communauté, E. (2002), 'Directive 2002/91/ce du parlement européen et du conseil du 16 décembre 2002 sur la performance énergétique des bâtiments', *Journal officiel des Communautés européennes* . [15](#)
- Crawley, D. B. (2008), 'Estimating the impacts of climate change and urbanization on building performance', *Journal of Building Performance Simulation* **1**(2), 91–115. [19](#)
- Crawley, D. B., Hand, J. W., Kummert, M. and Griffith, B. T. (2008), 'Contrasting the capabilities of building energy performance simulation programs', *Building and environment* **43**(4), 661–673. [95](#)
- Cropp, R. A. and Braddock, R. D. (2002), 'The new morris method : an efficient second-order screening method', *Reliability Engineering & System Safety* **78**(1), 77–83. [74](#)
- CSTB (2008), Méthode de calcul th-c-e ex, Technical report, Centre Scientifique et Technique du Batiment. [15](#)
- CSTB (2012), Méthode de calcul th-bce, Technical report, Centre Scientifique et Technique du Batiment. [15](#)
- Daniel, C. (1973), 'One-at-a-time plans', *Journal of the American statistical association* **68**(342), 353–360. [74](#)
- De Wilde, P. (2014), 'The gap between predicted and measured energy performance of buildings : A framework for investigation', *Automation in Construction* **41**, 40–49. [16](#), [28](#), [33](#)
- de Wilde, P. and Tian, W. (2012), 'Management of thermal performance risks in buildings subject to climate change', *Building and Environment* **55**, 167–177. [19](#), [88](#)
- De Wit, S. and Augenbroe, G. (2002), 'Analysis of uncertainty in building design evaluations and its implications', *Energy and Buildings* **34**(9), 951–958. [33](#)
- Demanuele, C., Tweddell, T. and Davies, M. (2010), Bridging the gap between predicted and actual energy performance in schools, in 'World renewable energy congress XI', UAE Abu Dhabi, pp. 25–30.
- D'Oca, S., Fabi, V., Corgnati, S. P. and Andersen, R. K. (2014), Effect of thermostat and window opening occupant behavior models on energy use in homes, in 'Building Simulation', Vol. 7, Springer, pp. 683–694. [20](#)

- Duarte, C., Van Den Wymelenberg, K. and Rieger, C. (2013), 'Revealing occupancy patterns in an office building through the use of occupancy sensor data', *Energy and buildings* **67**, 587–595. [20](#)
- Duplessis, B. (2008), Mise en œuvre des contrats de performance énergétique pour l'amélioration des installations de production frigorifique, PhD thesis, École Nationale Supérieure des Mines de Paris. [23](#), [120](#)
- Echenagucia, T. M., Capozzoli, A., Cascone, Y. and Sassone, M. (2015), 'The early design stage of a building envelope : Multi-objective search through heating, cooling and lighting energy performance analysis', *Applied Energy* **154**, 577–591. [19](#)
- ENERTECH (2012), Evaluation par mesure des performances énergétiques des 8 bâtiments construits dans le cadre du programme européen concerto – rapport de synthèse,, Technical report, ENERTECH, Grenoble – ZAC de Bonne. 145 p., Félines sur Rimandoule. [17](#)
- Fabi, V., Andersen, R. V., Corgnati, S. and Olesen, B. W. (2012), 'Occupants' window opening behaviour : A literature review of factors influencing occupant behaviour and models', *Building and Environment* **58**, 188–198. [20](#)
- FBE (Octobre 2014), Garantie de performance energetique, rédaction du dispositif de mesure et de verification, taches 5.2, Technical report, Fondation Batiment Energie. [25](#)
- Franco, J. (2008), Exploratory Designs for Computer Experiments of Complex Physical Systems Simulation, Theses, Ecole Nationale Supérieure des Mines de Saint-Etienne. [85](#), [86](#)
- Frey, H. C., Mokhtari, A. and Danish, T. (2003), 'Evaluation of selected sensitivity analysis methods based upon applications to two food safety process risk models', *Dept. of Civil, Construction, and Environmental Eng., North Carolina State Univ., Raleigh, NC* . [72](#)
- Gang, W., Wang, S., Shan, K. and Gao, D. (2015), 'Impacts of cooling load calculation uncertainties on the design optimization of building cooling systems', *Energy and Buildings* **94**, 1–9. [88](#)
- Goffart, J. (2013), Impact de la variabilité des données météorologiques sur une maison basse consommation. Application des analyses de sensibilité pour les entrées temporelles., PhD thesis, Université Grenoble Alpes. [19](#), [33](#)
- Gorse, C., Brooke-Peat, M., Parker, J. and Thomas, F. (2016), Building simulation and models : Closing the performance gap, in 'Building Sustainable Futures', Springer, pp. 209–226. [16](#)
- GPE, R. (2013), La garantie de performance énergétique, rapport GPE Version 5, Plan Batiment Durable, Cabinet Michel Huet, Bellenger, Blandin, APOGEE. Encadrement

- légal du risque de mise en jeu de la garantie decennale. Méthodologie pratique de la garantie de performance énergétique intrinsèque (GPEI). [17](#)
- GUM (1995), *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization, Bureau International des Poids et Mesures and Commission électrotechnique internationale and Organisation internationale de normalisation. [35](#), [45](#), [57](#), [76](#)
- Haldi, F. and Robinson, D. (2011), ‘The impact of occupants’ behaviour on building energy demand’, *Journal of Building Performance Simulation* **4**(4), 323–338. [62](#)
- Hamby, D. (1994), ‘A review of techniques for parameter sensitivity analysis of environmental models’, *Environmental monitoring and assessment* **32**(2), 135–154. [72](#)
- Heiselberg, P., Brohus, H., Hesselholt, A., Rasmussen, H., Seinre, E. and Thomas, S. (2009), ‘Application of sensitivity analysis in design of sustainable buildings’, *Renewable Energy* **34**(9), 2030–2036. [72](#)
- Heo, Y., Graziano, D. J., Zavala, V. M., Dickinson, P., Kamrath, M. and Kirshenbaum, M. (2013), ‘Cost-effective measurement and verification method for determining energy savings under uncertainty’, *ASHRAE Transactions* **119**, EE1. [21](#), [79](#)
- Heo, Y. and Zavala, V. M. (2012), ‘Gaussian process modeling for measurement and verification of building energy savings’, *Energy and Buildings* **53**, 7–18. [21](#), [25](#)
- Hoes, P., Hensen, J., Loomans, M., De Vries, B. and Bourgeois, D. (2009), ‘User behavior in whole building simulation’, *Energy and Buildings* **41**(3), 295–302. [17](#), [62](#)
- Hopfe, C. J. and Hensen, J. L. (2011), ‘Uncertainty analysis in building performance simulation for design support’, *Energy and Buildings* **43**(10), 2798–2805. [19](#), [72](#)
- Huang, P., Huang, G. and Wang, Y. (2015), ‘Hvac system design under peak load prediction uncertainty using multiple-criterion decision making technique’, *Energy and Buildings* **91**, 26–36. [19](#), [88](#)
- IEA (2015), Energy and climate change, world energy outlook special report, Technical report, International Energy Agency. [13](#)
- Iooss, B. and Saltelli, A. (2015), ‘Introduction to sensitivity analysis’. [72](#)
- IPMVP, E. V. (2010), ‘International performance measurement and verification protocol (ipmvp)’, *Concepts and options for determining energy and water savings* **1**. [22](#), [23](#), [24](#), [25](#), [31](#), [103](#), [161](#)
- JCGM (2008), *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*. [34](#), [37](#), [39](#)

- Kashif, A., Ploix, S., Dugdale, J. and Le, X. H. B. (2013), ‘Simulating the dynamics of occupant behaviour for power management in residential buildings’, *Energy and Buildings* **56**, 85–93. 20
- Kibert, C. J. (2016), *Sustainable construction : green building design and delivery*, John Wiley & Sons. 19
- Lee, K. and Braun, J. E. (2004), ‘Development and application of an inverse building model for demand response in small commercial buildings’. 21
- Macdonald, I. and Strachan, P. (2001), ‘Practical application of uncertainty analysis’, *Energy and Buildings* **33**(3), 219–227. 44
- Magnier, L. and Haghghat, F. (2010), ‘Multiobjective optimization of building design using trnsys simulations, genetic algorithm, and artificial neural network’, *Building and Environment* **45**(3), 739–746. 21
- Mahdavi, A. and Tahmasebi, F. (2015), ‘Predicting people’s presence in buildings : An empirically based model performance analysis’, *Energy and Buildings* **86**, 349–355. 20
- Marrel, A., Iooss, B., Jullien, M., Laurent, B. and Volkova, E. (2011), ‘Global sensitivity analysis for models with spatially dependent outputs’, *Environmetrics* **22**(3), 383–397. 72, 73
- Masoso, O. and Grobler, L. J. (2010), ‘The dark side of occupants’ behaviour on building energy use’, *Energy and buildings* **42**(2), 173–177. 20, 62
- McKay, M. and Beckman, R. (1979), ‘Wj, a comparison of three methods for selecting values of input variables in the analysis of output from a computer codetechnometrics’, *Am Stat Assoc Am Soc Qual* **21**, 239–245. 85
- Menezes, A. C., Cripps, A., Bouchlaghem, D. and Buswell, R. (2012), ‘Predicted vs. actual energy performance of non-domestic buildings : Using post-occupancy evaluation data to reduce the performance gap’, *Applied Energy* **97**, 355–364. 17, 19, 20, 62
- Merheb, R. (2013), *Fiabilité des outils de prévision du comportement des systèmes thermiques complexes*, PhD thesis, Université Sciences et Technologies-Bordeaux I. 80
- Meteonorm (n.d.), ‘Meteonorm : Irradiation data for every place on Earth’.
URL: <http://www.meteonorm.com/en/>
- Morris, M. D. (1991), ‘Factorial sampling plans for preliminary computational experiments’, *Technometrics* **33**(2), 161–174. 74
- Neto, A. H. and Fiorelli, F. A. S. (2008), ‘Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption’, *Energy and buildings* **40**(12), 2169–2176. 21

- Nguyen, A.-T. and Reiter, S. (2015), A performance comparison of sensitivity analysis methods for building energy models, *in* 'Building Simulation', Vol. 8, Springer, pp. 651–664. [72](#)
- Nguyen, T. A. and Aiello, M. (2013), 'Energy intelligent buildings based on user activity : A survey', *Energy and buildings* **56**, 244–257. [20](#), [62](#)
- Nouy, A. (2010), 'Identification of multi-modal random variables through mixtures of polynomial chaos expansions', *Comptes Rendus Mécanique* **338**(12), 698–703. [116](#)
- Oldewurtel, F., Parisio, A., Jones, C. N., Gyalistras, D., Gwerder, M., Stauch, V., Lehmann, B. and Morari, M. (2012), 'Use of model predictive control and weather forecasts for energy efficient building climate control', *Energy and Buildings* **45**, 15–27. [19](#)
- O'Neill, Z. and O'Neill, C. (2016), 'Development of a probabilistic graphical model for predicting building energy performance', *Applied Energy* **164**, 650–658. [19](#), [21](#), [28](#)
- Page, J., Robinson, D., Morel, N. and Scartezzini, J.-L. (2008), 'A generalised stochastic model for the simulation of occupant presence', *Energy and buildings* **40**(2), 83–98. [20](#)
- Peuportier, B., Thiers, S. and Guiavarch, A. (2013), 'Eco-design of buildings using thermal simulation and life cycle assessment', *Journal of Cleaner Production* **39**, 73–78. [20](#)
- PREBAT (2015), Batiments démonstrateurs à basse consommation d'énergie : Enseignements opérationnels tirés de 60 constructions et rénovations du programme prebat, Technical report, Cerema. [16](#), [17](#), [43](#), [58](#), [59](#), [61](#), [66](#), [125](#), [131](#)
- Pérez-Lombard, L., Ortiz, J. and Pout, C. (2008), 'A review on buildings energy consumption information', *Energy and buildings* **40**(3), 394–398.
URL: <http://www.sciencedirect.com/science/article/pii/S0378778807001016>
- Rabl, A. (1988), 'Parameter estimation in buildings : methods for dynamic analysis of measured energy use', *Journal of Solar Energy Engineering* **110**(1), 52–66.
- Rabouille, M. (2014), Recherche de la performance en simulation thermique dynamique : Application à la réhabilitation des bâtiments, PhD thesis, Université de Grenoble. [80](#), [88](#)
- RAGE (2012), Retours d'expérience (rex) – bâtiments performants et risques – résultats 2012, programme d'accompagnement des professionnels, Technical report, Règles de l'art grenelle Environnement. [17](#)
- Réglementation thermique, G. E. (2012), *Réglementation thermique 2012*, cstb editions edn, 84, Avenue Jean Jaurès, Champs sur Marne, 77447 Marne-la-vallée Cedex 2. [14](#), [27](#), [58](#), [65](#), [66](#)
- Riahi, H. (2013), Analyse de structures à dimension stochastique élevée : application aux toitures bois sous sollicitation sismique, PhD thesis, Clermont-Ferrand 2. [79](#), [133](#)

- Rivalin, L. (2016), Vers une démarche de garantie des consommations énergétiques dans les bâtiments neufs : Méthodes d'évaluation des incertitudes associées à la simulation thermique dynamique dans le processus de conception et de réalisation., PhD thesis, PSL Research University. [72](#), [80](#), [116](#)
- Robillart, M. (2015), Etude de stratégies de gestion en temps réel pour des bâtiments énergétiquement performants, PhD thesis, Paris, ENMP. [21](#)
- Ryu, S. H. and Moon, H. J. (2016), 'Development of an occupancy prediction model using indoor environmental data based on machine learning techniques', *Building and Environment* . [20](#), [62](#)
- Saltelli, A., Chan, K., Scott, E. M. et al. (2000), *Sensitivity analysis*, Vol. 1, Wiley New York. [72](#), [75](#)
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008), *Global sensitivity analysis : the primer*, John Wiley & Sons. [73](#), [85](#), [89](#)
- Sanchez, D. G., Lacarrière, B., Musy, M. and Bourges, B. (2014), 'Application of sensitivity analysis in building energy simulations : Combining first-and second-order elementary effects methods', *Energy and Buildings* **68**, 741–750.
- Spitz, C. (2012), Analyse de la fiabilité des outils de simulation et des incertitudes de métrologie appliquée à l'efficacité énergétique des bâtiments, PhD thesis, Université Grenoble Alpes. [33](#), [54](#), [61](#)
- Spitz, C., Mora, L., Wurtz, E. and Jay, A. (2012), 'Practical application of uncertainty analysis and sensitivity analysis on an experimental house', *Energy and Buildings* **55**, 459–470. [60](#), [72](#)
- Stefanoiu, A.-M., Woloszyn, M., Jay, A., Wurtz, E. and Buhé, C. (2015), 'A methodology to assess the ambient temperature of a building using a limited number of sensors', *Energy Procedia* **78**, 1944–1949.
- Sudret, B. (2008), 'Global sensitivity analysis using polynomial chaos expansions', *Reliability Engineering & System Safety* **93**(7), 964–979. [90](#)
- Sun, Y. (2014), Closing the building energy performance gap by improving our predictions, PhD thesis. [16](#), [33](#)
- Sun, Y., Gu, L., Wu, C. J. and Augenbroe, G. (2014), 'Exploring hvac system sizing under uncertainty', *Energy and Buildings* **81**, 243–252. [19](#), [88](#)
- Tahmasebi, F. and Mahdavi, A. (2016), 'Stochastic models of occupants' presence in the context building systems control', *Advances in Building Energy Research* **10**(1), 1–9. [20](#)

- Teodosiu, C., Hohota, R., Rusaouën, G. and Woloszyn, M. (2003), ‘Numerical prediction of indoor air humidity and its effect on indoor environment’, *Building and Environment* **38**(5), 655–664. [20](#)
- Tian, W. (2013), ‘A review of sensitivity analysis methods in building energy analysis’, *Renewable and Sustainable Energy Reviews* **20**, 411–419. [67](#), [72](#)
- Tian, W. and De Wilde, P. (2011), ‘Uncertainty and sensitivity analysis of building performance using probabilistic climate projections : A uk case study’, *Automation in construction* **20**(8), 1096–1109. [19](#), [72](#), [73](#)
- Titikpina, F., Caucheteux, A., Charki, A. and Bigaud, D. (2015), ‘Construction of temporal data-based files for building energy simulations : The missing data issue’. [42](#)
- Trčka, M. and Hensen, J. L. (2010), ‘Overview of hvac system simulation’, *Automation in Construction* **19**(2), 93–99. [16](#), [19](#), [95](#)
- VAD (2013), Retours d’expérience sur 21 bâtiments performants de la région rhône-alpes – enquête menées de mars à aout 2013, en partenariat avec l’agence qualité construction, la région rhône-alpes et l’ademe, Ville et Aménagement Durable, Centre d’échanges et de ressources pour la qualité environnementale des bâtiments et des aménagements en Rhône-Alpes (2013). [17](#)
- Wang, L., Mathew, P. and Pang, X. (2012), ‘Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building’, *Energy and Buildings* **53**, 152–158.
- Warren, P. and Parkins, L. (1984), ‘Window-opening behaviour in office buildings’, *Building Services Engineering Research and Technology* **5**(3), 89–101. [20](#)
- Wiener, N. (1938), ‘The homogeneous chaos’, *American Journal of Mathematics* **60**(4), 897–936. [80](#)
- Wong, J. K., Li, H. and Wang, S. (2005), ‘Intelligent building research : a review’, *Automation in construction* **14**(1), 143–159. [20](#)
- Yan, D., O’Brien, W., Hong, T., Feng, X., Gunay, H. B., Tahmasebi, F. and Mahdavi, A. (2015), ‘Occupant behavior modeling for building performance simulation : Current state and future challenges’, *Energy and Buildings* **107**, 264–278. [20](#), [62](#)
- Yang, J. and Peng, H. (2001), ‘Decision support to the application of intelligent building technologies’, *Renewable Energy* **22**(1), 67–77. [20](#)
- Yu, W., Li, B., Jia, H., Zhang, M. and Wang, D. (2015), ‘Application of multi-objective genetic algorithm to optimize energy efficiency and thermal comfort in building design’, *Energy and Buildings* **88**, 135–143. [19](#)

Uncertainty assessment in building energy performance with a simplified model

Fally Titikpina^{1,*}, Antoine Caucheteux², Abderafi Charki¹, and David Bigaud¹

¹ L'UNAM, LARIS, Systems Engineering Research Laboratory, University of Angers, 62 avenue Notre Dame du Lac, Angers, France

² Cerema, risk, environment, mobility and development research Center, 23 avenue de l'Amiral Chauvin, 49136 Les Ponts-de-Cé, France

Received: 14 January 2015 / Accepted: 7 September 2015

Abstract. To assess a building energy performance, the consumption being predicted or estimated during the design stage is compared to the measured consumption when the building is operational. When valuing this performance, many buildings show significant differences between the calculated and measured consumption. In order to assess the performance accurately and ensure the thermal efficiency of the building, it is necessary to evaluate the uncertainties involved not only in measurement but also those induced by the propagation of the dynamic and the static input data in the model being used. The evaluation of measurement uncertainty is based on both the knowledge about the measurement process and the input quantities which influence the result of measurement. Measurement uncertainty can be evaluated within the framework of conventional statistics presented in the Guide to the Expression of Measurement Uncertainty (GUM) as well as by Bayesian Statistical Theory (BST). Another choice is the use of numerical methods like Monte Carlo Simulation (MCS). In this paper, we proposed to evaluate the uncertainty associated to the use of a simplified model for the estimation of the energy consumption of a given building. A detailed review and discussion of these three approaches (GUM, MCS and BST) is given. Therefore, an office building has been monitored and multiple temperature sensors have been mounted on candidate locations to get required data. The monitored zone is composed of six offices and has an overall surface of 102 m².

Keywords: Building energy performance, Uncertainty evaluation, GUM Method, Bayesian Approach, Monte Carlo

Nomenclature

n	Number of observations
N	Number of parameters
f	Functional relationship between parameters
R_{ay}	Horizontal Solar radiations
$\Delta\Psi$	Difference between the outdoor and indoor temperature
T_{out}	Outdoor Temperature
T_{in}	Indoor Temperature
M	Number of Monte Carlo trials
$Q_{building}$	Building heat loss
Q_{cv}	Heat loss due to conduction and ventilation
Q_{inf}	Heat loss due to infiltrations
Q_I	Heat loss due to building inertia
A_o	Heat gain due to occupation
A_s	Heat gain due to solar radiations
ξ	Building Inertia coefficient

U_b	Average Conduction coefficient through the building envelop
Φ	Envelop Solar radiations absorption coefficient
β	Envelop conduction coefficient
x_i, x_j	Input parameters
$u(y)$	Output uncertainty
γ	Infiltration coefficient due to wind
v	Wind Speed
S	Surface of the zone
y	Output parameter
q	Air flow rate in the building
c_p	Air Specific Heat
ω	Week of the year
X	A selected input quantity

1 Introduction

The regulation of energy consumption in all area has become today a worldwide concern. In a context of energy resources increasing scarcity, especially oil and gas, it has become essential and vital to limit the energy demand of

* Correspondence: fally.titikpina@univ-angers.fr

some sectors especially in the area of buildings energy consumption. According to the International Energy Agency (IEA), at the current consumption trend, oil reserves will run out in less than 50 years, natural gas in 60 years and coal in 120 years while the energy demand is increasing substantially [1].

The building sector is the largest energy consumer in the European Union. In 2008, it engages 40% of the total energy consumption and gives off 36% of CO₂ emissions. Among all economic area in Europe, the construction industry is the area of huge energy demand. It is responsible of more than 20% of greenhouse gas emissions [2]. With the continuous demographic growth, these numbers are in continual progression and are making the building energy consumption matter an international issue, which is managed by public policies [3].

A building energy performance is defined as the amount of energy actually consumed or estimated to meet the different needs associated with a standardized use of the building, which may include heating, cooling, hot water, ventilation and lighting [3]. This standardized use is somewhat unpredictable and usually lead to a gap in the energy predicted during the building design stage and the real energy consumption when the building is operational [4].

The challenge in the construction industry is to ensure that the predicted energy performance of a given building is achieved with a minimum of uncertainty at the operational stage [4]. This situation comes to provide high performance in thermal comfort and reliability at lowest cost possible. This task can be achieved primarily by finding the sources of the performance gap and later by estimating the amount of uncertainty associated to the performance being predicted [5].

The parameters that describe a building such as the static parameters (wall absorption, inertia coefficient, conduction and infiltration coefficients due to wind) and dynamic parameters (weather and occupation) are relevant sources of uncertainty but does not represent the entire causes when estimating building energy consumption. De wit [6] introduced various causes of uncertainty in the building performance assessment: the specification uncertainty coming from inaccurate description of the building and its systems, the modelling uncertainty which comes from the physical model itself, the errors introduced in the model during simulation, and those arising from external conditions and occupants [7]. To better sense uncertainty and its influence on building performance, it is important that these sources of uncertainty are known. Several research works were focused on this problem, although few case studies included this aspect into their analyses. In this paper, the uncertainties coming from measured data are evaluated in different ways.

Most of time, the uncertainty associated to measured data comes not only from the random errors but also from the data itself. These parameters are affected by the measurement methods, the equipment being used, the quality of the data and also by the assumptions made during the analysis process. A set of methods are used to es-

timate uncertainty in measurements. Among them, the GUM (Guide to the expression of Uncertainty in Measurements) [8], the Monte Carlo Simulation [9], and the Bayesian inference [10, 11] are the common ones.

These methods are based on both the knowledge about the measurement process and on the impact of parameters being involved. The model equation represents the knowledge about the measurement process and the relation between the quantities while their degree of belief is expressed by probability density functions.

In this paper, we tried to estimate the energy demand of a selected zone of an office building with a simplified model, (the BEECHAM model [12] which combines physical principles and statistical method and its associated uncertainty with the three different approaches cited above. The content is organized as follows; Section 2 reviews the uncertainty estimation methods, each one with its base principles and the presentation of the BEECHAM model, followed in Section 3 by the application made and in Section 4 by the results and discussions.

2 Materials and methods

2.1 GUM principle

The Guide to the expression of Uncertainty in Measurements [8] is frequently used to evaluate measurement uncertainty. It has been published in 1993 and constitutes one of the reference documents in the area of measurement uncertainty. When repeated measurements are made on a variable, or when information is available on the data or on the measurement process, the Guide specifies the estimation of the associated uncertainty. Then, given the estimates and the associated uncertainties of the input quantities, the GUM describes how to determine the best estimation of the measurand and its associated uncertainty as well as a confidence interval.

According to the GUM, uncertainty is a parameter associated to the result of a measurement, which characterizes the dispersion of the values that could be reasonably attributed to the quantity being measured (measurand). Two types of methods are presented. Among them, the type A method of evaluation of uncertainty, carried out by the statistical analysis of series of observations. In most cases, the measurand y is not measured directly, but is determined from the measurement of N other quantities x_1, x_2, \dots, x_N through a functional relationship f , given in equation (1).

$$y = f(x_1, x_2, \dots, x_N). \quad (1)$$

The GUM method is based on the concept of a measurement equation (Eq. (1)), which represents the process to determine the best estimate of y and its associated standard uncertainty $u(y)$ from the estimates and associated standard uncertainty of the input quantities x_1, x_2, \dots, x_N [8]. In this equation the variables are the input quantities and the measurement result is obtained by introducing the estimates of the inputs in

the model. Then, the uncertainty of the output quantity is calculated from the measurement model by propagation in equation (2). The GUM Uncertainty framework is then said to be a method of propagation of uncertainties. When the input quantities are correlated, the appropriate expression for the uncertainty associated with the result of a measurement is:

$$u^2(y) = \sum_{i=1}^N \frac{\partial f}{\partial x_i}^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left(\frac{\partial f}{\partial x_i} \right) \left(\frac{\partial f}{\partial x_j} \right) u(x_i, x_j), \quad (2)$$

where x_i and x_j are the estimates of the parameters X_i and X_j and $u(x_i, x_j)$ is the estimated covariance associated to x_i and x_j . To quantify the uncertainty of a measurement using the GUM approach, one does not need to compute any complicated calculation. The methodology is clearly stated in the GUM and simple to apply.

2.2 Monte Carlo method for uncertainty estimation

An updated supplement to the GUM, GUM S1 [13], specifies uncertainty evaluation in terms of a probability density function. It describes the numerical propagation of probability density function through the model mentioned before. The Monte Carlo methodology [14], as presented by the GUM S1, induces the propagation of the input quantities distributions through the model in order to provide the output probability distribution [15].

When the inputs density functions are defined, a number of Monte Carlo trials is selected. Generally, the greater the number of simulation is, the better are the results. The evaluation of the measurement uncertainty from Monte-Carlo simulation is done through the following steps.

After setting up the relationship between the analytical result y and the input quantities x_i , which is somewhat the model equation for the measurement process, the probability density function $p(x_i)$ of each input quantity x_i , have to be identified according to all the uncertainty sources. A number M of Monte Carlo trials is selected. Then, one generates M samples $x_{i1}, x_{i2}, \dots, x_{iM}$ of each x_i . From the estimated set of result, $y_{i1}, y_{i2}, \dots, y_{iM}$ the combined uncertainty $u(y)$ can be easily calculated by the propagation of the generated data through the functional relationship f as stated in equation (3). The Monte Carlo Method (MCM) is then said to be a method of propagation of distributions. However, when the inputs quantities are correlated, their probability density function are suitably estimated by a joint probability density function. Then, a confidence interval of the output y is estimated after the choice of a coverage probability p .

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = f \left(\begin{bmatrix} x_{11} & x_{21} & \dots & x_{N1} \\ x_{12} & x_{22} & \dots & x_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1M} & x_{2M} & \dots & x_{NM} \end{bmatrix} \right). \quad (3)$$

The Monte Carlo Simulation is easy to apply and returns the information about the model distribution. However, it also has some limitations: the model simulation runtime could be long in some complex cases. Furthermore, the selection of the proper probability distribution functions for the model parameters may be difficult due to inaccurate data. In addition, the accuracy of the numerical simulation depends on the number of simulations.

Regarding the Monte Carlo Method, it is sometimes difficult to reliably specify the probability density function of each input variable. If some inputs are known to be correlated, it will be difficult to specify their joint probability distribution. Most of time, an approximation is made from the available data and information. Furthermore, related to the ISO-GUM approach, the Monte Carlo Simulation does not make the distinction between the uncertainties estimated by the statistical approach and those arising from other source such as sensors specifications, previous tests results or from the operator judgment. The difference should be done to estimate the probability density function of the input quantities.

2.3 Bayesian uncertainty quantification

For Bayesian Uncertainty Quantification, all the information available on a quantity is updated by the measurement in order to assign a value to this quantity. Bayesian statistics is based on Bayes Theorem [10]. This theorem comes from the axiom which states that given two propositions X and Y , the probability of both of them being true is equal to the product of the probability of X being true by the probability of Y being true given that X is true (Eq. (4)).

$$p(X/Y) = \frac{p(X)p(Y/X)}{p(Y)}. \quad (4)$$

Let us consider a sample of independent and identically distributed (i.i.d), x_1, x_2, \dots, x_N random variables, which depends on an unknown parameter θ , with a probability density function $p(x/\theta)$ $\pi(\theta)$ is the probability density function of θ and $L(x/\theta)$ the measurement likelihood function (Eq. (5)). The posterior distribution of θ denoted by $\pi(\theta/x)$, is the conditional distribution of θ given the observations of x , with $x = (x_1, x_2, \dots, x_N)$, the i.i.d variables considered above.

$$L(x/\theta) = \prod_{i=1}^n p(x_i/\theta). \quad (5)$$

The posterior distribution $\pi(\theta/x)$, gives the information about θ after measurement is done (Eq. (6)). This posterior distribution is found through the Bayesian inference.

$$\begin{aligned} \pi(\theta/x) &= \frac{L(x/\theta) \pi(\theta)}{\int_{\theta} L(x/\theta) \pi(\theta) d\theta} \\ \pi(\theta/x) &= \frac{\prod_{i=1}^n p(x_i/\theta) \pi(\theta)}{\int_{\theta} \prod_{i=1}^n p(x_i/\theta) \pi(\theta) d\theta}. \end{aligned} \quad (6)$$



Fig. 1. Data-driven method flow chart.

Equation (6) can then be written as:

$$\pi(\theta/x) \propto \prod_{i=1}^n p(x_i/\theta) \pi(\theta). \quad (7)$$

Bayesian inference about a quantity [11], is made through the probability density function that describes the information acquired from measurement and the knowledge about the quantity before the measurement is performed [16, 17]. Bayes formula is a mechanism that combines the prior information on the parameters and the information provided by the measured data. The Bayesian approach applied with probability density functions is given in equation (8) by [18].

$$p(y/x_1, x_2, \dots, x_N) \propto p(x_1, x_2, \dots, x_N/y) p(y). \quad (8)$$

When estimating the uncertainty of a quantity by the Bayesian method, it is often difficult to assign prior information to the quantity being measured. To construct the prior density function, one should make use of all available prior information such as calibration data, data from other related experiments, personal judgment and other available evidences about the quantity. Still, since that information comes from a personal opinion and are unreliable, it can result to significant uncertainties [19].

2.4 A simplified building energy model: the BEECHAM model

One of the major concerns for the building sector nowadays is the evaluation and the prediction of a building energy performance. Many factors influence the energy behaviour of a building. Weather conditions, the occupants and their activities, the components such as lighting, the building thermal properties, the HVAC systems performance are the relevant parameters. These elements induce several variables, which are involved in the building energy model and can be split into two groups: the dynamic and the static parameters. The dynamic factors are known to change with time like the weather, the occupancy and the static parameters are the non-variable factors such as the building construction and thermal property. The energy model of buildings is then said to be complex and when predicting its energy consumption one might take into account all these factors.

Several methods rather complex or simplified are used to model the thermal performance of a building and its heat gains. Different techniques, going from statistical methods such as simple regression model to engineering methods based on physical principles are used. Artificial intelligence methods for example Artificial Neural Networks (ANN) [20, 21] or Support Vector Machines

(SVMs) [22] are also used to forecast the building energy consumption. In this paper, we will focus on an engineering and statistical based method especially a data driven model (Fig. 1) which uses physical and thermal principles to estimate the energy behaviour of a building.

According to Coakley et al. [23], data-driven models can be used to model a system behaviour with a minimum set of adjustable inputs. The statistical methods are used to construct physical model from measured data as illustrated in Figure 1. This kind of methods mainly focuses on the thermal phenomena occurring between a building and its surroundings to calculate precisely, at each time step the building heat load. The most influential and the most important factors to determine a building energy demand are the weather conditions. The outdoor temperature, the solar radiations and the wind speed have a non-negligible impact on the conduction through the building envelop and on the infiltrations through the building walls.

Caucheteux et al. [12] developed a model, the BEECHAM (Building Energy Efficiency Characterization Method) with which the heating requirement of a given building can be estimated based on its thermal balance.

The energy balance of a given building is assessed as previously said, by summing the heat losses and the heat gains under operational conditions. Then, the heat loss due to conduction through the building envelop Q_C , the heat loss due to ventilation Q_V , the heat loss due to infiltrations by wind and the stack pressure effect Q_{inf} , the heat gain due to solar radiations A_S , the heat load due to occupancy A_o (occupation, the use of electrical equipment's, lighting, ...), and the heat gain or loss (depending on the case) of energy in the building construction materials, which is called inertia Q_I , are summed in order to find the required amount of energy so that the building is balanced on a thermal basis. For a given period, the heating requirement of the building can be estimated by equation (9). Each element in these equations is denoted in the nomenclature part.

$$Q_{building} = Q_{cv} + Q_{inf} + Q_I - A_s - A_o \quad (9)$$

$$\Delta\Psi = \sum_{24} (T_{out} - T_{in}). \quad (10)$$

At a specified time step, the conduction and ventilation heat loss through a building envelop Q_{Cv} , depends on the outdoor and indoor temperature difference $\Delta\Psi(K)$, like specified in equation (10), on the ventilation average flow rate and on the conduction coefficient β (kW/(K)). β represents the irradiative and convective exchange of the walls in the overall zone and also the ventilation flow rate. Then, the conduction and ventilation heat loss through the building envelop can be expressed by equation (11).

$$Q_{cv} = \beta\Psi. \quad (11)$$

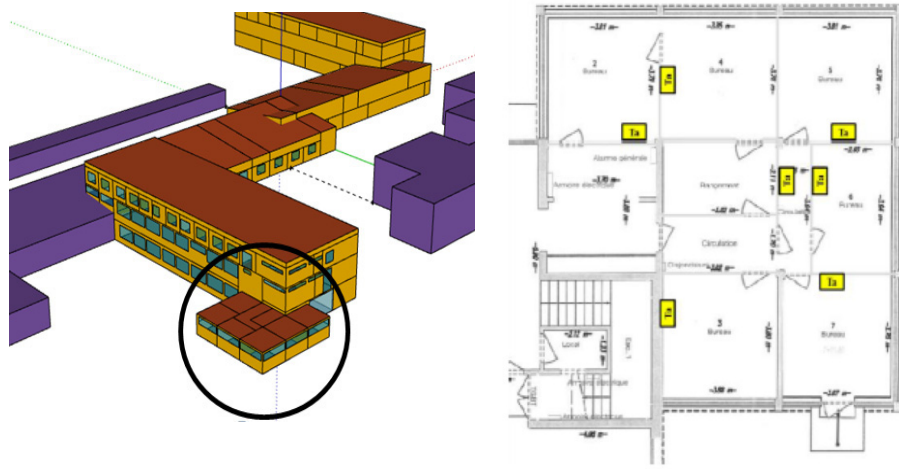


Fig. 2. (a) A 3-dimension view of the monitored building. (b) The monitored zone.

Two effects are mainly responsible of infiltrations: the stack effect and the wind pressure. Due to the simplicity of the model, the stack effect was not considered (the model is supposed to be applied to small-rise buildings). The infiltrations are then expressed by equation (12) where γ represents the infiltrations coefficient due to wind.

$$Q_{inf} = \gamma \left((\vartheta^2)^\alpha \right) \Psi, \quad \text{with } \alpha = 2/3. \quad (12)$$

The airflow through the building envelop is related to the difference of pressure created in the building within its environment. The parameter α is usually called the flow exponent and depends on the nature of the airflow (turbulent or laminar). By default, one retains usually a value of 2/3 [24]. An exponent close to 1 characterize laminar flow while a value of 1/2 represents a turbulent airflow.

The capability of a given building to charge and discharge the intake energy is called inertia. It is assumed that it depends essentially on the difference in temperature from one day to the next one. For example, if the considered day temperature is lower than the previous one, the building will give in the energy from the day before. This heat transfer can be expressed as:

$$Q_I = \xi \frac{\partial \Psi}{\partial t}, \quad (13)$$

where ξ is the building inertia coefficient and t is the considered time step. Usually, the heat gain due to solar radiations is estimated with an accurate description of the building (surfaces and windows being exposed to the sun, their orientations and titles angles ...). The BEECHAM model aims to avoid the description of the overall building which is time consuming and which parameters are often unknown. With these assumptions, the heat gain due to solar radiations is supposed to be constant during a time scale of one week and is stated in equation (14) where Φ represents the building geometry and R_{ay} the average solar radiations on horizontal.

$$A_s = \Phi R_{ay}. \quad (14)$$

Unlike the other coefficients β, γ, C_m which are stationary, Φ can be represented over the year, by a polynomial of degree equal or greater than 2 [12]. The degree of the polynomial depends on the presence or not of solar masks adjacent to the building. The time scale used to represent this polynomial is the week. A polynomial regression is then performed to estimate the coefficients a, b and c in equation (15). ω represents the week of the year and the resulting model for the estimation of the building energy demand is given by equation (16).

$$\Phi = a\omega^2 + b\omega + c \quad (15)$$

$$Q_{building} = \beta\Psi + \gamma \left(\vartheta^{4/3} \right) \Psi - (a\omega^2 + b\omega + c) R_{ay} + \xi \frac{\partial \Psi}{\partial t} - A_o. \quad (16)$$

The BEECHAM model, in equation (16), is somewhat simplified and can be qualified as a data-driven approach among the building energy models [23]. It uses the building monitored data through the model to predict the system thermal behaviour. Thus, with a minimal set of dynamic inputs (the outdoor and indoor temperature, the solar radiations, the wind speed), the user is able to estimate the heating load of a building. It also simplifies the measurements, as there is no analysis of the user behaviour or the solar gain evaluations. In this way, it allows to reduce the errors coming from the operator judgement.

3 Building description

The purpose of this work is to evaluate the heating requirement of a building zone and its associated uncertainty with the three methods described before: the GUM, the Monte Carlo Method and the Bayesian Method for uncertainty quantification. The monitored zone is a part of the ground floor of an administrative building located in Angers, France (Fig. 2). It is composed of six offices with a circulation area and a total surface of 102 m² (Fig. 2a).

Table 1. Estimation of the static parameters distributions.

Static parameters	Distributions
Envelop Conduction Coefficient β (kWh/°K)	$N(0.4, 0.057)$
Infiltration coefficient due to wind γ (kW/°K(km/h))	$N(9.57 \times 10^{-7}, 6.28 \times 10^{-7})$
Building inertia Coefficient ξ (kW/°K)	$N(-0.139, 0.1559)$
Solar radiation Coefficient $a\omega^2 + b\omega + c$ (m ²)	$0.07\omega^2 - 0.29\omega + 1.39$

Most of the building occupancy occurs between 8:00 and 18:00. The sensors have been fixed in all the offices to get the temperature data of the overall zone. The outdoor temperature, the solar radiations and the wind speed data for the monitored period are provided by a meteorological station implanted near the building being experimented. The temperature data are collected during one year from January to December 2013 in each office (Fig. 2b) with a time step of one hour. The average wind speed, horizontal solar radiation, and outdoor temperature are also measured each hour. For the occupancy rate, we assumed that there is one person in each office during work time in order that the occupancy rate is proportional to the electrical consumption being measured. Then, with the static parameters of the building estimated above, the heating requirement of the zone was calculated.

Regarding the test-bed, each sensor can be independent and works with a battery, which allows them to have their own memory to record the data. The sensors have been put on walls, at half height and away from openings (doors and windows).

4 Results and discussion

4.1 Inputs quantities uncertainty quantification

As the BEECHAM model is known to be a data-driven method, it requires prior data in order to model the building energy behaviour. For the static parameters (the envelop conduction coefficient, the infiltration coefficient due to wind, the building inertia coefficient and the solar radiation coefficient), as they are not changing in time, their daily uncertainty cannot be measured. A suitable methodology is required to assign them a variation range, in which they could be included, in order to do the uncertainty analysis. For this purpose, we used three months heating consumption measurements results, going from January to mid-March 2013 to get very wide upper and lower bonds. We realised 10 000 random samples from an uniform distribution and we calculated the daily energy consumption with these values.

To select the parameters ranges, we assumed that the model is calibrated. The ASHRAE guidelines [3] suggested the use of standardised statistical indices to sense the performance of a model. Among these criterion, the Coefficient of Variation of Root Mean Square Error, *CVRMSE* (Eq. (17)), allows to determine how well the energy demand model being used fits the data. For a model to be calibrated, ASHRAE sets the *CVRMSE* on a monthly basis to 15% and on an hourly basis to 30%. Thus, only

the set of static parameters with a *CVRMSE* < 20% are picked up (Tab. 1 and Fig. 3).

Equation (17) gives the *CVRMSE* of a model, with m_i and s_i , the measured and simulated data points, N_i the number of data points at regular interval (hourly, daily or monthly) and \bar{m} the mean of the measured data points.

$$CVRMSE = \frac{\sqrt{\left(\sum_{i=1}^{N_i} \frac{(m_i - s_i)^2}{N_i}\right)}}{\bar{m}}. \quad (17)$$

The probability density plots of the static model parameters β, γ, ξ, a, b and c are determined according to the previous results (Fig. 3). These probability density results give us information on the variation range of each parameters and the probability to obtain each value, which will be useful for the computation of the GUM, MCM and BST approaches. For the estimation of the building solar radiation equation please refer to the experiments in reference [12].

On the other hand, for the dynamical parameters involved such as the weath24 hourser, the occupancy and the indoor temperature, measurements data are available during the selected period. Still, the user needs to find the adequate probability distribution of each parameter and the uncertainty associated to the measured data.

After having performed several distributions fittings to the measured data, we assumed that each dynamic variable X , follows a normal distribution with its associated distribution coefficients. The related uncertainty $u(X)_k$, is obtained by adding the uncertainty of the measurement sensors $u(X)_s$ and those arising from the data itself [8]. Then, in order to take into account the variability of the data, an uncertainty $u(X)_k$ is assigned to the variable at each time step k (Eq. (18)).

$$u(X)_k^2 = u(X_k)^2 + u(X)_s^2. \quad (18)$$

For example, in Figure 4 is represented the temperature histogram in the six offices of the monitored zone and the circulation area. The represented temperature is the sum on 24 hours of the difference between the outdoor and indoor temperature (Eq. (10)). Then, we assumed that the temperature data follows a normal distribution in all the offices and the same procedure is used for the solar radiations and the wind speed data.

4.2 Propagation of uncertainty in the model

The purpose of this research work is to estimate the uncertainty associated to the daily energy consumption of a

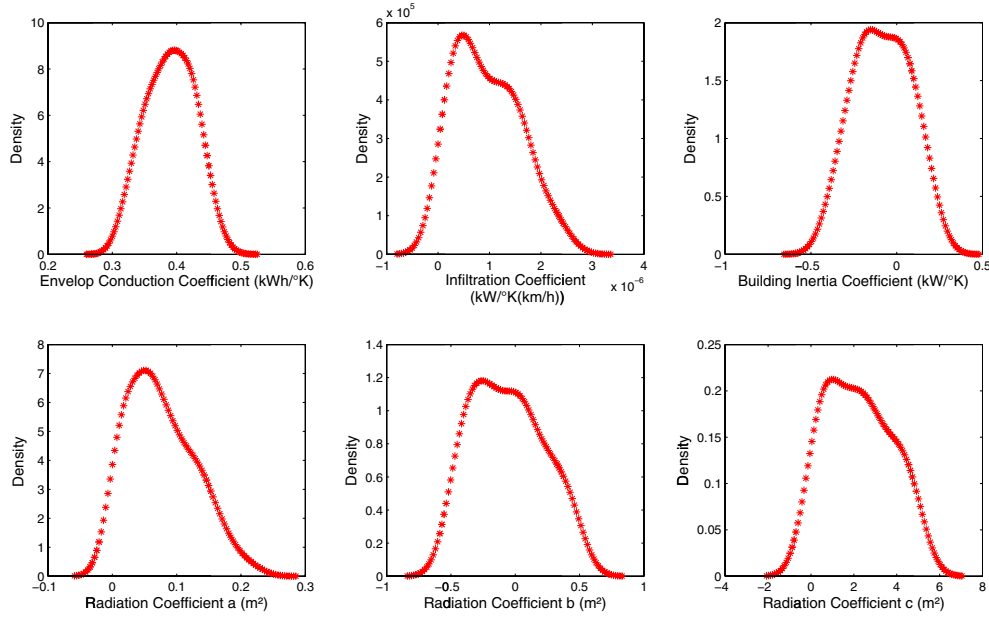


Fig. 3. Static parameters density plots obtained from the 10 000 samples realised.

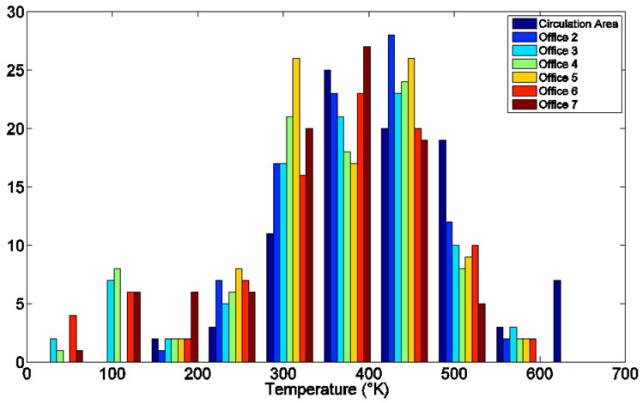


Fig. 4. Sum on a day basis of the difference between the outdoor and indoor temperature distribution ($^{\circ}\text{K}$) in each part of the monitored zone.

building estimated by a given model. The uncertainty will be estimated by the three different methods cited above in order to find the better or the more suitable one for the estimation of a daily energy consumption.

As stated above, the BEECHAM model dynamic parameters are:

- the difference between the indoor and outdoor temperature over a day Ψ ;
- the sum of the solar radiations over a day Φ ;
- the wind speed v found by taking the average on a day basis also.

Knowing that the weather conditions and the occupancy within the building energy demand model are changing at each iteration, all the uncertainty analysis have been done with the assumption that their daily variability is equal to zero which means that the weather conditions and the

occupancy rate for example are the same from one day to another.

Then, to get the uncertainty associated to these three dynamic parameters, we combined the uncertainty of the sensor being used (depending on the measured parameter) and the uncertainty coming from the data itself (depending on the selected method) like stated in equation (16). If it is the GUM method, the user will have an amount of uncertainty associated to its data; if it is the Monte Carlo method, the uncertainty will be on the form of a probability density function and if it is the Bayesian Analysis, then according to the prior information available on the parameter under consideration and the data recorded the user will have a posterior distribution to perform the uncertainty analysis.

Then, for the GUM method and referring to equation (2), we assumed that the input quantities are not correlated and we estimated thanks to the BEECHAM model, the value $\frac{\partial Q_{\text{building}}}{\partial X}$ of each input quantity X and we obtained the uncertainty associated to each variable $u(X)$ thanks to equation (18). The results are presented in Tables 1 and 2.

To give a better interpretation to the uncertainty obtained, the GUM [8] proposed the calculation of a confidence interval in which the best estimate of the measurand will fall in. A lower and upper bounds are estimated with a coverage factor α like stated in the following equation.

$$X - \varphi * u(X)_k \leq X \leq X + \varphi * u(X)_k. \quad (19)$$

With the 60 days data sets, we estimated the probability to obtain the energy consumption of the monitored zone with its upper and lower bounds within a confidence level of 95% ($\varphi = 2$).

The estimation of the uncertainty with the Monte Carlo Method was quite easier after knowing the distribution functions of all the input quantities. We performed

Table 2. Estimation of the dynamic parameters distributions with measured data.

Dynamic parameters	Distributions
Outdoor and indoor temperature Difference Ψ ($^{\circ}\text{K}$)	$N(386.76, 59.04)$
Solar Radiations over a day R_{ay} (W/m^2)	$N(1.19 \times 10^{+03}, 695.49)$
Wind Speed ϑ (km/h)	$N(10.84, 6.18)$
Occupancy rate A_o (Wh)	$N(94.4, 47.7)$

Table 3. The daily energy consumptions and its uncertainty estimated by the three approaches and compared to the measured energy consumption in the monitored zone.

Method	Estimated Value (kWh)	Associated Uncertainty (kWh)	Proportion of Uncertainty (%)
GUM Principle	1.493×10^2	29.56	19.79
Monte Carlo Method	1.491×10^2	29.91	20.06
Bayesian Uncertainty Quantification	1.567×10^2	24.99	15.95
Measured Energy Consumption	1.363×10^2	32.97	24.18

a number of simulations of 10^6 in order to have a large dataset for the output estimation. We generated 10^6 data for each input quantity according their distribution functions and the data obtained from the generation are spread through the BEECHAM model like stated in equation (3). The energy consumption result obtained on the 60 days basis is then compared to the GUM results in Figures 6 and 7.

However, for the Bayesian analysis the process was quite different from the last two methods. We have to define the prior knowledge for each parameter and the likelihood functions. Having performed measurements on the dynamic parameters, the likelihood functions was known as well as their probability density functions and the prior was made of the weather data used when performing a building simulation. In TRNsys 17 [25] for example, to perform a building energy simulation, TMY (Typical Meteorological Year) files are opened for users and represent a typical weather data. The files are built from the weather data measured during the last 20 years and are available in several cities. Then, we chose two months measurement files to make the weather parameters (temperature, solar radiations and wind speed) prior functions (Fig. 5).

Nevertheless, for the static parameters, we decided to use normal distributions to build the prior knowledge of each parameter and the dataset built previously for the GUM and Monte Carlo Method for the likelihood functions. The uncertainty associated to the building zone estimated by the Bayesian method is obtained by the propagation of each posterior distribution through the BEECHAM model and the results are compared to the previous methods (Figs. 7 and Tab. 3).

5 Discussion and conclusion

In order to assess the BEECHAM model consistency, we decided to compare the two months real energy consumption to the energy consumption estimated by the model for the monitored zone presented before and during the same period. The heat load of the building zone was measured during January and February providing a 60 days

dataset. Table 3 shows the results comparison. One can notice that the measured consumption is much closed to the consumption estimated by the BEECHAM model but the real daily energy consumption is overestimated by an average of 27 kWh in the whole period. This difference can be explained by the fact that a building energy consumption is badly estimated and sometimes lead to a gap between the estimated energy performance (by simulation or by calculation), of a building and the real one. De Wilde [4] pointed out this issue and stated that it appears mainly at three stages of the building life: the design stage, the construction stage and the operational stage. The real energy consumption can go from two to three times the predicted energy, which is simply huge to be acceptable.

However, the amount of uncertainties arising from the data analysis can be explained in different ways. The first issue is the presence of uncertainties in the model itself. A building model, either a simulation or an analytical model, aims to represent the building in real conditions and as true as possible. The use of improper methods or incorrect input data can lead to increase the amount of uncertainty. However, even if the model is supposed to be correct, the user needs to have the proper qualifications and knowledge to apply it.

Uncertainties could also come from the measurement procedures and the devices used for the data acquisition. Sometimes, these instruments are the root causes of data deviation and the operator judgement and assumptions can have a huge influence on the results in different situations. For example, in the Bayesian analysis case, the user sometimes has to make hypotheses in order to have a consistent prior knowledge on a quantity, which will influence the results on the posterior distributions.

Finally, the last cause could be the variability in the dynamical data. It is known that the outdoor temperature at a moment t for example, depends on the temperature at $t - 1$ and will be different from the temperature at time $t + 1$. This variability in time of the dynamical parameters is fairly unpredictable and cannot be quantify on a regular basis. This dynamical feature is a non-negligible source of uncertainty and should be taken into account

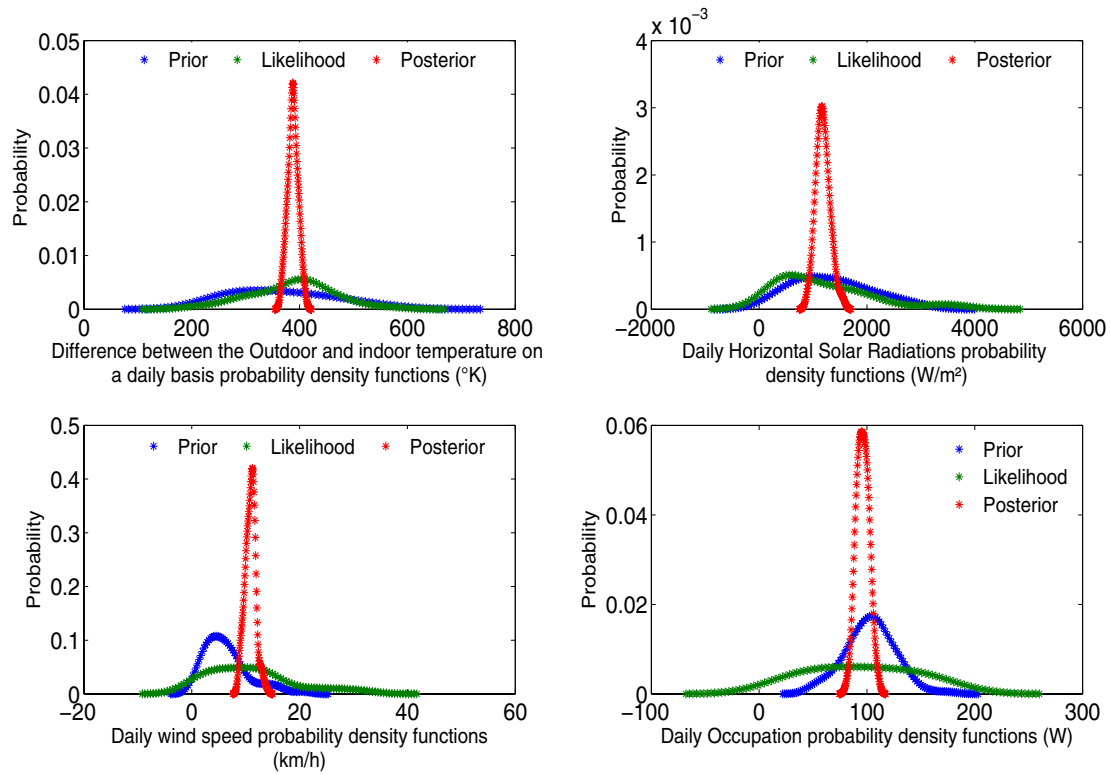


Fig. 5. Daily probability density plots (prior, likelihood and posterior) for the dynamic input parameters.

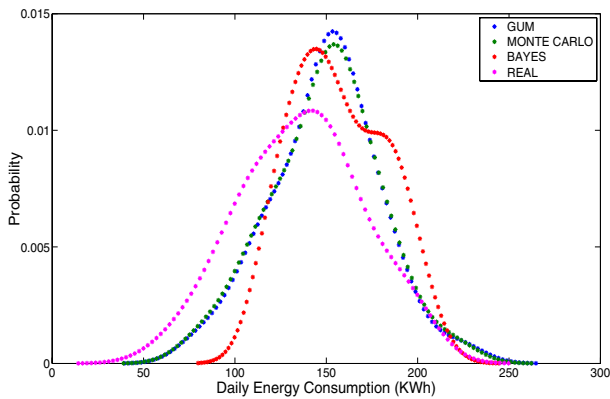


Fig. 6. Daily energy consumption probability plots obtained from the gum, the Monte Carlo method, the bayesian principle and the measured consumption.

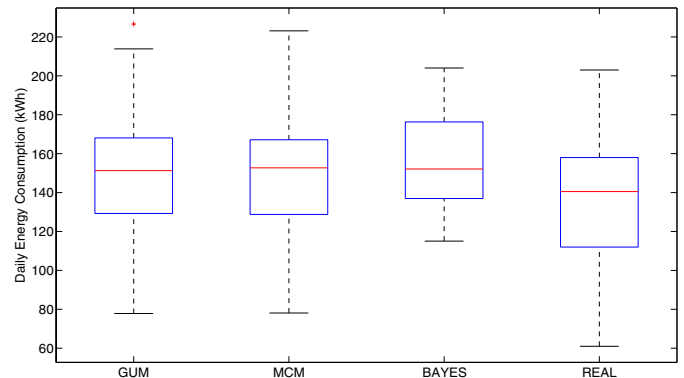


Fig. 7. Dispersion of the daily energy consumption data estimated by the three approaches and compared to the measured data.

when quantifying the uncertainty related to the dynamic parameters.

The main goal of this research work was to compare three approaches of uncertainty quantification: the Guide to the expression of Uncertainty in Measurement, the Monte Carlo Method and the Bayesian Uncertainty Quantification with a simplified model. These three approaches allow quantifying the uncertainty associated to the energy consumption of a given building with low variation among them. They give almost the same results and have affirmed the research which have already been made in metrology [19,26]. Each method has certainly its own limitations

and its advantages, thus, the user will have to choose according to the data and the information he already has, the convenient one. This paper has also pointed out by its results, the existence of a gap between the predicted and the actual energy consumption in a given building and has highlighted the causes of such discrepancies. Moreover, this analysis was done with the assumption that all the input parameters are uncorrelated. A logical continuation of this research work will be to model the relationship between them and study its influence on the uncertainty analysis. Then, the analysis can be extended to the optimization of the measurement plan.

Acknowledgements. This study was funded by the grants from the National Association of Research (ANR) in France within the OMEGA project and the data was provided by the Cerema.

References

1. IEA, "Key World Energy Statistics 2014 – Key World 2014.pdf," 2014. [Online]. Available: <http://www.iea.org/publications/freepublications/publication/KeyWorld2014.pdf>. [Accessed: 10-Dec-2014].
2. L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, *Energy Build.* **40**, 394–398 (2008)
3. ASHRAE, ASHRAE Guideline 14, Measurement of energy and demand savings (2002)
4. P. de Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, *Autom. Constr.* **41**, 40–49 (2014)
5. A.C. Menezes, A. Cripps, D. Bouchlaghem, R. Buswell, Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap, *Appl. Energy* **97**, 355–364 (2012)
6. S. De Wit, G. Augenbroe, Analysis of uncertainty in building design evaluations and its implications, *Energy Build.* **34**, 951–958 (2002)
7. S. Attia, M. Hamdy, W. O'Brien, S. Carlucci, Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design, *Energy Build.* **60**, 110–124 (2013)
8. GUM, Guide to the Expression of Uncertainty in Measurement (1995)
9. O. Bodnar, G. Wübbeler, C. Elster, On the application of Supplement 1 to the GUM to non-linear problems, *Metrologia* **48**, 333 (2011)
10. K. Weise, W. Woger, A Bayesian theory of measurement uncertainty, *Meas. Sci. Technol.* **4**, 1 (1993)
11. I. Lira, G. Kyriazis, Bayesian inference from measurement information, *Metrologia* **36**, 163 (1999)
12. A. Caucheteux, A Building Energy Efficiency Characterisation Method (BEECHAM) to Assess Existing Buildings Performance (2012)
13. C. Elster, W. Wöger, M.G. Cox, Draft GUM Supplement 1 and bayesian analysis, *Metrologia* **44**, L31 (2007)
14. M. Ángeles Herrador, A.G. González, Evaluation of measurement uncertainty in analytical assays by means of Monte-Carlo simulation, *Talanta* **64**, 415–422 (2004)
15. M.G. Cox, B.R. Siebert, The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty, *Metrologia* **43**, S178 (2006)
16. C. Elster, B. Toman, Bayesian uncertainty analysis under prior ignorance of the measurand versus analysis using the Supplement 1 to the Guide: a comparison, *Metrologia* **46**, 261 (2009)
17. I. Lira, D. Grientschnig, Bayesian assessment of uncertainty in metrology: a tutorial, *Metrologia* **47**, R1 (2010)
18. I. Lira, W. Wöger, Bayesian evaluation of the standard uncertainty and coverage probability in a simple measurement model, *Meas. Sci. Technol.* **12**, 1172 (2001)
19. R. Kacker, A. Jones, On use of Bayesian statistics to make the Guide to the Expression of Uncertainty in Measurement consistent, *Metrologia* **40**, 235 (2003)
20. S.A. Kalogirou, C.C. Neocleous, C.N. Schizas, Building heating load estimation using artificial neural networks, in *Proceedings of the 17th international conference on Parallel architectures and compilation techniques*, (1997), Vol. 8, p. 14
21. T. Olofsson, S. Andersson, Long-term energy demand predictions based on short-term measured data, *Energy Build.* **33**, 85–91 (2001)
22. Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Applying support vector machine to predict hourly cooling load in the building, *Appl. Energy* **86**, 2249–2256 (2009)
23. D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, *Renew. Sustain. Energy Rev.* **37**, 123–141 (2014)
24. AFNOR, Norme Européenne, Norme Française, Performance Energetique des batiments, Calcul des besoins d'énergie pour le chauffage, Batiments résidentiels. Norme AFNOR, Aout-1999.
25. A. Trnsys, Transient System Simulation Program, University. Wis. (2000)
26. C. Elster, B. Toman, Bayesian uncertainty analysis for a regression model versus application of GUM Supplement 1 to the least-squares estimate, *Metrologia* **48**, 233 (2011)

CONSTRUCTION OF TEMPORAL DATA-BASED FILES FOR BUILDING ENERGY SIMULATIONS: THE MISSING DATA ISSUE

Fally Titikpina¹, Antoine Caucheteux², Abderafi Charki¹, David Bigaud¹,

¹LARIS, Laboratoire Angevin de Recherche en Ingénierie des Systèmes, Université d'Angers, FRANCE

²CEREMA, Centre d'Etudes et d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement, FRANCE

ABSTRACT

To quantify the energy performance of a given building, the most uncertain and probably the most influent parameters are the dynamical parameters known as the occupancy rate, the weather data and the temperature set point. Users sometimes encounter gaps in these data, and techniques are needed to estimate variables when data are missing. To obtain most of the dynamic data, measurements are made at least during one year. These measurements are expensive and time consuming. In this paper, we proposed a methodology by means of statistical approaches to build a full year dynamic data files in order to fill gaps according to the measured data features. The methodology is based on statistical approaches, (Iman and Conover 1982) and on the methodology used by *J. Goffart (2013)* to generate weather data. Thus performance indicators are used to evaluate the methodology and its consequences on energy consumption estimation. An office building has been monitored and multiple sensors have been mounted on candidate locations to get needed data. In this paper a methodology to address the problem of missing data in measurement data such as temperature set point data is proposed.

INTRODUCTION

A building energy performance is defined as the amount of energy actually consumed or estimated to meet the different needs associated with a standardized use of the building, which may include heating, cooling, hot water, ventilation and lighting (EU Directive 2002).

To assess this performance, simulation tools and measurements are used. The simulation tools allow the user to represent the building in real conditions and to perform dynamic simulations (BES) in order to obtain the building behavior over a year. The physical phenomena, under which the building is, like the conduction, the convection, the solar radiations as well as its internal loads are all represented and linked together. To ensure that the results match the real energy demand, measurements are made to check if the model represents as close as possible the reality: this process is called calibration.

Calibration consists of comparing measured and simulated data. This process is iterative and tends to improve the model and is making it more reliable. It

consists of the comparison of the model outputs to monitored data to enhance the building model analysis of the results. The first difficulty arises with the intrinsic quality of measured data. It often presents missing data or data are only available on a short period or do not exactly corresponds to calculated ones. In order to analyze the differences between measurements and calculations, two different and complementary methods should be used: graphical and statistical.

The statistical approach consists of calculating mathematical indicators that evaluate if BES results matches with measurements. According to ASHRAE guidelines, three indices can be used to describe the performance of a model: the coefficient of variation of the standard deviation (*CVSTD*), the coefficient of variation of the root mean square error (*CVRMSE*) and the normalized mean bias error (*NMBE*) but for calibrated simulations, only *CVRMSE* and *NMBE* are used.

The Normalized Mean Bias Error (*NMBE*), as stated in EQUATION 1, is a non-dimensional bias measure which shows how close the total energy uses estimated by the simulations match the real energy consumption on an hourly or monthly basis. However, this feature is affected by the cancellation effect, it is then necessary to compute another performance indicator. By evaluating the Coefficient of Variation of Root Mean Square Error (*CVRMSE*) of a data (EQUATION 2), one can determine how well a model fits the data. This indicator gives the offsetting errors between measured and simulated data and does not suffer from the cancellation effect. As stated by, ASHRAE guideline (ASHRAE 2002) and IPMVP (IPMVP 2002), models are admitted to be calibrated if they provide *NMBE* and *CVRMSE* indices respectively within $\pm 10\%$ and $\pm 30\%$ when performing hourly simulations.

$$NMBE (\%) = 100 \cdot \frac{\sum_{i=1}^N (m_i - a_i)}{\sum_{i=1}^N (m_i)} \quad (1)$$

$$CVRMSE (\%) = 100 \cdot \frac{\sqrt{(\sum_{i=1}^N (m_i - a_i)^2 / N)}}{\bar{m}} \quad (2)$$

Where m_i and a_i are respectively measured and simulated data points at each instance i , N is the

number of independent data points and \bar{m} is the average of the measured data point.

Sensitivity and uncertainty analysis are important steps to reach good results when doing a calibration procedure (Tian 2013; Macdonald 2002). Sensitivity analysis is used to identify the key variables affecting the building thermal performance and uncertainty analysis is used to represent the output and the inputs parameters data dispersion. Every physical quantity has an amount of uncertainty associated to its value. In order to obtain representative results one needs to consider all the associated uncertainties in line with the missing data points in the files.

We talk about missing data when we do not have observations about a given variable at some periods in a data file. The presence of missing data is a very big issue and cannot be ignored when assessing a building performance and according to their proportions and types, different solutions will be used to encounter these gaps. The user can take out the variables with missing data in a file or develop methods to handle them. Various methods are used to deal with missing data in Building Energy Simulation (BES). The goal of this research work is to find a suitable methodology to fill the gaps by considering the impact on the energy consumption in order that the selected indicators (*CVRMSE* and *NMBE*) are as low as possible and also to add a minimum of uncertainty to the data being measured.

In this paper, the temperature data coming from a monitored building zone have been used for calibration purpose. Gaps are present in the data files and a suitable method is used to encounter the gaps in order to conserve the autocorrelation of the temperature data over the day and also the inter-correlation between the four offices of the building zone. The content is organized as follow: in the first section is exposed the method used to fulfill the missing gaps. In the following section is presented the monitored building and the zone used for this research work followed in the last sections by the data analysis and the outcomes of the developed methodology.

METHODOLOGY

It has been said above that to evaluate a building energy performance, one need to perform a measurement campaign and to construct a numerical model. The comparison of the measurements to the simulation results helps to ensure of the model consistency. Then, by setting some parameters such as the weather or the temperature set point, the building performance can be evaluated. This confrontation step is essential.

Encounter missing data points in a calibration procedure is a big concern in building sector nowadays (Coakley, Raftery, and Keane 2014). A very easy way to fill the missing data points is to take the mean of the available data by hour and by day

type. This technique can give acceptable results but is not taking into account the variability of a dynamic data file which means that two filled data points with the same hour and the same day in different weeks would have the same value, which is not possible on a realistic point of view.

The idea here is to find a consistent technique to fill the gaps efficiently by considering the variability in time of the temperature data (autocorrelation) and the space dependency of the temperature between offices (inter-correlation). It is known that dynamic data are auto correlated. Temperature at a moment t depends on the temperature at $t - 1$ and can be used to determine the temperature at time $t + 1$. This relevant feature is used in the technique presented here, with other statistical features such as the mean, the data correlation and the distribution function at each time step to generate suitable data to fulfill the measurement files and perform the model calibration. The generation method used here is the Latin Hypercube Sampling (LHS). The LHS method (Stein 1987) is a sampling technique which provides an efficient way of sampling variables from their distributions. The aim of this method is to parse the data into groups to do a representative sampling of the space parameter with a minimum of trials. The parameter range of variation is split into n intervals with the same probability. Then a value is picked up randomly within each interval. Even with a limited number of trials, the generated data will cover the overall parameter space (Iman and Conover 1980; McKay, Beckman, and Conover 1979). This sampling technique is used in this work because it is easy to apply and it offers the possibility to generate correlated data.

To introduce correlation in the sampling data, we used Iman and Conover method (Iman and Conover 1982). It introduces correlation among the variables by restricting the way variables are paired based on the ranked correlation of some target values. For example, to generate consistent dynamic files, Anstett-Collin et al. 2015 considered a random process $x(\theta, t)$ where t represents the time and θ the stochastic part of the variable. The random process can then be associated to its mean quantity $\bar{x}(t)$, to which can be added a random variable $\varepsilon(\theta, t)$ such as stated in EQUATION 3. The mean value is supposed to be deterministic and the random variable is defined by the autocorrelation function and the distribution function extracted before (EQUATION 4). Knowing that the mean value is constant for each hour, the generation of consistent data will depend on the generation of the random value ε (J Goffart 2014).

$$x(\theta, t) = \bar{x}(t) + \varepsilon(\theta, t) \quad (3)$$

$$\varepsilon(\theta, t) \begin{cases} C(t, t - 1) \\ r(t) \end{cases} \quad (4)$$

where $C(t, t - 1)$ is the correlation function and $r(t)$ the distribution function at each time step t . To

generate the values associated to $\varepsilon(\theta, t)$, one has to compute the Cholesky decomposition of the correlation matrix. Matrix C can be written as $C = PP'$ where P is the lower triangular matrix. After having generated the matrix G with Latin Hypercube Sampling, one calculates the correlation matrix T of G . The objective here is to rearrange the generated values closed to the correlation matrix (Iman and Conover 1982). Then, the result is obtained by adjusting the elements of G according to its correlation matrix T in order that G and the prior data set have the same correlation matrix $\varepsilon(\theta, t)$. The process is repeated until the user obtained a closed correlation matrix.

Finally, a data point is picked up randomly among the generated data to fill the missing data points until all the gaps are fulfilled.

BUILDING DESCRIPTION AND SIMULATIONS

The monitored building is part of the Cerema lab in Angers. The building has been partially built in 1961 (Figure 1). Since its construction, it has been extended and renewed, a second floor has been added in 2009. Due to the cost of measurements, the sensors have not been placed in the whole building, the building was split in different monitored zones.

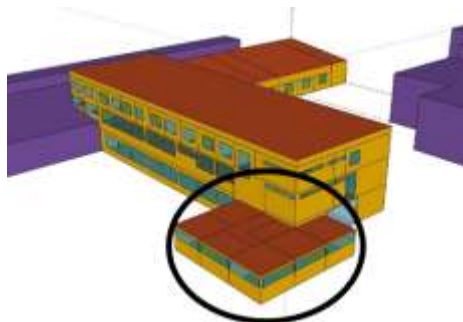


Figure 1: 3D Representation of the monitored building

The monitored zone (Figure 2) is a part of the ground floor of this administrative building. It is constituted of 6 offices with a circulation area and a gross floor area of 102 m². This part of the building is made of concrete block and is isolated only at the roof. Most of the building occupancy occurs between 8:00 and 18:00. The sensors have been fixed in all the offices to get the temperature data of the overall zone. The temperature data are collected during one year from January to December 2013 in each office with a time step of one hour.

The average wind speed, horizontal solar radiations, and outdoor temperature are also measured each hour. Occupancy has been measured through doors and windows openings lightning and plugs loads by the installation of infrared motion detection. Regarding the test-bed, each sensor is independent



Figure 2: The monitored zone

and works with batteries which allow them to have their own memory to record the data. The sensors have been placed on carefully chosen locations, on a half way up and away from openings (doors and windows).

The building has been modeled in TRNsys 17 (Trnsys 2000). The building geometry has first been drawn in Sketchup 3D and has been integrated in type 56 under TRNsys. The model zones was defined accordingly to the measurement process and took into account the airflow exchanges with Contam, the solar masks, the single flow ventilation system and the heat exchanges with the ground. The dynamic parameters such as the occupancy rate, lightning and plug loads have also been integrated to the model.

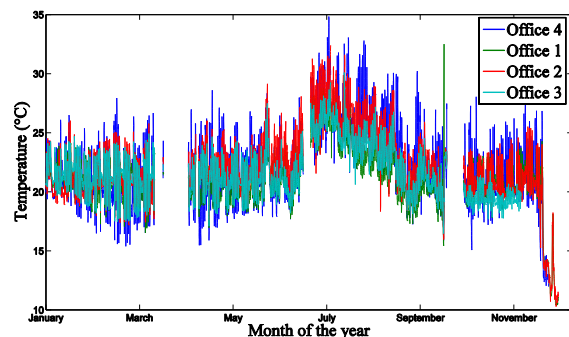


Figure 3: Missing Temperature data in office 4, 5, 6 and 7

DATA ANALYSIS

In the measurement process, temperature, electrical

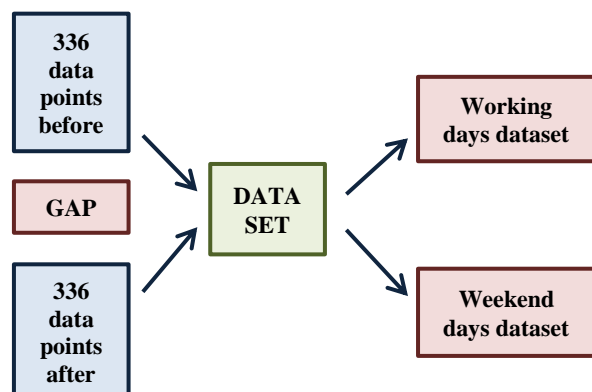


Figure 4: Dataset construction for LHS sampling

Table 1
The missing data points in each office

Date	Number of data point	Office 4	Office 5	Office 6	Office 7
From 06/02 to 07/02	11	X	X	X	X
From 10/03 to 11/03	10	X	X	X	X
From 19/03 to 25/03	144	X	X	X	X
From 25/03 to 12/04	424	X	X	X	X
From 17/04 to 19/04	43	X	X	X	X
From 03/07 to 08/07	122	X	X	X	X
From 22/07 to 23/07	11	X	X	X	X
From 06/08 to 07/08	12	X	X	X	X
From 13/10 to 25/10	282	X	X	X	X
From 05/12 to 01/01	164			X	X

consumption and occupation data were missing at some periods, in the monitored area. Due to some concerns in the measurement process (occupant's vacation period and for some unknown causes), we find out that data are missing in offices 2 and 3 from April to December and partially missing in the other offices. The missing data issue in the first two offices is due to measurement process and due to time issue; a suitable method was not been find yet to fill the missing data files in these offices. Then, the analysis which has been done is just focused on the temperature data in the office 4, 5, 6 and 7 (Figure 3). In Table 1 is dressed the list of the intervals gaps in each office. Data are on hourly basis and a cross is placed in a case when data are missing in an office during the referred interval.

The first step in this technique is to find all the gaps in the files. The offices 4 and 5 have 9 consistent gaps and the offices 6 and 7 have one bigger gap at the end of the year. To make the calculations easier, the gaps were homogenized in 10 gaps for all the offices and the gaps were filled only when the data are missing.

Thus, for each missing data point inside each gap, we evaluate if the missing point is a working day or not in order to separate working days from weekends. Then, with the selected gap, we construct two temperature data files (one for the working days and the other one for the week ends) with two weeks after and two weeks before the gap in order to have a representative data set for the extraction of the statistical features (Figure 4).

Table 2

Inter-correlation between offices before data correction

	Office 4	Office 5	Office 6	Office 7
Office 4	1	0.7249	0.7310	0.6809
Office 5	0.7249	1	0.8328	0.7894
Office 6	0.7310	0.8328	1	0.8426
Office 7	0.6809	0.7894	0.8426	1

This data set will then be representative enough to extract the needed features. The features we extracted are the mean temperature M_j ; the standard deviation S_j and the autocorrelation matrix C_{jj} at each time

step (EQUATIONS 5, 6 and 7). This stage is essential so that with these features, one can generate daily random data files based on the selected period and their appropriate characteristics.

$$[M_4 \ M_5 \ M_6 \ M_7] \quad (5)$$

$$[S_4 \ S_5 \ S_6 \ S_7] \quad (6)$$

Then, depending on the hour of the day, and with the extracted statistical features, one generates n correlated data (EQUATION 8).

$$\begin{bmatrix} C_{44} & C_{45} & C_{46} & C_{47} \\ C_{54} & C_{55} & C_{56} & C_{57} \\ C_{64} & C_{65} & C_{66} & C_{67} \\ C_{74} & C_{75} & C_{76} & C_{77} \end{bmatrix} \quad (7)$$

$$\begin{bmatrix} T_{4_1} & T_{5_1} & T_{6_1} & T_{7_1} \\ T_{4_2} & T_{5_2} & T_{6_2} & T_{7_2} \\ \vdots & \vdots & \vdots & \vdots \\ T_{4_i} & T_{5_i} & T_{6_i} & T_{7_i} \end{bmatrix} \quad (8)$$

The above process is resumed in Figure 5. In the first try, we attempted to fill the temperature files of the offices separately but we noticed that the inter-correlation was not well-maintained but data autocorrelation was intact. An example of temperature autocorrelation during a day in office 4 is shown in Figure 6. The correlation is decreasing

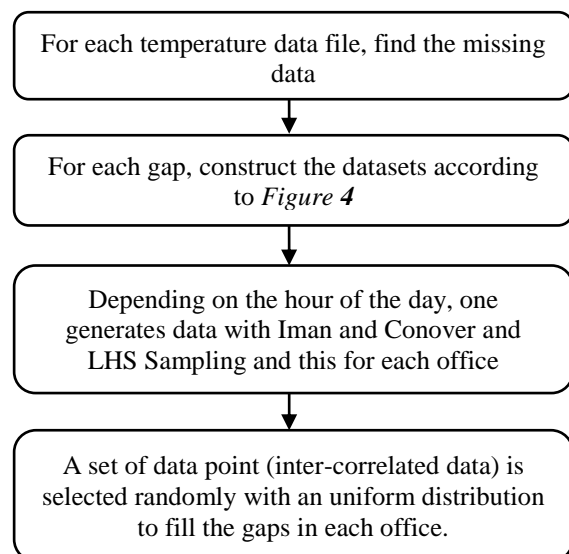


Figure 5: Missing data filling flow chart

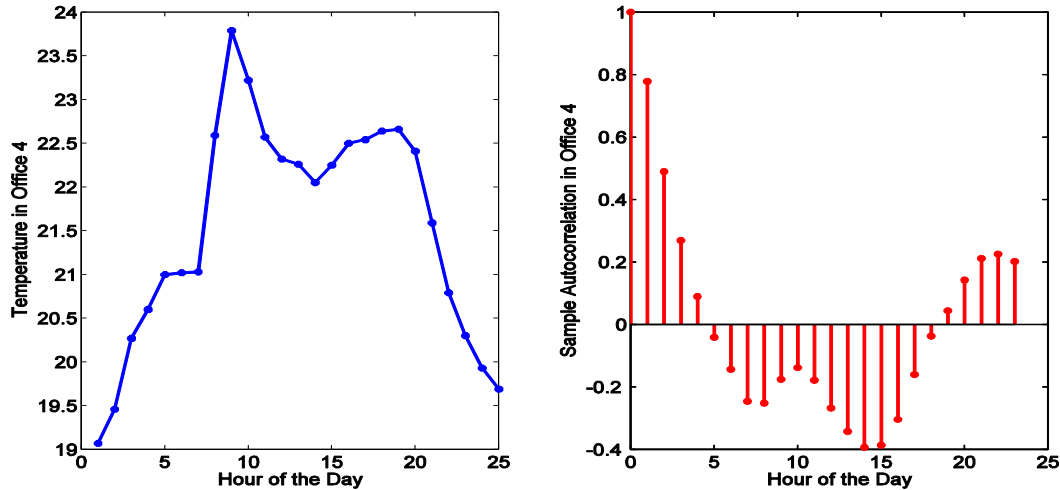


Figure 6: Temperature auto-correlation of a sampled day in office 4

and kept in the data during 5 hours. Correlated data in dynamic systems are usually correlated during few hours. Then, in the second test, the four offices gaps are filled simultaneously in order to keep the inter-correlation (Table 3 and Figure 9).

RESULTS AND DISCUSSION

Visual verification of the generated data is not enough to affirm that the generated data are consistent. Temperature inter-correlation matrix (TABLE 2 and Table 3) is compared to the initial values in order to prove that the generated data is in agreement with the prior data set extracted features. The correlation coefficients after the data correction are little bit higher than the prior data, which could be explained by the developed algorithm consistency and by the use of the autocorrelation features at once for the data generation.

$$I_1 (^\circ\text{C}) = \sum_{i=1}^N (a_i - m_i) \quad (9)$$

$$I_2 (^\circ\text{C}) = \sqrt{\left(\sum_{i=1}^N (m_i - a_i)^2 / N\right)} \quad (10)$$

In order to view the data dispersion and consistency and by the way its influence and propagation on the energy consumption by simulations, the indicators I_1 and I_2 (EQUATIONS 9 and 10) have been calculated for the generated temperature data. To better sense the success of the methodology used, the selected indicators have been computed only on the winter period (From October 1st to March 31st of 2013).

Table 3
Correlation between offices after temperature data correction

	Office 4	Office 5	Office 6	Office 7
Office 4	1	0.8538	0.8258	0.7354
Office 5	0.8538	1	0.8870	0.8374
Office 6	0.8258	0.8870	1	0.8163
Office 7	0.7354	0.8374	0.8163	1

The generated temperature data I_1 is in between $[-100, 100]^\circ\text{C}$ and the Coefficient of Variation I_2 does not exceed 0.03°C in each office (Figure 7 and Figure 8).

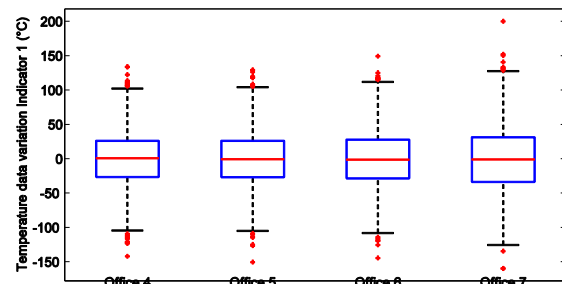


Figure 7: Temperature data Indicator 1 in the winter period ($^\circ\text{C}$)

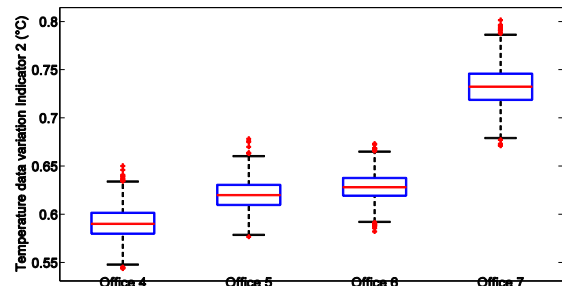


Figure 8: Temperature data Indicator 2 in the winter period ($^\circ\text{C}$)

To evaluate the influence of the generated data on the energy consumption of the monitored zone, 4000 simulations have been performed on TRNSys with in input the 4000 rectified data files for the offices 4, 5, 6 and 7. The other parameters are unchanged during the simulations. Then the NMBE and the CVRMSE have been estimated for the simulations for each office during all the interval gaps within the heating period.

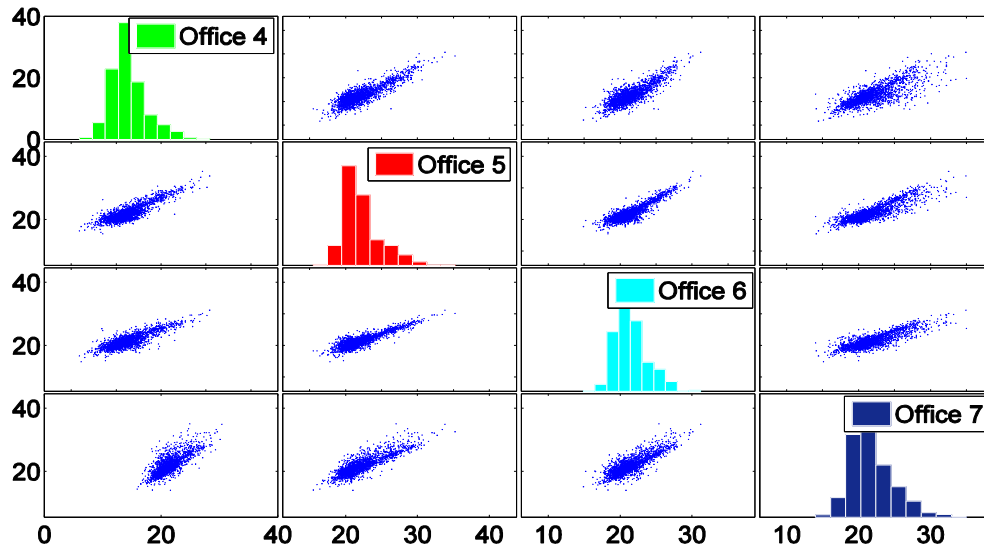


Figure 9: Inter-correlation between offices

The simulations process is represented in EQUATION 11, where j , is the number of simulations, i the number of hours, a_{ij} the considered energy consumption for a given simulation j and m_i the measured data point at a given hour i . To obtain the

NMBE of a given simulation j , the difference between the j^{th} simulated energy use and the mean energy use over the 4000 simulations for all the hours is computed, summed and divided by the sum of this energy use on the same period like stated in EQUATION 1. Regarding the *CVRMSE*, the error between the considered simulation and the mean value is calculated then squared and added on the whole period and divided by the respective number of points (EQUATION 2).

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1,4000} \\ a_{21} & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \vdots & \vdots & a_{ij} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{8760,1} & \dots & \dots & a_{8760,j} & \dots & a_{8760,4000} \end{bmatrix} \begin{bmatrix} m_1 \\ \vdots \\ m_i \\ \vdots \\ m_{8760} \end{bmatrix} \quad (11)$$

In Table 4 is presented the results of the energy consumption NMBE. The NMBE is essentially near zero in all the considered intervals in the zone. The deviations of the NMBE gives also very low results which can be explained by the compensation effect with the prior data.

The Figure 11 represents the energy consumption based on the *CVRMSE* of the 4000 simulations in the four offices. The *CVRMSE* is decreasing in all the offices when the amount of energy consumption is growing. This implies that the data correction gives better results in heating period and is not very preferment when the building is not being heated.

These results can similarly be observed in Table 5 in the office 4 for example. The *CVRMSE* is lower in winter intervals (1, 2 and 6) and is higher in the intervals within the warmer seasons.

Table 4
Energy Consumption NMBE of the overall zone

N°	Interval	NMBE (%)	NMBE Standard deviation
1	06/02 to 07/02	-0.12	3.51
2	10/03 to 11/03	-0.4	6.43
3	19/03 to 25/03	-0.025	1.58
4	25/03 to 12/04	-0.005	0.72
5	13/10 to 25/10	-0.028	1.67
6	05/12 to 31/12	-0.0012	0.34
7	30/03 to 01/10	$-9.82 \cdot 10^{-5}$	0.09

Table 5
Energy Consumption *CVRMSE* by intervals of each office (%)

N°	Office 4	Office 5	Office 6	Office 7
1	17.58	11.1	15.71	19.33
2	25.73	24.31	26.43	35.25
3	40.55	28.22	32.21	46.8
4	36.4	25.1	27.7	41.18
5	72.96	39.57	69.98	63.33
6	23.49	18.54	33.67	21.7
7	16.76	16.73	18.01	16.81

CONCLUSION

The use of metered data in building energy simulation is necessary in order to calibrate a model but sometimes, some parts of the data are not recorded well or data are totally missing during some periods. Then, the user has to find an adequate method to fulfill the gaps in order to have more reliable data files with the lowest uncertainty as

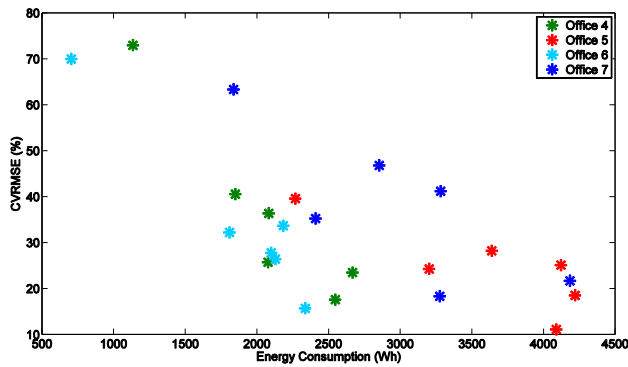


Figure 11: CVRMSE vs. Energy Consumption in each office by intervals

possible. The method proposed above takes into account the statistical features of the prior data i.e. the hourly dependence of the data, its inter-correlation with the other variables and produce suitable data for the missing parts.

In Figure 10 is represented the NMBE and the CVRMSE of the overall zone during the heating periods over a year. With the fact that the energy consumption NMBE indicator is always around zero and the CVRMSE within [10, 20] %, the method could be considered to be efficient, but it seems to not being efficient in the seasons without heating. When the performance indicator is the CVRMSE, the developed methodology can be assumed to not being reliable: then, the user needs to do measurements during the whole year in order to have consistent data. But if we consider only the NMBE as the performance indicator, the results are assumed to be acceptable. Still, since the NMBE is affected by the cancelation effect (ASHRAE 2002), we can assume that this approach cannot be used in order to do measurements during some periods and fill the other parts of the year with generated data. Another perspective will be to use the correlation effect between the building parameters such as the occupants, the indoor temperature and the internal loads, to generate data from one variable to another one.

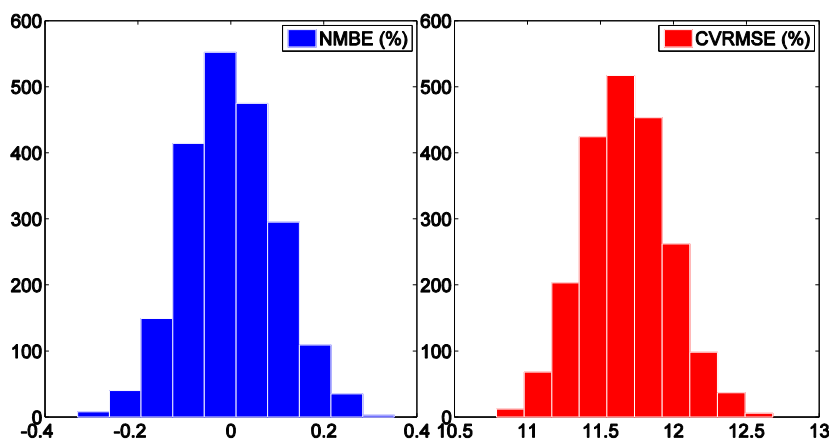


Figure 10: Energy Consumption NMBE and CVRMSE of the overall zone on the heating period

NOMENCLATURE

a	= simulated data
m	= measured data
n	= number of intervals
N	= number of data points
t	= time
$\bar{x}(t)$	= a mean quantity
ε	= random variable
θ	= stochastic part of the variable x
C	= correlation function
r	= distribution function
P	= lower triangular matrix
G	= generated matrix with LHS
T	= correlation matrix of G
M_j	= mean of data in office j
S_j	= standard deviation of data in office j
T_{ij}	= i^{th} generated temperature point in office j
NMBE	= Normalized Mean Bias Error
CVRMSE	= Coefficient of Variation of Root Mean Square

ACKNOWLEDGMENT

This research work is funded by the grants from the National Association of Research (ANR) in France within the OMEGA project and the Research Center of Risk, Expertise, Environment and Mobility (CEREMA) provides the data.

BIBLIOGRAPHY

- Anstett-Collin, F., J. Goffart, T. Mara, and L. Denis-Vidal. 2015. "Sensitivity Analysis of Complex Models: Coping with Dynamic and Static Inputs." *Reliability Engineering & System Safety* 134: 268–75. doi:10.1016/j.res.2014.08.010.
- ASHRAE. 2002. "ASHRAE Guideline 14, Measurement of Energy and Demand Savings."

- Coakley, Daniel, Paul Raftery, and Marcus Keane. 2014. "A Review of Methods to Match Building Energy Simulation Models to Measured Data." *Renewable and Sustainable Energy Reviews* 37: 123–41.
- EU. 2002. "Directive of the European Parliament and of the Council of the 16 December 2002." <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2003:001:TOC>.
- Iman and Conover. 1982. "A Distribution-Free Approach to Rank Correlation.pdf."
- Iman, Ronald L., and W. J. Conover. 1980. "Small Sample Sensitivity Analysis Techniques for Computer Models with an Application to Risk Assessment." *Communications in Statistics - Theory and Methods* 9 (17): 1749–1842.
doi:10.1080/03610928008827996.
- IPMVP. 2002. "International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings; Volume I: Revised March 2002 - 31505.pdf."
- J Goffart. 2014. "Méthode D'échantillonnage Pour La Prise En Compte Des Incertitudes Des Données Météorologiques Pour L'analyse de La Sensibilité Globale Sur La Performance Énergétique." *IBPSA*.
- Macdonald, Iain Alexander. 2002. "Quantifying the Effects of Uncertainty in Building Simulation." University of Strathclyde.
- McKay, M. D., R. J. Beckman, and W. J. Conover. 1979. "Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics* 21 (2): 239–45.
doi:10.1080/00401706.1979.10489755.
- Stein, Michael. 1987. "Large Sample Properties of Simulations Using Latin Hypercube Sampling." *Technometrics* 29 (2): 143–51.
doi:10.1080/00401706.1987.10488205.
- Tian, Wei. 2013. "A Review of Sensitivity Analysis Methods in Building Energy Analysis." *Renewable and Sustainable Energy Reviews* 20: 411–19. doi:10.1016/j.rser.2012.12.014.
- Trnsys, A. 2000. "Transient System Simulation Program." *University of Wisconsin*.

Thèse de Doctorat

Fally TITIKPINA

Optimisation des plans de mesure dans un contexte de garantie de performance énergétique des bâtiments

Caractérisation des incertitudes de mesure et propagation au sein de modèles dynamiques

Measurement plans optimization in a context of buildings energy performance warranty

Measurement uncertainties characterization and propagation within dynamic models

Résumé

Ces travaux de thèse s'inscrivent dans le contexte des contrats de garantie de performance énergétique développés afin de réduire l'écart entre les consommations d'énergie effective des bâtiments et celles prévues en phase de conception. Cette garantie exige l'évaluation de la performance énergétique en tenant compte des incertitudes associées aux grandeurs d'entrée notamment celles liées à l'usage. Des mesures sont ainsi mises en place pour affiner la connaissance sur ces grandeurs, mais aussi pour ajuster les grandeurs dynamiques telles que les variables météorologiques et celles liées à l'usage. Plusieurs scénarios de mesure peuvent être envisagés, et, tout en souhaitant en minimiser le coût, on cherchera à dimensionner un plan de mesure avec un niveau de complexité adapté à la précision des résultats escomptés. L'objectif de cette thèse est de proposer une méthodologie de sélection d'un plan de mesure et de vérification adapté, en phase d'exploitation du bâtiment. Les incertitudes des grandeurs d'entrée considérées sont tout d'abord estimées en tenant compte de leurs singularités et de la spécificité des mesures mises en place pour les quantifier. Un modèle de simulation thermique dynamique a été construit à cet effet. Les grandeurs les plus influentes sont ensuite déterminées pour réduire la dimension stochastique du modèle et déterminer le poids de chaque source d'incertitude sur la consommation calorifique annuelle. Une méthode de propagation d'incertitudes et d'évaluation rapide du modèle physique et probabiliste par les polynômes du chaos est alors proposée. Puis, les différents plans de mesure considérés ainsi que l'incertitude associée à leur mise en place respective sont exposés. Pour finir, une stratégie de sélection de l'option optimale basée sur le risque visé dans un contrat de garantie de performance énergétique est proposée. Afin d'illustrer la démarche proposée, la méthodologie a été déployée sur un cas réel.

Mots clés

Garantie de performance énergétique, plans de mesure, incertitudes, sensibilité.

Abstract

These research works fall in line with the energy performance warranty leases developed to reduce the performance gap between the real energy consumption and the predicted energy consumption at the building design stage. This warranty required the evaluation of the energy performance with the uncertainties of each input variables involved, especially those related to occupancy. Measurements are made to improve knowledge about them on one hand, and to adjust the dynamic variables such as the meteorological variables and those linked to occupancy on the other hand. Multiple measurement scenarios can be considered, and, due to their cost, they will be designed with a level of complexity adapted to the accuracy of expected results. The aim of this research work is to provide a methodology to select the suitable measurement and verification plan during the occupancy period. The uncertainty associated to the involved input variables are first estimated by taking into account their features and the specificity of the measurement campaign. For this purpose, a dynamic simulation model was built. The key variables are identified to reduce the size of the stochastic model and also to determine the contribution of each source of uncertainty on the annual heat consumption. Then, a method of propagation of uncertainty based on the rapid evaluation of the physical and probabilistic model by polynomial chaos expansion is proposed. The considered measurement plan and the uncertainty associated to their implementation are presented. Finally, a selection strategy based on the targeted risk in an energy performance warranty contract is suggested. In order to illustrate the proposed approach, the methodology was applied on a real case.

Key Words

Energy performance warranty, measurements, uncertainties, sensitivity.