



**HAL**  
open science

# Etude d'outils de modélisation des systèmes socio-techniques

Charlie Dworaczek

► **To cite this version:**

Charlie Dworaczek. Etude d'outils de modélisation des systèmes socio-techniques. Modélisation et simulation. 2020. hal-03024784

**HAL Id: hal-03024784**

**<https://inria.hal.science/hal-03024784>**

Submitted on 26 Nov 2020

**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.

*Inria*



---

---

# ÉTUDE D'OUTILS DE MODÉLISATION DES SYSTÈMES SOCIO-TECHNIQUES

---

---

Charlie DWORACZEK

Encadrants : Guillaume Mandil, Jean-Yves Courtonne

Avril-Septembre 2020

Rapport de stage M2  
soutenu le 16 septembre 2020



# Table des matières

<b>1</b>	<b>État des connaissances sur le sujet</b>	<b>4</b>
1.1	Métabolisme social . . . . .	4
1.1.1	L'approche MuSIASEM . . . . .	5
1.2	Vulnérabilité . . . . .	7
1.2.1	Analyse de vulnérabilité . . . . .	7
1.2.2	Discussion sur les définitions possibles et éléments de modélisation . . . . .	8
1.2.2.1	Un exemple . . . . .	9
1.2.2.2	Éléments à modéliser . . . . .	9
1.2.2.3	Distinction entre les types d'aléas . . . . .	10
1.2.2.4	Exposition, sensibilité, résilience . . . . .	10
1.2.2.5	Retour sur l'exemple . . . . .	12
1.3	Ingénierie système : Langage SysML et SADT . . . . .	12
1.3.1	SADT . . . . .	12
1.3.1.1	Éléments de modélisation . . . . .	13
1.3.2	Langage SysML . . . . .	14
1.3.2.1	Diagrammes fonctionnels . . . . .	15
1.3.2.2	Diagramme structurels . . . . .	15
1.3.2.3	Diagramme comportementaux . . . . .	16
<b>2</b>	<b>Contribution personnelle : outils de représentation systémique pour les SST</b>	<b>18</b>
2.1	Inspiration des diagrammes de Sankey et SADT . . . . .	18
2.1.1	Les diagrammes de Sankey . . . . .	18
2.1.2	Les diagrammes SADT . . . . .	19
2.2	Adaptation du langage SysML pour l'étude de la vulnérabilité . . . . .	21
2.2.1	Les diagrammes structurels . . . . .	21
2.2.2	Les diagrammes fonctionnels . . . . .	23
2.2.3	Le diagramme d'états . . . . .	24
2.2.4	Le diagramme de séquence . . . . .	27
2.3	Discussion, perspectives . . . . .	29
<b>3</b>	<b>Réconciliation de données</b>	<b>30</b>
3.1	Contraintes linéaires et erreurs gaussiennes . . . . .	30
3.1.1	Cadre théorique . . . . .	30

<i>TABLE DES MATIÈRES</i>	3
3.2 Contraintes non-linéaires et erreur non-gaussienne . . . . .	35
3.2.1 Cadre et hypothèses . . . . .	36
3.2.2 Rappels de géométrie différentielle et riemannienne . . . . .	36
3.2.3 Explication de la méthode . . . . .	36
<b>A Rappels de géométrie différentielle et riemannienne</b>	<b>41</b>
<b>Bibliographie</b>	<b>45</b>

## Résumé

Un système socio-technique est un réseau liant des acteurs avec des liens de type économiques, matériels, culturels ou encore sociaux. Le changement climatique nous invitant à une grande modification de notre société, l'imagination de nouveaux systèmes, appelées alternatives, socio-techniques est une problématique particulièrement actuelle. L'équipe STEEP, dans laquelle j'ai effectué mon stage, s'intéresse à l'étude des alternatives socio-techniques (AST) et notamment de leur valeur selon plusieurs critères. J'ai pu travailler sur l'un d'eux, la vulnérabilité d'une AST qui désigne la propension de cette dernière à subir des dommages. Afin d'étudier cet aspect, la représentation et modélisation des systèmes est une étape importante sur laquelle j'ai pu largement me consacrer.

Mon travail a été constitué de deux parties toutes deux portant sur la modélisation des AST, une première partie qualitative et une deuxième partie quantitative. J'ai premièrement effectué une revue des notions de l'approche systémique comme celle du métabolisme social ou de la vulnérabilité. J'ai ensuite pu adapté certains outils rencontrés dans l'ingénierie système pour la représentation de la vulnérabilité des AST. Ce travail a permis de réfléchir à des questions modélisation, de savoir ce qui peut facilement être représenté et ce qui l'est un peu moins. Dans un second temps, j'ai également pu m'intéresser au problème de la réconciliation des données, domaine à la jonction entre les mathématiques appliquées et l'approche systémique, connexes à des branches des mathématiques très intéressantes comme la géométrie riemannienne, branches avec lesquels j'ai du me familiariser afin de comprendre les outils utilisés. J'ai effectué une synthèse d'articles que j'ai pu lire où je détaille le formalisme mathématique des méthodes présentés par ces articles.

# Introduction

Face à l'urgence climatique, les Accords de Paris ont dressé des exigences pour ne pas atteindre la barre des + 1,5°C d'ici à 2100. En effet, un tel réchauffement aurait des conséquences désastreuses sur la faune et la flore, nos rendements agricoles et plus généralement nos sociétés. Parmi ces exigences, l'empreinte carbone annuelle moyenne d'un Français (regroupant nos émissions directes de gaz à effet de serre mais aussi nos émissions indirectes comme celles qui proviennent de nos choix de consommation) devrait, d'ici 2050, afin d'atteindre la neutralité carbone nationale, être au maximum de 2 tonnes de CO<sub>2</sub>e<sup>1</sup> par an. Étant actuellement autour de 10,8 tonnes, cette consommation individuelle doit baisser de 80% d'ici 30 ans. Seulement peut-on vraiment s'attendre à une telle baisse ?

Certaines actions individuelles, comme le changement du type de locomotion ou un régime alimentaire plus végétal, sont clairement réalisables mais même avec les hypothèses les plus optimistes, le rapport du cabinet de conseil Carbone 4 de 2019 dresse une baisse moyenne maximale de l'ordre de 20% et une baisse moyenne plus probable de l'ordre de 5 à 10% pour un individu. Selon ce rapport, les actions individuelles ne permettront donc pas d'atteindre les objectifs des Accords de Paris. Que faire alors pour respecter ces objectifs ?

Il n'est vraisemblablement pas réaliste d'atteindre ces objectifs en se reposant seulement sur les efforts individuels des citoyens, bien qu'ils soient d'importance non-négligeable. La mise en place de politiques pourraient permettre de rendre cette transition plus facile, par exemple la facilitation de la rénovation des logements du parc immobilier français qui représente aujourd'hui

---

1. CO<sub>2</sub> équivalent. Dans nos émissions de gaz à effet de serre, il n'y a pas seulement du dioxyde de carbone, mais aussi du méthane, du protoxyde d'azote... Ces différents gaz ont un effet différent sur le réchauffement climatique via leurs durées de séjour dans l'atmosphère et leurs forçages radiatifs (qui dépend de la longueur de la période considérée). Par des processus biophysiques, ces gaz sont ensuite réabsorbés par la biosphère, permettant ainsi un refroidissement de la température globale. Le dioxyde de carbone CO<sub>2</sub> et le protoxyde d'azote N<sub>2</sub>O ont une durée de séjour dans l'atmosphère de l'ordre du siècle, le méthane CH<sub>4</sub> de l'ordre de la décennie et la vapeur d'eau de l'ordre de quelques jours seulement. Néanmoins tous ces gaz, n'ont pas le même forçage radiatif, il est donc nécessaire de pouvoir comparer les différents "pouvoirs de réchauffement global", aussi appelé PRG. Si F<sub>g</sub>(t) est le forçage radiatif d'un gaz g au temps t et T la durée de la période considérée, ce pouvoir de réchauffement global se calcul par la formule suivante :

$$PRG_g(T) = \frac{\int_0^T F_g(t) dt}{\int_0^T F_{CO_2}(t) dt}$$

Pour obtenir le CO<sub>2</sub>e/an d'un gaz, il faut calculer 0.2727 × PRG<sub>g</sub>(T) × m<sub>e</sub>(g) où T est en général égal à 100 ans et m<sub>e</sub>(g) est la masse émise du gaz g. La constante correspond quant à elle à la masse de carbone dans 1 kg de CO<sub>2</sub>.

autour une émission de 3 tonnes de  $CO_2e$  par an<sup>2</sup> (en 2006) et qui pourrait être réduite de 1,2 tonnes de  $CO_2e$  par an. Plus généralement, des mesures nationales comme la réduction des importations (représentant 57% des émissions en 2018<sup>3</sup>) représenteraient un grand pas en avant vers le respect des Accords de Paris.

Plus généralement, d'après le rapport Carbone 4, l'État est le seul acteur, plus que les citoyens et les entreprises, à pouvoir réguler ces émissions. Il y possède un rôle crucial par sa capacité à inciter financièrement et réglementairement les choix des acteurs en faveur d'une société avec moins d'émissions. Il est de plus le seul à pouvoir négocier avec l'Europe, acteur également crucial. Les solutions existent et il est nécessaire d'employer tous les moyens à notre disposition ainsi que de rendre possible les efforts de chacun des acteurs. Pour cela, il faut envisager le problème d'un point de vue systémique.

***”Chaque individu est limité par le système socio-technique<sup>4</sup>, c’est à dire l’environnement social et technique dont il dépend”<sup>5</sup>.***

Le problème du changement climatique est d'ordre systémique. Il ne peut pas être uniquement résolu par une somme d'action individuelles. On peut prendre l'exemple de la mobilité pour le montrer. Même si, dans les années à venir le covoiturage systématique et l'utilisation principale des transports en commun comme moyens de locomotion sont mis en place, notre société et son système de transport sont toujours dépendants d'énergies fossiles très polluantes. Il est nécessaire de supprimer notre dépendance à ces énergies carbonées sans quoi aucun réel progrès ne pourrait vraisemblablement être réalisé. Ainsi une modification de notre système de transport, et plus généralement de notre société toute entière, se doit de voir le jour.

L'atteinte de ces objectifs passe donc par l'imagination et la conception d'alternatives socio-techniques répondant à certains critères, notamment un niveau d'émission de GES<sup>6</sup> bien plus bas et donc à beaucoup d'autres critères d'ordre technologique, économique ou encore social. Ces critères comprennent par exemple l'utilisation de moins d'énergies carbonées et de plus de technologies telles que les dispositifs de captage et de stockage de  $CO_2$ <sup>7</sup>, ou encore une économie moins tournée vers les échanges internationaux et une consommation plus composée de production locale. L'éradication de la pauvreté et la réduction des inégalités pourraient aussi représenter des passages obligés pour cette transition vers ces alternatives<sup>8</sup>.

Le monde de la recherche scientifique a clairement un rôle, sinon le rôle principal, à jouer dans l'analyse de ces systèmes car elle est capable de fournir un cadre rigoureux d'évaluation des alternatives. L'équipe STEEP<sup>9</sup>, au sein de laquelle j'ai effectué mon stage, concentre une partie de ses nouvelles recherches sur «les modalités de construction et d'évaluation d'alternatives

2. *Les émissions directes de  $CO_2$  des ménages selon leur localisation*, Commissariat général au développement durable 08/2012

3. *L'empreinte carbone des Français reste stable*, Commissariat général au développement durable, 01/2020

4. Réseau d'acteurs liés par des types de liens variés, techniques comme socio-culturels

5. *Faire sa part ? Pouvoir et responsabilité des individus, des entreprises et de l'État face à l'urgence climatique*, Carbone 4, 06/2019

6. Gaz à effet de serre

7. Connu sous le nom de BECCS, BioÉnergie avec Captage et Stockage du Carbone

8. *Résumé à l'intention des décideurs : Réchauffement planétaire de 1,5°C*, GIEC, 2018, p23

9. Sustainability, Transition, Environment, Economy and local Policy



socio-techniques, en se basant sur ses travaux antérieurs dans le domaine de la modélisation et de l'évaluation environnementale». Les alternatives étudiées se situent à des échelles nationales et infranationales et leur but est double. Il est premièrement, d'assister les décideurs en fournissant des outils de simulation, d'optimisation et de visualisation, des outils qui permettraient de fournir une part de la réponse à la question «*comment atteindre les objectifs environnementaux à l'échelle locale ?*», et deuxièmement de contribuer à l'information d'un public plus large sur les conséquences de diverses alternatives et les compromis devant être tranchés.

J'ai pu, lors de mon stage, travailler sur des questions de modélisations de SST. Une des problématiques principales de mon stage fut de développer et d'adapter des outils permettant, à des experts d'un domaine, de fournir une description détaillée, claire, fidèle à la réalité et en représentant le plus grand nombre d'aspects propres au système, notamment les aspects sociaux. Un aspect original de mon stage fut justement d'adapter des outils de sciences de l'ingénieur à des systèmes non-purement techniques, comprenant une dimension sociale importante.

Un autre aspect de mon travail, en complément de la démarche qualitative précédente, fut de m'intéresser à la notion de réconciliation des données, un outil de modélisation quantitative des flux traversant un ST. La démarche de réconciliation de données est assez naturelle lorsque l'on effectue des mesures (avec donc une incertitude) et que l'accès à la totalité des grandeurs décrivant notre système n'est pas permise.

Dans la première partie, j'effectue une synthèse des connaissances que j'ai pu acquérir sur les notions de métabolisme territorial et d'analyse systémique, notamment : d'analyse de vulnérabilité d'un système et des outils d'ingénierie système que j'ai utilisé.

Dans la deuxième partie, je décris les outils sur lesquels j'ai travaillé pour pouvoir modéliser un système socio-technique.

Dans la troisième partie, je décris mon travail sur deux articles présentant des méthodes de réconciliation des données. Le premier article présente un cadre linéaire gaussien, permettant de réconcilier les données à l'aide d'outils d'algèbre linéaire. L'autre article présente un cadre beaucoup plus général, autorisant un cadre non linéaire avec des incertitudes non-nécessairement gaussiennes, en s'appuyant sur des outils de géométrie riemannienne et des méthodes MCMC. Mon travail a consisté à décrire le formalisme mathématique derrière ces articles.

Enfin j'effectue une conclusion générale sur mon stage.

# Chapitre 1

## État des connaissances sur le sujet

Dans cette partie, je vais effectuer une synthèse de ce que j'ai pu apprendre sur les notions d'analyse systémique, ce qui a été une part très important de mon stage.

Je présente les outils d'analyse systémique sur lesquels j'ai travaillé. Je présente en premier l'approche MuSIASEM pour étudier le métabolisme social, puis je traite du critère de vulnérabilité pour évaluer une alternative socio-technique. Enfin je termine par la présentation d'outils tirée de l'ingénierie système, les langages SysML et SADT. Ces outils ont tenu un rôle central dans mon travail. Ils m'ont notamment permis d'apporter une contribution personnelle que je détaille dans le deuxième chapitre.

### 1.1 Métabolisme social

Commençons par définir la notion de métabolisme social. Il s'agit de tous les flux de matière et d'énergie, qui ont lieu entre une société ou système avec son environnement, réalisés sous une forme culturelle spécifique. Cette notion est à la frontière entre les sciences sociales, notamment les sciences économiques ; les sciences technologiques et les sciences de l'environnement. Comprendre le métabolisme d'un système, c'est avoir une bonne compréhension du fonctionnement de celui-ci. En effet, on peut alors se figurer de quoi le système "est-il fait", sur quoi base t-il son fonctionnement ? Sur quelles ressources ? Sur quels flux ? Cela permet aussi de comprendre quelles sont les fonctions de ce système. Quelles sont ces forces ? Quelles sont ces faiblesses ?

Pouvoir répondre à ces questions permet alors d'évaluer le système et de le comparer à d'autres sur des critères fixés au préalable. Ceci est, comme mentionné dans l'introduction, la raison pour laquelle l'analyse systémique est utilisée dans ce cadre, elle nous permet de pouvoir préférer rationnellement et objectivement un système socio-technique à un autre (penser au taux d'émission de GES d'une société humaine ou le rendement d'un système agricole par rapport à un autre).

Mais comment faire une telle analyse d'un système ?

### 1.1.1 L'approche MuSIASEM

Pour effectuer une telle analyse des outils ont été développés, j'ai pu étudier l'approche MuSIASEM pour me familiariser avec l'analyse systémique d'un système.

#### Définition 1.1.1

L'approche **MuSIASEM**, signifiant *Multi-Scale Integrated Analysis of Societal and Ecological Metabolism*, est une approche d'analyse systémique intégrant des données quantitatives (obtenues généralement avec des modèles) et les différentes échelles d'analyse (eg *micro*, *méso*, *macro*).

L'approche MuSIASEM se base sur le modèle de flux/fonds qui est détaillé ci-dessous.

**Distinction flux/fonds :** Pour obtenir une description plus détaillée du système, une distinction entre les différents constituants du systèmes doit être faite. On appelle **flux** tout élément du système disparaissant lors de la représentation du système. On peut penser au carburant lors de la représentation du fonctionnement d'une voiture qui est bien consommé lors de son fonctionnement. D'un point de vue plus sociétal, on peut penser à toute ressource énergétique. On oppose souvent aux flux, la notion de **fond** (ou *fund* en anglais). Cette notion englobe tout élément du système qui ne disparaît pas, qui reste intact, durant l'échelle de temps fixé par la représentation du système voulue. Les fonds d'un système ont souvent un rôle de convertisseur, ils permettent la transformation d'un flux entrant en un flux sortant. Les fonds sont utilisés à un certain rythme et doivent être périodiquement renouvelés. On peut penser notamment à la main d'oeuvre d'une entreprise qui est toujours disponible d'un jour sur l'autre. On ne possède qu'une certaine quantité de travail et les employés ne reste pas éternellement dans l'entreprise (retraites, départs pour raisons diverses). La citation suivante offre une synthèse pertinente de ces deux notions :

"[...] fund coordinates represent the set of attributes used by the analyst for defining *what the system is* and flow coordinates represents [...] "*what the system does*", when interacting with its context." [MG08]

Ces deux notions relèvent (et révèlent) des choix et des hypothèses de modélisation effectués pour rendre compte du système. En effet, il n'est pas toujours très objectif d'affirmer que tel ou tel élément du système reste intact ou disparaît sur l'échelle de temps que l'on s'est fixé.

**Fonctionnement de MuSIASEM :** Il est détaillé ci-dessous.

- Étape 1 : La première étape, lorsque l'on souhaite utiliser MuSIASEM, consiste à dresser une description de notre système, en compartiments hiérarchisés qui sont essentiels à la survie de celui-ci.
- Étape 2 : La répartition quantitative des fonds au sein des différents compartiments fonctionnels du système. Typiquement la quantité d'activité humaine, vue comme un fond, que l'on va répartir dans chaque compartiment d'une société.
- Étape 3 : La définition quantitative des flux utilisés par les différents compartiments du système.

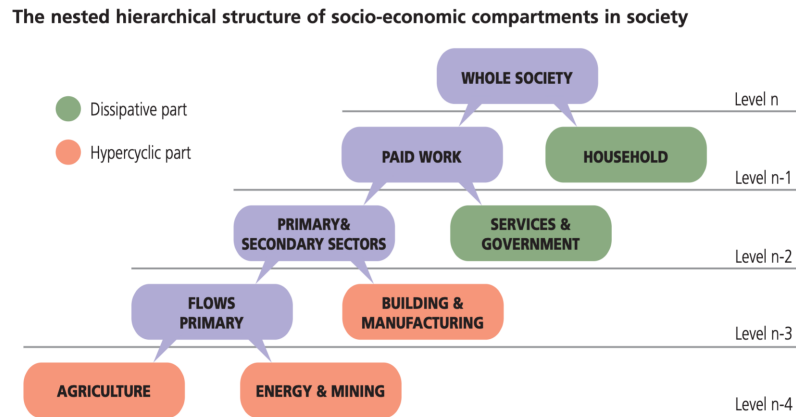


FIGURE 1.1 – Étape 1 : représentation hiérarchiques de la société en compartiments socio-économiques. Les cadres rouges définissent la partie *hypercyclique*, soit la partie qui produit les ressources. Les cadres verts définissent la partie dissipative qui consomme le surplus produit par la partie hypercyclique.

- Étape 4 : Design de la matrice multi-dimensionnelle décrivant le métabolisme du système. Elle permet une description systémique précise.

**Multi-level matrix characterizing the metabolic pattern of Mauritius**

		Flow elements			Fund elements			Money (billion USD)
		Food (PJ)	Energy (PJ-GER)	Water (hm <sup>3</sup> extraction)	HA (Mhr)	PC (GW)	Land (Kha)	
Consumption	HH	5.9	16	98	10,000	1.5	28	n/a
	PW*	0.8	37	11	600	1.1		8,200
	AG	1.3	negligible	100	39	negligible	21	210
	EM	n/a	2.2	260	8	0.03	negligible	180
	exp <sub>pw*</sub>	n/a	n/a	3	590	n/a	n/a	50% GDP
	exp <sub>ag</sub>	negligible	0.36	1,100	39	0.02	51	2.5% GDP
	Whole	8	56	1,700	11,300	6.0	100	9,800 (GDP)
Supply	Imports	6.7	19	n/a	n/a	n/a	n/a	63% (GDP)
	Domestic supply	1.3	7	1,700	11,300	6.0	100	9,800

FIGURE 1.2 – Étape 4 : matrice multi-dimensionnelle du système [Gia14]. Pour tout compartiment du système (colonne de gauche), sont renseignés les quantités de chaque flux (*flow elements*) et de chaque fond (*fund elements*) utilisé. Dans les compartiments de consommations, *HH* désigne le secteur des foyers (*household sector*), *PW*, le secteur du travail (*paid work sector*), *AG* le secteur agricole et *EM*, le secteur énergétique (*energy and mining sector*) les *exp* désignent les parties exportation. Pour les fonds, *HA*, désigne l'activité humaine, *PC* la puissance disponible (*power capacity*).

- Étape 5 : Conclusion sur la viabilité, désirabilité et faisabilité du système. Un système est dit *viable* s'il parvient à produire les flux qu'il consomme. Il est dit *désirable* si la

quantité de flux et de fonds accordées à la partie dissipative (voir. Figure 1.1) du système est grande. Cela permet à la société des standards de vie élevés et il existe une corrélation entre la désirabilité d'un système et des niveaux d'indicateurs de développement hauts. Enfin, la *faisabilité* désigne la capacité du système à satisfaire les conditions "de bord" requises. Ces conditions de bords désignent les échanges existants entre le système et son environnement, *eg* la quantité d'énergie entrante, la quantité de déchets sortantes etc...

**Conclusion :** L'approche MuSIASEM constitue alors un outil de diagnostic et de scénario pour évaluer un système. Il permet de prendre en compte beaucoup de facteurs simultanément et à plusieurs échelles différentes. Il est également critiquable, notamment car il repose en grande partie sur l'acquisition de données pas toujours faciles à obtenir, et c'est pourquoi il en existe bien d'autres : on peut citer *MEFA*, *Material and Energy Flow Analysis* par exemple (on pourra en retrouver d'autres ainsi qu'une analyse des atouts et inconvénients de chacun dans [GS18]).

Cet outil illustre bien à quoi ressemble un outil systémique et quelles peuvent être ses applications. C'est avec celui-ci que j'ai pu me familiariser avec ces notions de métabolisme social et obtenir une vision plus systémique des choses.

## 1.2 Vulnérabilité

Idéalement, un SST doit pouvoir être évalué selon plusieurs critères, comme la satisfaction des besoins de la population, l'inscription ou non dans les limites environnementales locales et globales, la performance socio-économique et enfin les vulnérabilités du système. J'ai étudié plus précisément ce dernier critère lors de mon stage. Il est particulièrement utile lorsqu'il s'agit de décrire les impacts du changement climatique et son étude m'a permis de me familiariser encore plus avec le paradigme systémique. Mais qu'est-ce que la vulnérabilité d'un système ?

Il existe beaucoup de définitions de la notion ainsi que plusieurs manières de l'utiliser. Nous rediscuterons de cette polysémie mais pour l'instant, contentons-nous de la définition suivante donnée par [Del17].

**Définition 1.2.1** *La vulnérabilité d'un système décrit la propension du système à subir un dommage.*

Nous pouvons alors nous poser les questions suivantes : comment rendre compte de la vulnérabilité d'un système ? Comment faire son analyse de vulnérabilité ? Quels éléments devons-nous inclure ? Comment comparer la vulnérabilité de deux systèmes différents ?

### 1.2.1 Analyse de vulnérabilité

Pour effectuer une analyse de vulnérabilité [Del17] propose de tenir compte de l'environnement physique et naturel du système ainsi que son contexte socio-culturel. Pour ce faire, l'idée est d'étudier le système socio-technique à différentes échelles et d'en comprendre, en situation de crise, les logiques et les représentations de chaque sous-système. Cela permet d'en comprendre

les comportements de réaction aux aléas, les actions prioritaires et la manière dont elles sont coordonnées par les différents acteurs notamment. On peut alors relier chaque sous-système avec les autres en utilisant les relations existant entre eux.

Il est aussi pertinent d'étudier les flux du système socio-technique car ils caractérisent son fonctionnement, et ce, de manière dynamique, aspect qui est très important lorsqu'il s'agit de le décrire en situation de crise.

*"[...] l'intérêt des flux provient du fait qu'ils sont particulièrement révélateurs de dysfonctionnements de système urbain en situation de crise, et ce pour toutes les échelles. Lors d'une crise, [...], [la circulation des flux], peu[t] se trouver dégradée, réduite voire interrompue. Le moment à partir duquel la circulation des flux commence à être dégradée fournit également des informations quant à l'entrée en crise du système et de ses composants, à la durée de la crise, à la mise en route des actions d'urgence et au retour à une situation dite normale "* [Del17]

La temporalité lors d'une situation de crise est également assez importante car elle donne beaucoup de renseignements via la description de différentes étapes (voir Figure 1.3). On peut, en effet, définir plusieurs phases qui caractérisent bien la vulnérabilité du système en crise : une phase préliminaire à la crise, de prévention, ensuite une phase de subissement de l'aléa et enfin une phase de résilience ou de guérison. On en profite pour définir la **résilience** comme la capacité d'un système endommagé à revenir à son état dit "normal". Il est également possible de le voir comme la mesure des dégâts et des perturbations qu'il peut endurer.

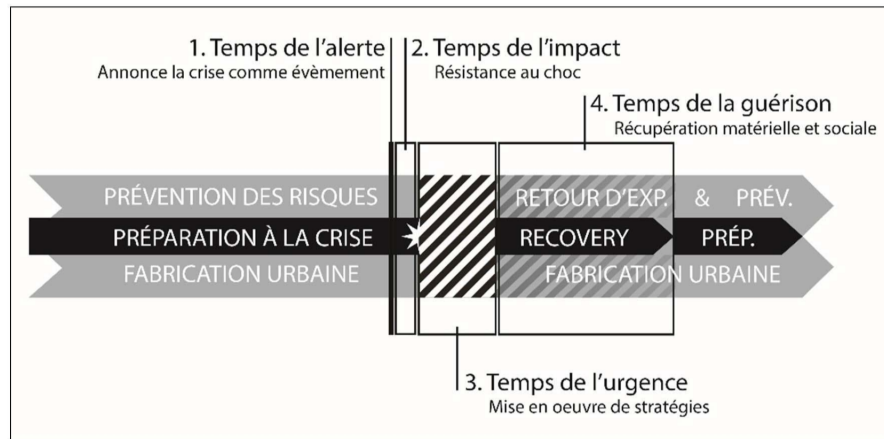


FIGURE 1.3 – Différentes phases lors d'une situation de crise. [Del17]

### 1.2.2 Discussion sur les définitions possibles et éléments de modélisation

Il existe une variété de définitions possibles pour le concept de vulnérabilité. Cela est dû à son utilisation dans des domaines très variés comme l'écologie, ingénierie, la climatologie, les sciences sociales ou de l'environnement.

Pour se rendre compte de la diversité des définitions, nous pouvons nous poser la question suivante : comment comparer la vulnérabilité de deux systèmes différents ?

### 1.2.2.1 Un exemple

Comme le souligne [Fü07], deux systèmes évoluant dans des contextes climatiques, socio-économiques et politiques différents, exposés à des aléas différents sont difficiles à comparer. L'exemple du subissement du changement climatique entre :

- le Tibet, qui n'est exposé que modérément<sup>1</sup> au changement climatique, ne possède que peu de moyens pour lutter contre un aléa et subit des pressions politiques par la Chine,
- la Floride qui est exposée à une très probable augmentation du niveau de la mer<sup>2</sup>, dû à sa faible élévation. Cela pourrait augmenter la fréquence et les dégâts causés par les ouragans.

### 1.2.2.2 Éléments à modéliser

Nous pouvons exhiber quatre éléments très important pour décrire la vulnérabilité d'un système :

- le système en lui-même, avec tous les liens qu'il entretient avec son environnement,
- les attributs d'intérêts, comme les vies humaines, l'économie d'un pays, l'environnement, un éco-système etc...,
- l'aléa ou le danger, qu'on peut définir comme un événement probable pouvant avoir une influence sur un des attributs d'intérêts du système,
- l'échelle de temps considérée, qui décrit la période de référence (l'instant et la durée) au cours de laquelle on étudie la vulnérabilité de notre système.

Nous pouvons nous rendre compte de la pertinence de ces paramètres en les appliquant à la crise du COVID-19 en France : étudier la région de Bretagne<sup>3</sup> ou l'île de France, lors de la première vague, ne donne pas la même vulnérabilité. La précision du système étudié est alors importante.

Étudier la situation économique du pays ou sa situation sanitaire est différent. Quels sont les attributs d'intérêts ? Les vies humaines ou l'économie du pays ?

L'étude de l'épidémie ou de l'épisode de canicule de cet été n'est pas la même chose, la France est bien mieux préparé au dernier qu'au premier.

Enfin, l'aspect temporel est aussi très important car l'impact de la deuxième vague de contamination du COVID-19 ne sera sûrement pas le même que celui de la première vague, étant donné la meilleure gestion de l'épidémie durant l'été.

Une conclusion de cette discussion est la nécessité de précision dans l'énoncé de la situation de vulnérabilité, notamment vis-à-vis de ces quatre paramètres.

---

1. Mes recherches sur le sujet ne m'ont pas vraiment permises de nuancer cette phrase car je n'ai pas trouvé de document sourcé traitant de la gravité de la vulnérabilité du Tibet face au CC

2. *La montée des eaux noires*, John Young, "<http://archive.wikiwix.com/cache/index2.php?url=http%3A%2F%2Fwww.delaplanete.org%2FLa-montee-des-eaux-noires.html>

3. *Coronavirus. Pourquoi la Bretagne est-elle la région française la moins touchée ?*, Ouest-France, 05/20

### 1.2.2.3 Distinction entre les types d'aléas

[Fü07] effectue une synthèse des différentes distinctions d'aléas entre plusieurs articles et conclue sur la nomenclature suivante :

Sphere	Domain	
	Socioeconomic	Biophysical
Internal	Household income, social networks, access to information	Topography, environmental conditions, land cover
External	National policies, international aid, economic globalization	Severe storms, earthquakes, sea-level change

FIGURE 1.4 – Illustration des deux dimensions indépendantes pour le classification des **facteurs de vulnérabilité**, la sphère et le domaine.[Fü07]

La **sphère** désigne l'aspect endogène ou exogène de la vulnérabilité, cette dimension répond à la question : d'où provient ma vulnérabilité ? Provient-elle de l'intérieur de mon système comme un épisode de canicule (climat français) ? ou provient-elle de l'extérieur et de l'environnement du système comme la crise du COVID-19 ?

Le **domaine** de la vulnérabilité désigne lui, la nature même de la vulnérabilité, il permet de classer les vulnérabilités entre deux catégories distinctes : les vulnérabilités de nature biophysiques (phénomène climatiques, biologiques, physiques, géologiques etc...) et celles de nature socio-économiques (phénomènes sociétaux, économiques, politiques etc...).

Bien sûr, une vulnérabilité externe biophysique peut entraîner une vulnérabilité interne socio-économique (l'exemple du COVID-19 encore) mais ces dimensions sont bien indépendantes et permettent de définir précisément la vulnérabilité en question.

Une troisième dimension doit être précisée, l'aspect temporel de la vulnérabilité en question : est-elle présente, future ? Ponctuelle ou au contraire progressive ? Il est possible de répondre à cette question en caractérisant l'aléa de **présent** (un ouragan), **future** (une crise économique lointaine) ou encore **dynamique** (comme le changement climatique, qui est progressif et non ponctuel).

### 1.2.2.4 Exposition, sensibilité, résilience

Pour décrire une vulnérabilité, il est possible de décrire l'exposition, la sensibilité et la résilience du système pour obtenir une analyse plus précise. Définissons les deux premiers termes :

- L'**exposition** est la mesure des attributs d'intérêt qui peuvent être affecté par l'aléa. La quantité de vies humaines affectée, le nombre d'institutions affectées par une crise etc...
- La **sensibilité** est la mesure de la manière dont va être affecté le système. Cette notion décrit les conséquences immédiates, ainsi que leurs gravités, que va causer l'aléa. Typiquement, les conséquences d'une canicule sur une région, le système agricole va subir



des pertes de rendement tandis que certaines entreprises vont être moins sensibles à la hausse des températures.

On voit dans les exemples de sensibilité donnés, tous les sous-systèmes d'un même système ne sont pas sensibles de la même manière. On peut essayer de résumer le subissement d'un aléa par ces trois termes.

- À un instant donné, le système va être exposé d'une certaine manière.
- Face à cette exposition, le système va subir un dommage, décrit par sa sensibilité.
- Ensuite, le système va utiliser sa capacité de résilience pour se reconstruire, limiter son exposition et/ou sa sensibilité (phase 3 et 4 sur la Figure 1.3). On peut ainsi obtenir des boucles de rétroactions où le système, en restreignant sa sensibilité, restreint son exposition à l'aléa.

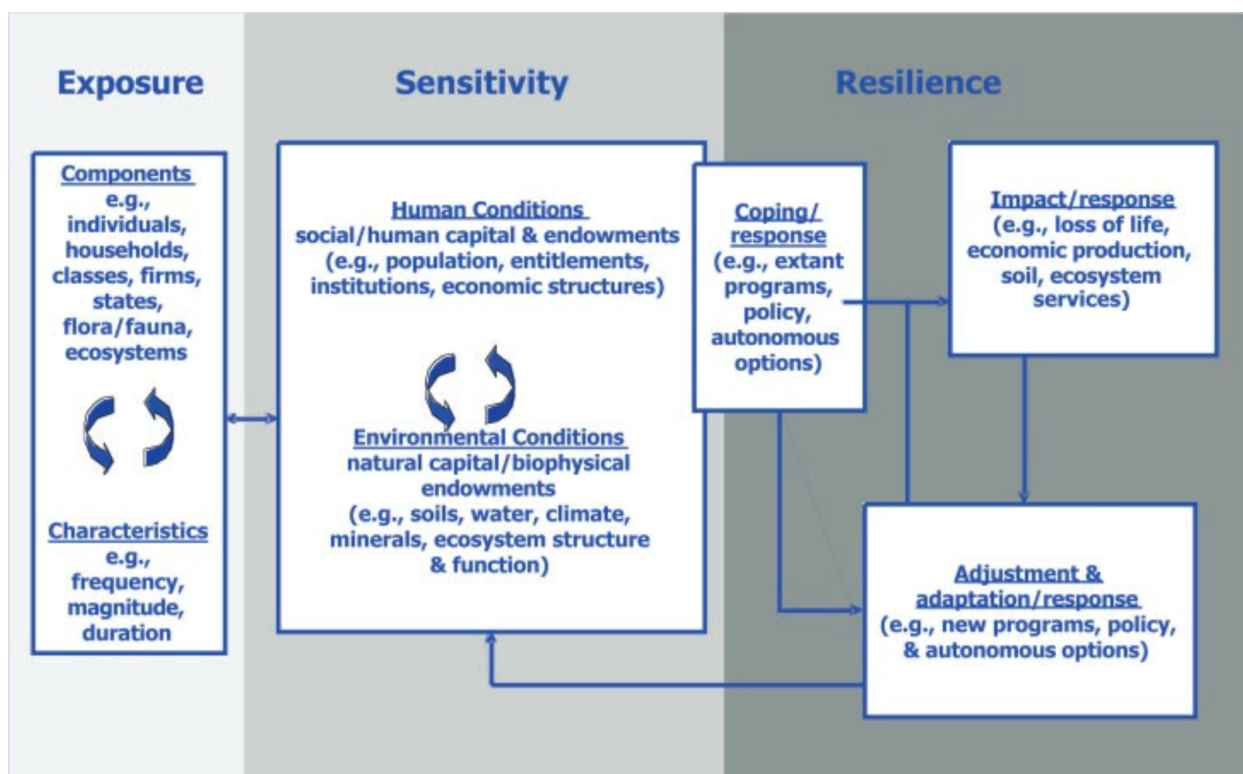


FIGURE 1.5 – Illustration d'une description de la vulnérabilité d'un système. Les trois concepts décrit ci-dessus y sont représentés. [Tur03]

Je me suis contenté de ces trois concepts principaux pour guider mon analyse de vulnérabilité même s'il est possible d'en utiliser d'autres ([Foe09]). On peut citer la *capacité-tampon* d'un système<sup>4</sup>, la robustesse<sup>5</sup> ou encore la redondance<sup>6</sup>

4. Mesurant la capacité d'un système à minimiser l'impact de l'aléa sur le maintien de ses fonctions essentielles

5. Mesurant l'aptitude d'un système à fonctionner correctement sous perturbations

6. Capacité d'un système à activer des systèmes de substitution lors de perturbations.

### 1.2.2.5 Retour sur l'exemple

Pour synthétiser notre propos, revenons sur l'exemple de la comparaison entre la situation du Tibet et celle de la Floride. Notre analyse n'est qu'anecdotique et représente surtout une occasion d'illustrer les concepts introduits.

- Le Tibet connaît, à l'heure actuelle ; des vulnérabilités de domaine plus socio-économiques (pressions de la Chine) que biophysiques. Ces perturbations sont majoritairement externes et dynamiques, elles menacent la population et l'autonomie politique de la zone. La région n'est que peu exposée au phénomènes biophysiques mais elle pourrait y être très sensible s'ils se produisaient, la région pourrait avoir une capacité de résilience assez limitée.
- La Floride, quant à elle, subit exclusivement des perturbations d'ordre biophysique interne (dues à sa topographie) mais aussi externe (à cause des ouragans notamment). La perturbation de la montée des eaux est d'ordre dynamique tandis que la menace d'ouragans est un événement future<sup>7</sup> La région est exposée et très sensible mais avec une très grande capacité de résilience<sup>8</sup>.

Nous n'avons toujours pas réellement répondu à la question, "*quel système est le plus vulnérable ?*" mais nous pouvons nous apercevoir, au contraire, à quel point ces situations ne sont pas comparables, tout en ayant une idée bien plus précise de la vulnérabilité de chacun de ces systèmes.

## 1.3 Ingénierie système : Langage SysML et SADT

Je présente maintenant des méthodes d'ingénierie système servant à décrire des systèmes purement techniques. Ils se révèlent d'une grande utilité pour décrire les systèmes incluant une dimension socio-culturelle. Un des enjeux de mon stage, a été de m'inspirer d'une utilisation d'une méthode d'ingénierie système, le langage SysML, pour représenter un SST. Cela avait déjà été fait via un autre langage, le langage SADT que je présente ci-dessous.

### 1.3.1 SADT

**SADT** signifie Structure Analysis and Design Technique est une méthode de représentation graphique des systèmes complexes en ingénierie système. C'est une approche descendante, au sens où il est possible d'avoir une description de chaque échelle du système, "*on peut zoomer*". Cet aspect, lorsque transposé à la description de SST, est très intéressant comme nous avons pu le voir dans la partie sur MuSIASEM.

---

7. Il serait possible de voir ce problème de manière plus dynamique via le fait que le changement climatique augmente la fréquence des événements extrêmes.

8. Cette résilience est très relative, 6 millions de personnes devrait migrer des côtes d'ici la fin du siècle selon le démographe Jason Evans, et ce malgré les investissements massifs de la ville dans des plans d'adaptation (sources dans l'article). *En Floride, les riches n'auront pas les pieds dans l'eau*, L. Raim, 2020 <https://www.monde-diplomatique.fr/2020/05/RAIM/61762>

### 1.3.1.1 Éléments de modélisation

La représentation d'un bloc en langage SADT est illustré sur la figure ci-dessous.

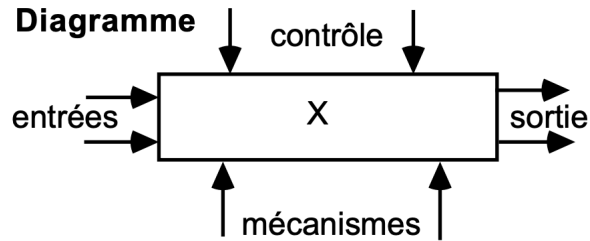


FIGURE 1.6 – Représentation d'un bloc fonctionnel en SADT

X représente la fonction du système (comme laver le linge). Les flux occupent une place centrale dans cette représentation. Les entrées et les sorties représentent les flux circulant dans le système X. Les flèches "contrôle" représentant les flux constants au cours du fonctionnement de X comme l'alimentation électrique ou en eau ou encore les réglages. Les flèches mécanismes indiquent, quant à elle, des précisions sur le dispositif utilisé pour réaliser fonction X. (typiquement machine à laver pour X laver le linge)

SADT possède également un aspect récursif très intéressant. Le langage offre la possibilité de représenter un bloc fonctionnel en fonction des sous-blocs fonctionnels, qui eux mêmes peuvent être représentés en fonction de leurs sous-blocs. C'est ce qu'on appelle un aspect descendant.

**Un exemple** Plaçons-nous dans le cadre d'un dispositif électrique pour couper du bois. Nous pouvons choisir de représenter le système de la manière suivante :

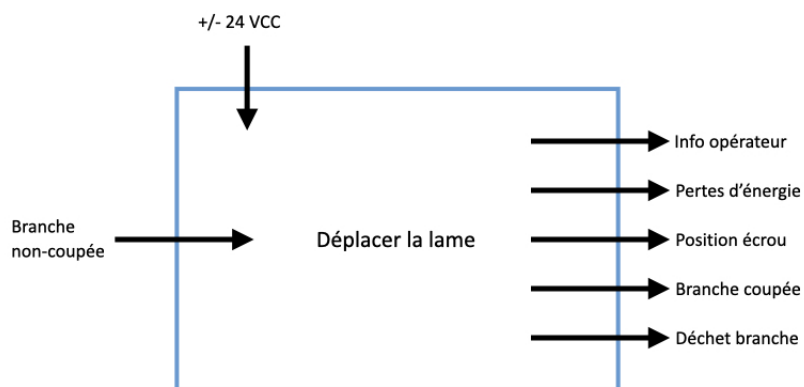


FIGURE 1.7 – Représentation simple du bloc fonctionnel

Mais il est également possible d'utiliser l'aspect récursif qu'offre le langage.

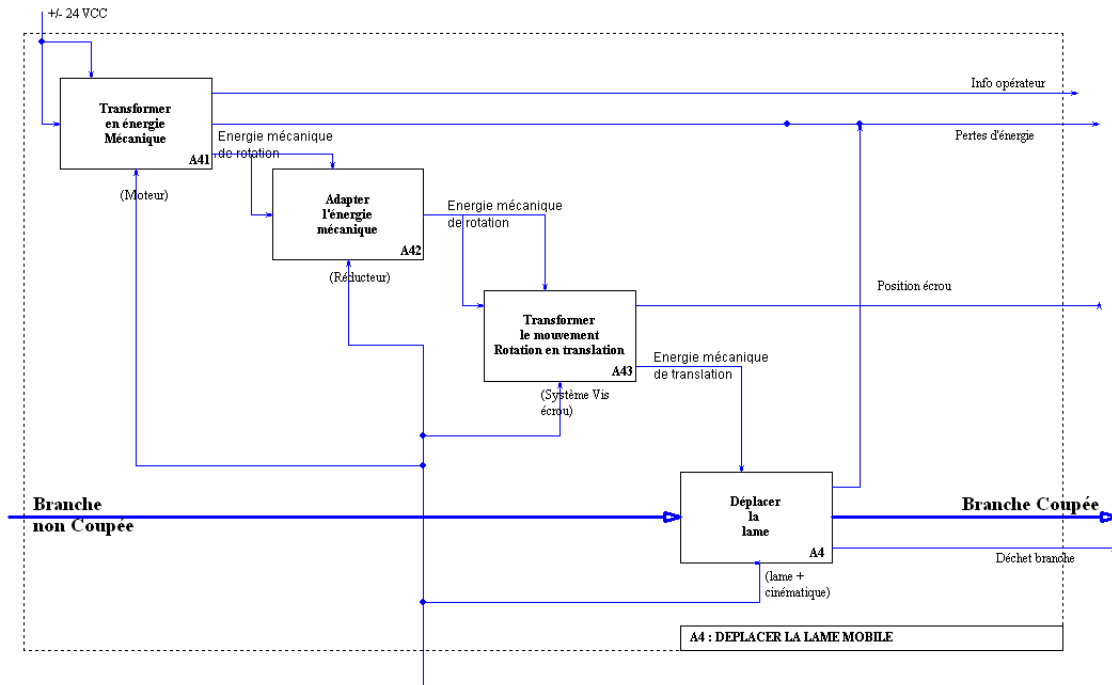


FIGURE 1.8 – Aspect descendant de la méthode de représentation SADT

La figure ci-dessus illustre comment il est possible "descendre" dans le niveau de précision avec lequel nous représentons notre système. Nous y voyons le bloc fonctionnel "transformer en énergie mécanique" précéder le bloc "adapter l'énergie mécanique" etc... donc des sous-fonctions se succéder pour constituer un plus gros bloc fonctionnel "Déplacer la lame mobile".

L'aspect descendant et centré sur les flux du langage SADT le rend particulièrement intéressant pour l'étude des SST. On verra dans le chapitre 2 comment il a été utilisé.

### 1.3.2 Langage SysML

Le langage SysML ou Systems Modeling Language est un langage basé sur le langage de modélisation logiciel UML. Il consiste comme son nom l'indique à modéliser des systèmes dans le domaine de l'ingénierie système, autant pour l'explication du fonctionnement d'un système que pour sa conception. Ce langage est composé de diagrammes qui sont regroupés par paquets en fonction des aspects qu'ils permettent de modéliser dans le système.

Une grande avancée dans notre travail a été d'essayer d'adapter les outils de représentation de SysML à la représentation des systèmes socio-techniques. Pour comprendre comment adapter ce langage à mes objets d'études, j'ai dû comprendre comment il fonctionnait.

Dans les parties ci-dessous, je décris le fonctionnement des parties de SysML qui m'ont été utiles du point de vue de la modélisation. Je me suis servi en grande partie d'un cours de [LG13] pour me familiariser avec ces concepts. J'ai choisi de présenter ces outils dans l'ordre

conventionnel, c'est-à-dire dans l'ordre de conception d'un produit. L'ingénieur cherchant à concevoir un produit en utilisant le langage SysML, va en premier lieu dresser les diagrammes fonctionnels. Ils permettent de définir les exigences et les cas d'utilisation du produit. Ces diagrammes permettent de répondre aux questions : "que doit faire le système?", "à quoi sert-il?". C'est bien la première étape lorsque l'on souhaite créer un produit.

Dans un second temps, la question naturelle qui vient à l'esprit est "de quoi est composé le système?". Les diagrammes structurels de SysML permettent de répondre à cette question.

Enfin dans un troisième temps, je présente les diagrammes comportementaux qui décrivent le comportement du système et répondent à la question "comment le système réalise-t-il les exigences énoncées?"

### 1.3.2.1 Diagrammes fonctionnels

Il est possible d'utiliser les diagrammes fonctionnels pour traduire les fonctions du système et les liens entre les fonctions et les acteurs du système.

**Diagramme des cas d'utilisation** Les diagrammes d'utilisation ont pour but de représenter les actions réalisables par le système et le lien avec les différents acteurs (individu, automate) interagissant avec le système. Elle permet de comprendre l'organisation de ces derniers autour des fonctions du système (voir Figure 1.9).

**Diagramme des exigences** Le diagramme des exigences sert lui, à représenter les différentes conditions requises pour le fonctionnement du système. Ces exigences sont organisées et explicités, reliées aux cas d'utilisation.

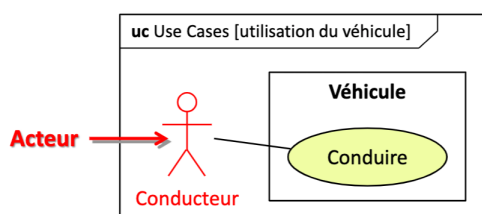


FIGURE 1.9 – Diagramme des cas d'utilisation d'une voiture où l'on a représenté un seul acteur, le conducteur, qui utilise le système voiture pour conduire. On aurait également pu représenter un acteur «Passager» utilisant la voiture pour «Être transporté».

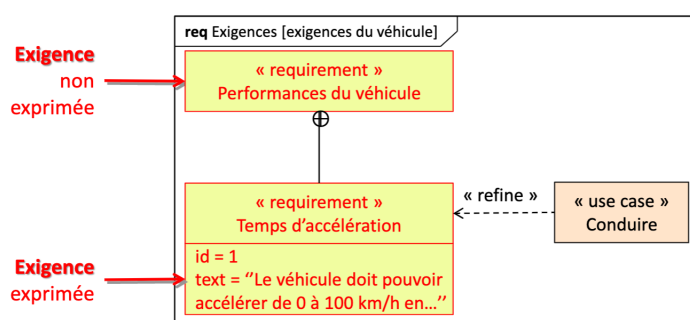


FIGURE 1.10 – Diagramme des exigences relatif au diagramme des cas d'utilisation de la voiture. Les performances du véhicule sont les seules exigences représentées ici et elles sont liées au cas d'utilisation «Conduire» représenté dans l'autre diagramme.

### 1.3.2.2 Diagramme structurels

Les diagrammes structurels permettent de décrire la structure et la composition du système.

**Diagramme de définition des blocs** Le diagramme de définition des blocs permet de définir les composant (ou blocs) du système. On peut y spécifier des valeurs, des paramètres...

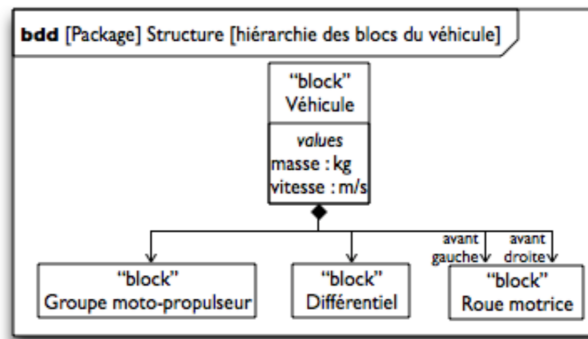


FIGURE 1.11 – Diagramme de définition des blocs du véhicule. Des valeurs sont assignées au bloc véhicule.

**Diagramme interne d'un bloc** Le diagramme interne d'un bloc modélise la structure d'un bloc et spécifie les flux existants entre les différents composant du système.

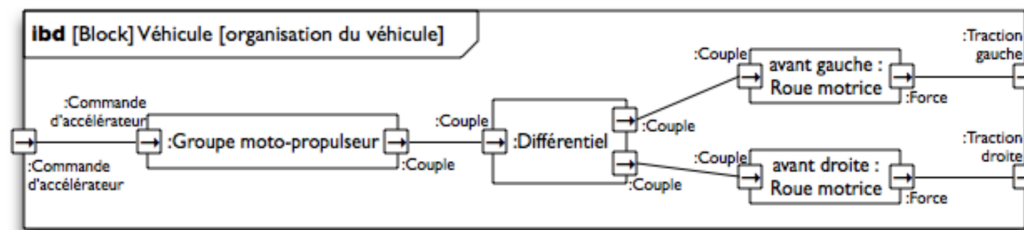


FIGURE 1.12 – Diagramme interne du bloc «véhicule». Les différents composants, définis dans le diagramme de définition des blocs, mis en relation par les flux.

### 1.3.2.3 Diagramme comportementaux

Les diagrammes comportementaux permettent de décrire la dynamique du système. Ils sont particulièrement intéressants car ils permettent de représenter des fonctionnements complexes et sont très adaptables.

**Diagramme d'états** Le diagramme d'états représente différents états du système. Ces états sont liés par des flèches représentant les événements qui font passer d'un état à un autre. Ce diagramme a été très utile pour les représentations sur lesquelles j'ai travaillé, je lui consacre une section dans le chapitre 2.

**Diagramme de séquences** Le diagramme de séquences décrit les interactions entre les différents composants du système par une suite de messages qui sont supposés instantanés.

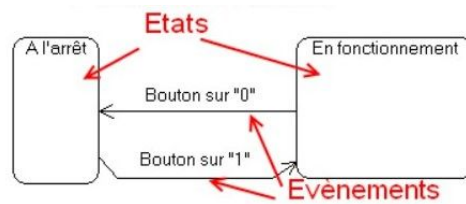


FIGURE 1.13 – Diagramme d'état d'un système générique.

à l'échelle de temps de la représentation. La dimension verticale du diagramme représente le temps. Les rectangles symbolisent la période d'action du composant.

Il est également possible de donner beaucoup de précisions quant à la transmission des messages. Il est par exemple possible d'utiliser des représentations algorithmiques *"si composant A fait truc, alors composant B fait truc, sinon fait machin"* ou encore de préciser que plusieurs messages différents s'envoient simultanément.

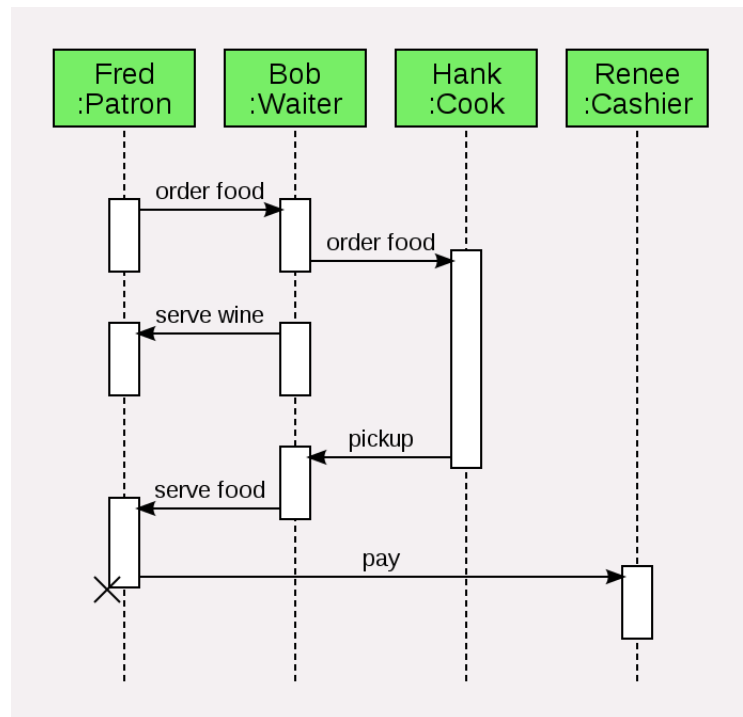


FIGURE 1.14 – Diagramme de séquence d'un restaurant.

# Chapitre 2

## Contribution personnelle : outils de représentation systémique pour les SST

Après avoir étudié de la bibliographie, j'ai commencé à appliquer les outils de mes lectures, présentés à la partie précédente, en représentant différents systèmes socio-techniques.

L'objectif n'était pas vraiment d'obtenir une représentation fidèle du système ; n'ayant pas les connaissances requises pour prétendre à une compréhension fine d'un système en particulier. Le but de mes travaux était de développer des outils de modélisation adaptables, par un expert, à la représentation d'un système générique et de montrer que les aspects sociaux d'un SST sont représentables à l'aide des outils de l'ingénierie système. Il est facile de voir que les aspects techniques sont facilement représentables à l'aide de ces langages, le défi a été ici de se poser la question moins évidente si cela était possible pour les aspects socio-culturels.

L'enjeu derrière tout cela est de proposer ces outils à des experts pour pouvoir créer des modélisations d'alternatives socio-techniques. Ces modélisations peuvent alors être présentées à des décideurs. Elles ont alors pour but de simplifier le processus de prise de décision.

Je n'ai pas réellement cherché à représenter des systèmes de manière très exacte et fidèle, j'ai cherché à m'exercer, à comprendre que telle situation ou tel aspect d'un système était représentable avec tel outil.

Je me suis aidé d'un travail déjà existant : les modélisations d'une doctorante, Michela Bevione, et celles d'une ancienne stagiaire de l'équipe Steep, Caroline Rivier.

### 2.1 Inspiration des diagrammes de Sankey et SADT

#### 2.1.1 Les diagrammes de Sankey

Ces représentations appelés *diagrammes de Sankey* sont des diagrammes représentant les flux circulant dans un système. Un élément caractéristique de ces diagrammes est que les flèches représentant les flux sont de taille proportionnelle à la valeur de la grandeur qu'elles représentent. Historiquement, ils ont d'abord été utilisés dans le domaine de l'énergie mais connaissent aujourd'hui des applications en économie ou en écologie. Ils permettent de facilement de comprendre comment circulent les flux dans le système, identifier les pertes du système.



On montre ici, ces diagrammes réalisés en 2019 par Caroline Rivier dans le cadre d'une étude économique sur la filière AOP du Beaufort dans la Vallée de la Maurienne.

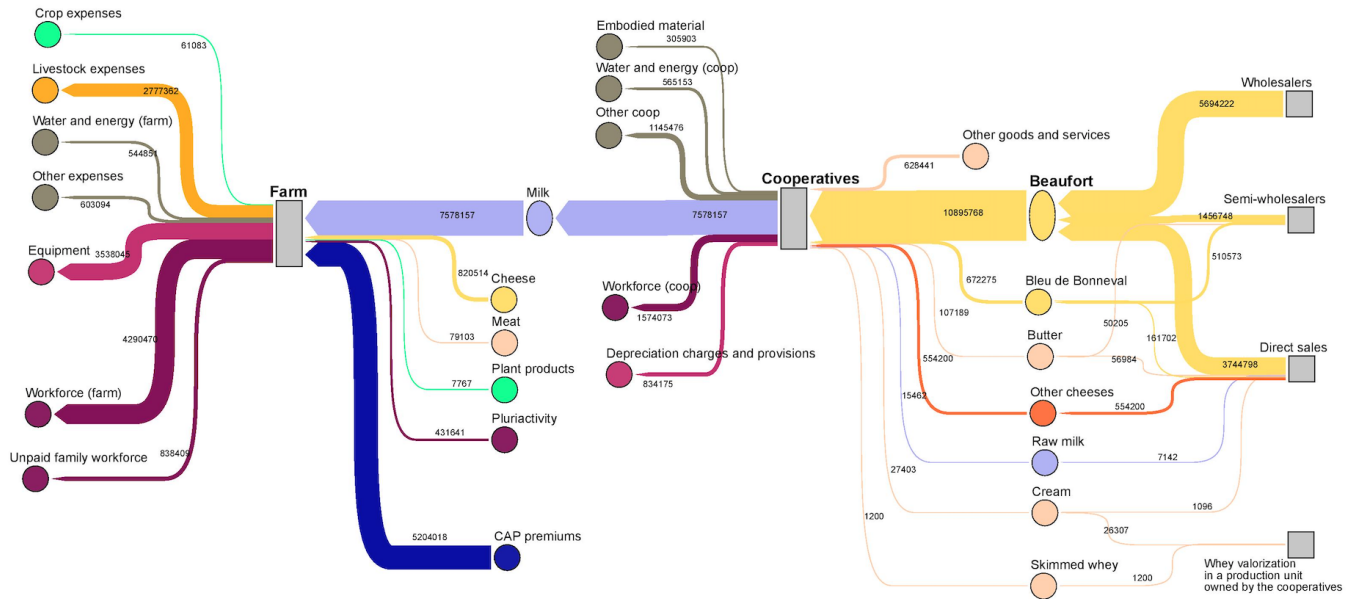


FIGURE 2.1 – Diagramme de Sankey des flux monétaires circulant dans la Vallée de la Maurienne. Les bloc gris représentent les acteurs, toutes les flèches représentent des flux monétaires. La couleur dépend de l'origine de ces flux qu'il n'est pas nécessaire de détailler ici.

Chaque flux est associé à une valeur et Caroline a étudié l'interdépendance des flux. Notamment l'arrêt d'exploitations fragiles dans la Vallée.

Je me suis inspiré de ce diagramme, notamment de la manière dont sont représentés les flux pour adapter SysML (voir la partie sur les diagrammes d'état).

## 2.1.2 Les diagrammes SADT

Michela Bevione, une thésarde s'était inspirée du formalisme SADT pour créer des représentations de SST. Cette utilisation de SADT est motivée par les raisons que nous avons énoncés dans la partie 1.3.1, à savoir l'aspect descendant (récursivité de la représentation) et la représentation des flux de la méthode. Un exemple de ces représentations est illustré sur la figure 2.2. La syntaxe (c'est à dire la manière de représenter les différents éléments du système) de SADT est bien respectée et permet une modélisation assez complexe du SST (seuls les aspects techniques sont représentés sur cette figure, les aspects socio-culturelles sont ajoutés sur la figure 2.3).

La production de foin : ressources et produits

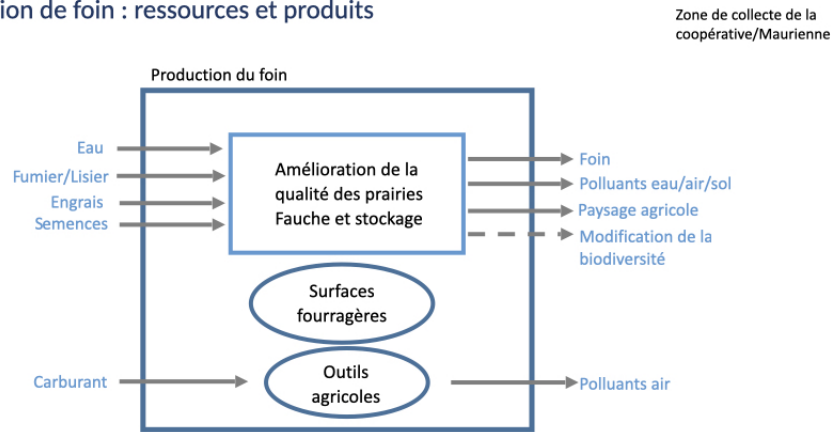


FIGURE 2.2 – Adaptation de la méthode SADT pour représenter la fonction ”production du foin” dans le SST de la vallée de la Maurienne. On y voit des flux entrants, des flux sortants, des sous-fonctions (l’amélioration de la qualité des prairies et la fauche et stockage) et les fonds du système (les surfaces fourragères et les outils agricoles).

La production de foin : acteurs

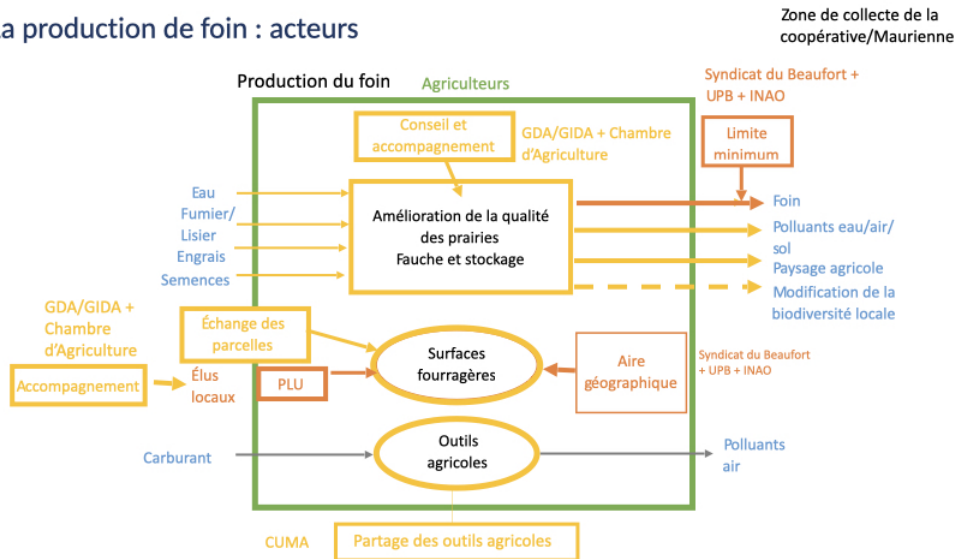


FIGURE 2.3 – Adaptation de la méthode SADT pour représenter la fonction ”production du foin” dans le SST de la vallée de la Maurienne. Les aspects socio-culturels sont représentés : les acteurs présentés, remplissent des fonctions, c’est pourquoi ils sont disposés autour des blocs fonctionnels (on utilise la structure descendante de SADT). L’idée n’est pas de comprendre en détail, c’est pourquoi la signification des flèches en pointillé et des codes couleurs n’est pas expliquée.

Ces diagrammes présentent tout de même des inconvénients. Ils n’incluent pas vraiment une description du comportement chronologique du système ; ceci est un aspect important de la modélisation car il permet de décrire l’évolution du système, sa dynamique.

Ces diagrammes ne permettent pas non plus de comprendre les liens entre les différents acteurs intervenant dans le système (là où le diagramme de Sankey fait légèrement mieux) et comment ils interagissent.

Les liens entre les différentes sous-fonctions du système, comme des liens de dépendance, ne sont pas explicités sur ces diagrammes.

Toutes ces remarques m’ont poussé à essayer d’inclure ces aspects à l’aide d’un nouvel outil : le langage SysML. Ses nombreux et divers diagrammes permettent de séparer les informations et de pouvoir modéliser les aspects mentionnés ci-dessus. Je présente mes résultats dans les parties ci-dessous.

## 2.2 Adaptation du langage SysML pour l’étude de la vulnérabilité

### 2.2.1 Les diagrammes structurels

Les diagrammes structurels permettent de répondre à la question «de quoi est composé le système?». J’ai eu l’idée d’utiliser ces diagrammes pour donner une description de l’ensemble des acteurs tenant un rôle dans le système.

Dans un diagramme de définition des blocs, censé lister les composants du système, j’ai listé les différents acteurs. Dans un second temps, dans un diagramme de bloc interne, qui doit normalement décrire les interactions entre les différents composants et les flux circulant entre eux, j’ai représenté les flux existant entre les acteurs, notamment les flux d’information. Cela permet d’illustrer la collaboration de certains acteurs et donne une idée des réactions possibles du système face à un aléa. En effet, s’il existe, par exemple, des clivages idéologiques entre différents acteurs, il est difficile d’espérer une collaboration rapide de ces différents acteurs lors du subissement d’un aléa. Les réactions que l’on peut prévoir seront alors plus des actions individuelles des acteurs ou des collaborations entre acteurs communiquant déjà entre eux.

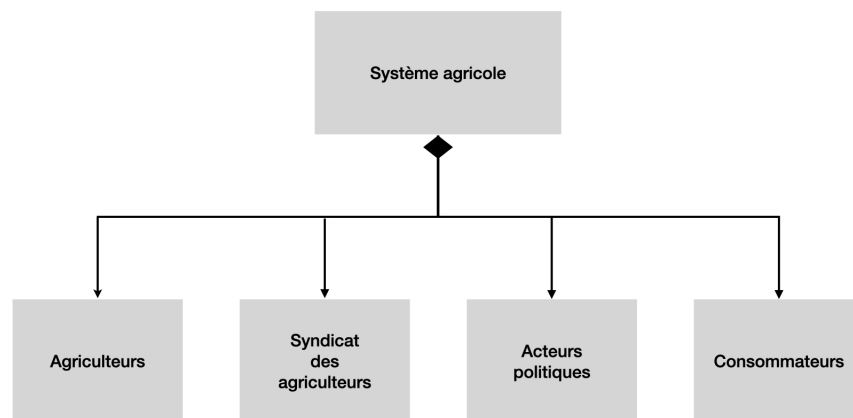


FIGURE 2.4 – Diagramme de définition des blocs d’un système agricole. Les différents acteurs composant le système sont représentés dans des sous-boîtes. Le losange en dessous du bloc “Système agricole” n’est là que pour respecter la syntaxe SysML.

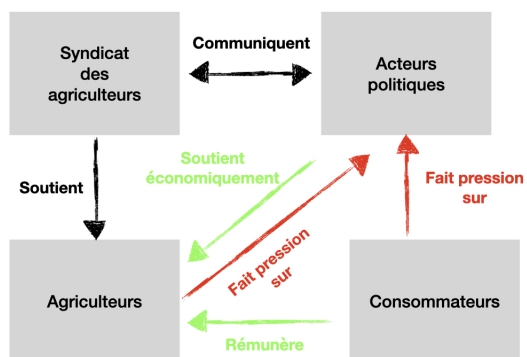


FIGURE 2.5 – Diagramme de bloc interne du même système. Sont représentés les interactions entre les différents acteurs. Un code couleur est utilisé pour décrire des types d’interactions particulières, il peut-être enrichi.

## 2.2.2 Les diagrammes fonctionnels

J'ai utilisé les diagrammes fonctionnels pour une visualisation des fonctions de chaque acteur (diagramme des cas d'utilisation) et des exigences sur chacune de ces fonctions (diagramme des exigences). Nous remarquons rapidement que dans un système, la fonction d'un acteur est souvent dépendante des fonctions des autres acteurs. Ainsi, on peut avoir une première idée des possibles réactions en chaîne lorsqu'une fonction est touchée. Nous obtiendrons alors les fonctions dépendante de la première qui seront affecté également et donc les acteurs associés à ses fonctions aussi. Ces aspects là se trouve principalement dans le diagramme des cas d'utilisation. Je n'ai utilisé le diagramme d'exigences que pour spécifier les fonctions. On peut alors se passer du diagramme d'exigences en première approximation dans le sens où il n'apporte pas d'information de nature différente dans la représentation si ce n'est des précisions.

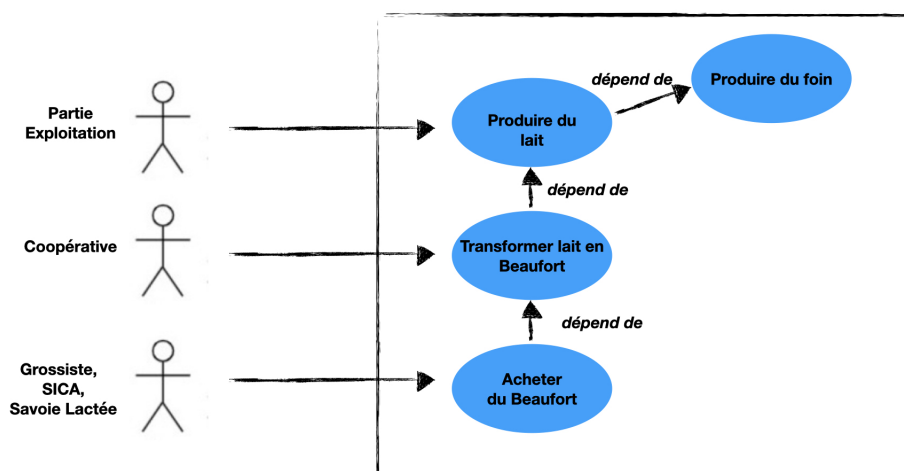


FIGURE 2.6 – Diagramme de cas d'utilisation de la filière Beaufort. On a distingué trois acteurs ayant chacun une fonction liée aux fonctions des autres acteurs. Les flèches entre acteurs et fonction signifie «a la fonction de».

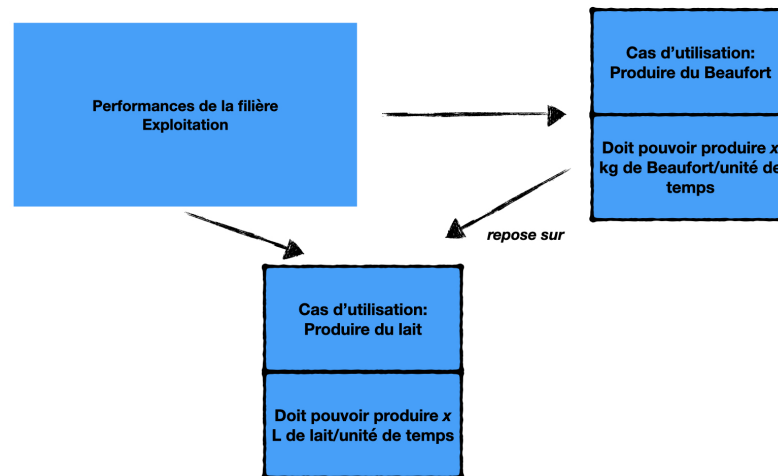


FIGURE 2.7 – Diagramme des exigences de la partie Exploitation de la filière Beaufort. Chaque cas d’utilisation est lié à une exigence précisé en dessous.

### 2.2.3 Le diagramme d’états

J’ai adapté les diagrammes d’états du langage SysML pour décrire la vulnérabilité d’un système. Ce diagramme, à l’origine, permet de décrire les différents états possibles du systèmes et d’expliquer les transitions entre états. J’ai trouvé ce diagramme très pertinent pour décrire la dynamique d’un système subissant un aléa. La figure 2.8 montre l’utilisation type du diagramme d’états dans le cas de la vulnérabilité du système. Pourquoi ce diagramme est-il pertinent lors d’une analyse de vulnérabilité ?

Comme expliqué dans la section 1.2, l’aspect chronologique et la description des différentes phases du subissement de l’aléa sont très importants, il permettent d’estimer la sensibilité et la résilience du système.

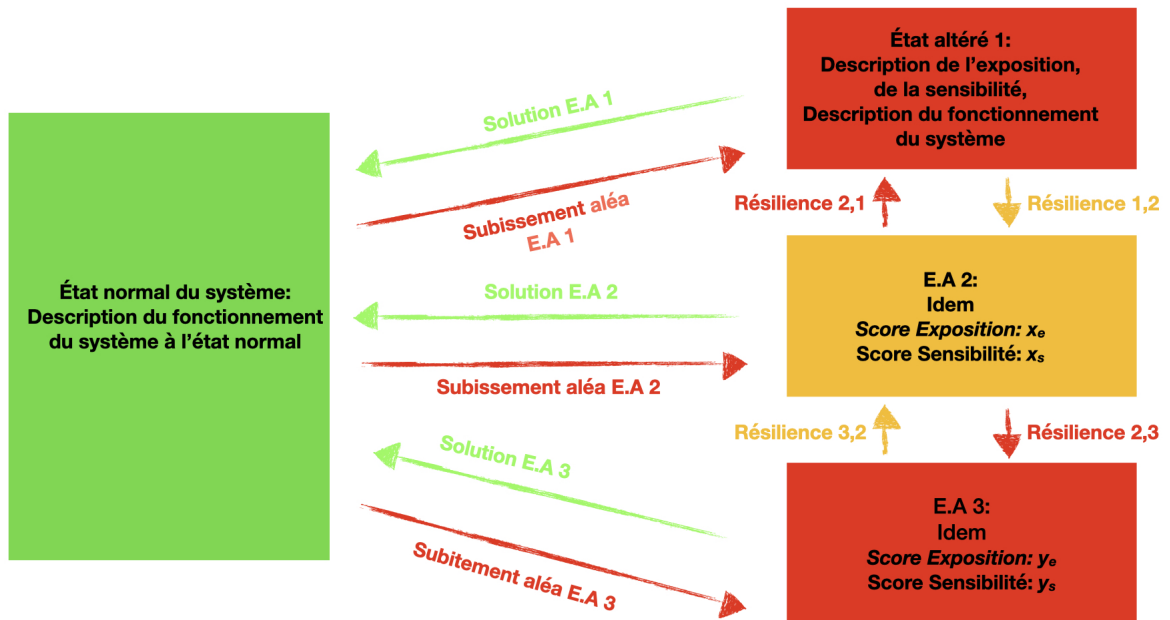


FIGURE 2.8 – Adaptation du diagramme d'états SysML pour modéliser la circulation du système entre différents états (un état normal et des états dits «altérés»). Les différents états représentent les conséquences des différentes manières de subir l'aléa. La circulation entre les différents états altérés et les manières de retrouver l'état normal, illustrent la résilience du système. Le code couleur des blocs : vert pour l'état normal, rouge pour un état altéré, orange pour un état considéré comme moins altéré. Il y a également la présence de score d'exposition et de sensibilité.

Cet outil de modélisation permet de comprendre comment le système circule d'un état à un autre et ce grâce aux flèches. En effet, lorsqu'utilisé par un expert, on pourrait voir sous quelle condition il serait possible de retourner à l'état normal (typiquement une mesure politique). Des scores d'exposition et de sensibilité peuvent être inclus dans cette modélisation, typiquement lorsqu'une quantité d'intérêt est exhibée et qu'elle n'a pas la même valeur dans chacun des états altérés. Illustrons par des exemples l'utilisation de ce diagramme.

**Exemple 2.2.1** La figure 2.9 décrit grossièrement la manière dont a été subi la crise du COVID-19 par la France. Les seules solutions pour revenir à l'état normal sont un médicament/vaccin trouvé rapidement ou l'atteinte de l'immunité collective. Pour atteindre l'immunité collective, une stratégie de «*stop and go*» pourrait être adoptée, consistant entre un cycle entre une situation de confinement (E.A 2) et de reconfinement (E.A 3).

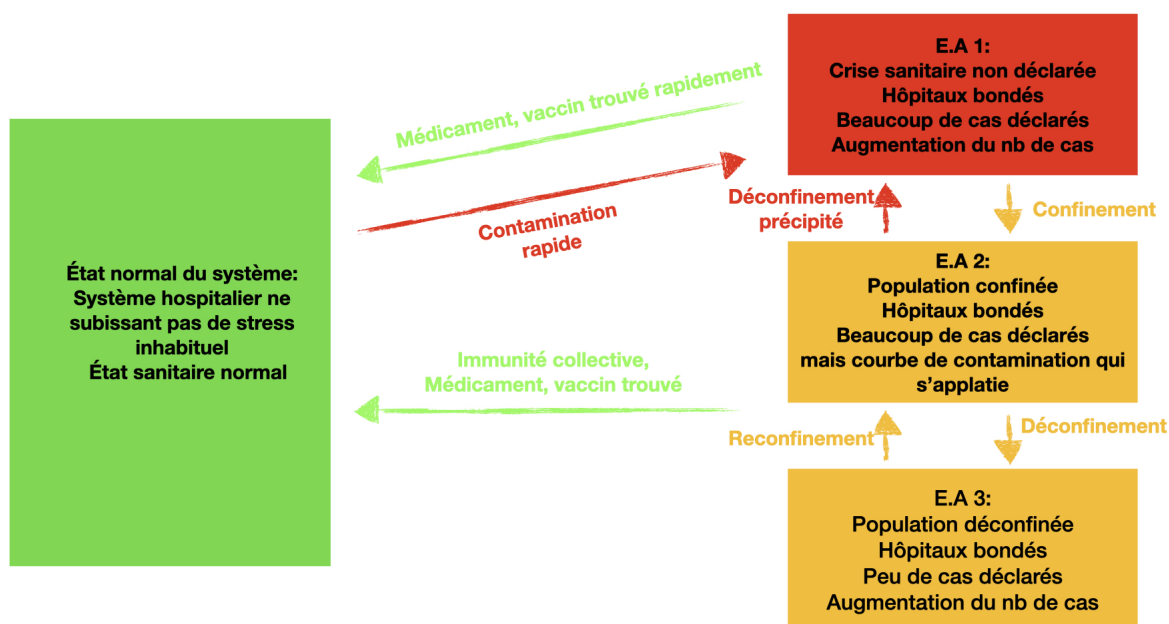


FIGURE 2.9 – Diagramme d'état d'une représentation de la crise du COVID-19 en France

**Exemple 2.2.2** Pour souligner que ces diagrammes sont souples à manipuler et modifiables, j'ai intégré les principes des diagrammes de Sankey dans les diagrammes d'états. J'ai utilisé l'épaisseur des flèches pour savoir comment sont modifiés les flux principaux sous l'effet de l'aléa. Ainsi nous pouvons voir sur ce diagramme que la réduction de la main d'œuvre et des terres utilisées, causée par un arrêt des exploitations fragiles, induit une baisse dans le litrage. Le système est alors considéré comme dans un état altéré car le système peine alors à remplir sa fonction de production de lait. Les chiffres donnés sont ceux présents dans l'étude de Caroline Vivier sur la filière du Beaufort.

L'aléa considéré ici étant l'arrêt des exploitations fragiles, le mot aléa doit se comprendre comme évènement affectant le système.



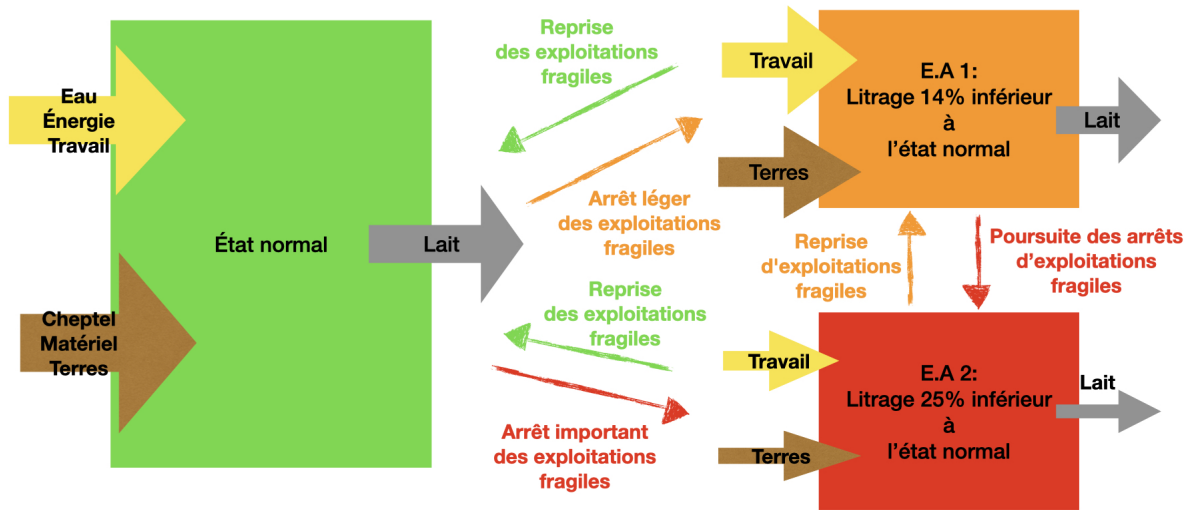


FIGURE 2.10 – Diagramme d'état de la partie Exploitation de la filière Beaufort soumis à l'aléa «arrêt des exploitations» (voir la partie Discussion ci-dessous sur cette notion d'aléa).

## 2.2.4 Le diagramme de séquence

J'ai utilisé le diagramme de séquence pour représenter des scénarios lorsque le système fait face à une vulnérabilité. Il est très riche car il permet de voir successivement les différentes actions avec les interactions entre acteurs. Il inclut également un aspect temporel assez complet.

L'autre avantage qu'il offre est qu'il permet de traiter les actions suivantes de manière probabiliste. En effet, il est possible de créer des boîtes "if", "elif", "else" exprimant le fait que des conditions sur le système implique des réactions différentes. Comme c'est de ces conditions que dépend l'évolution du système, il peut-être intéressant de donner les probabilités de réalisation de ces conditions. Un exemple est donné sur la figure 2.11.

Le diagramme de séquence permet également une assez bonne quantification du temps qui s'écoule lors du scénario. Les temps de transmission des messages sont supposés instantanés mais il est possible d'inclure des durées de réactions entre les messages. Ces durées peuvent être spécifiées dans le diagramme si elles sont connues.

L'exemple du diagramme de séquence de la filière du Beaufort est donné sur la figure 2.12.

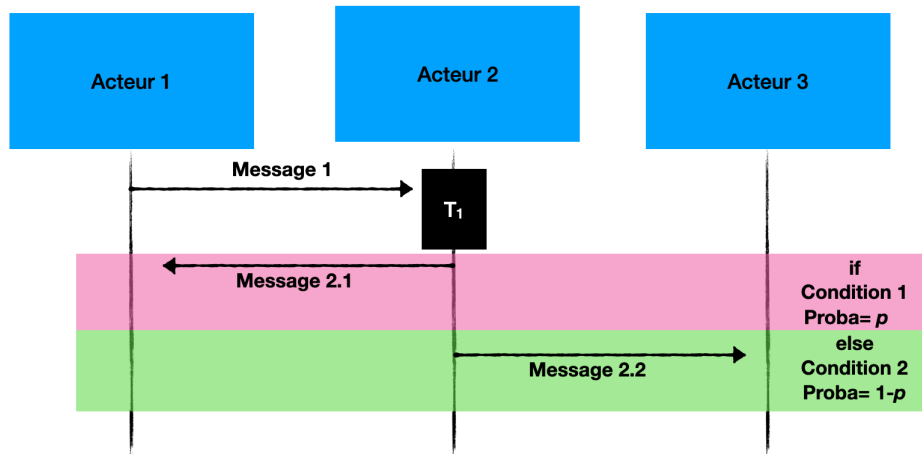


FIGURE 2.11 – Aspect probabiliste des diagrammes de séquence. Ici c'est le message 2.1 qui sera envoyé par l'acteur 2 si la condition 1 est réalisée, elle est réalisée avec une probabilité  $p$ . Ce sera le message 2.2 sinon.

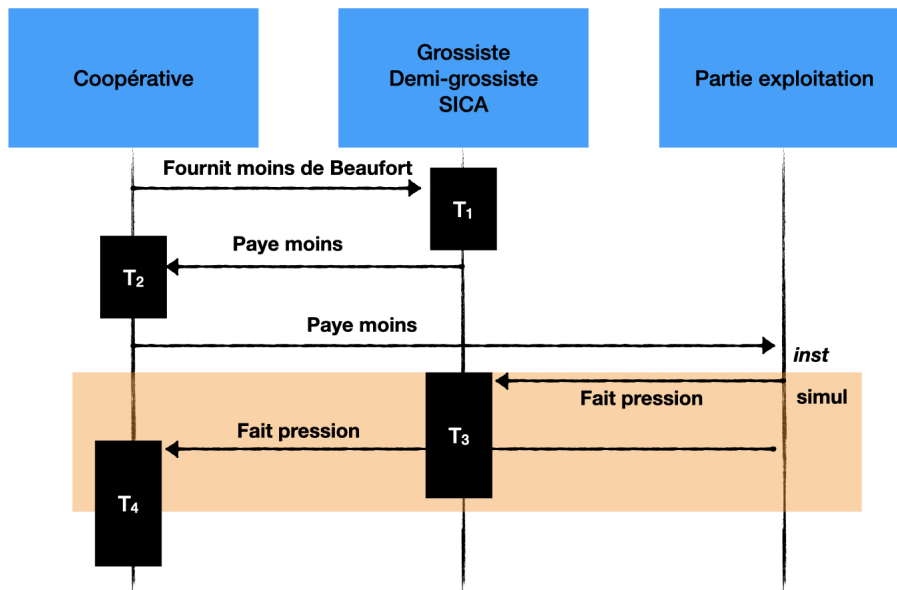


FIGURE 2.12 – Diagramme de séquence de la filière Beaufort face à la vulnérabilité des arrêts d'exploitations fragiles. Les blocs bleus indiquent les acteurs, les blocs noirs indiquent des temporalités  $T_i$  à renseigner, "inst" signifie que la réaction est instantanée et la boîte "simul" signifie que tous les messages faisant parti de cette boîte sont envoyés simultanément.

## 2.3 Discussion, perspectives

Ces diagrammes d'état révèlent bien les configurations possibles du système et on peut y inclure beaucoup d'informations complémentaires ; notamment plusieurs vulnérabilités différentes, y préciser la nature (exogène vs. endogène), le domaine (socio-culturel vs. biophysique), une échelle de temps sur laquelle on analyse notre système ; cependant le seul moyen que j'ai trouvé pour inclure l'ordre chronologique dans lequel on parcourt les états ou l'aspect probabiliste de ces scénarios, est le diagramme de séquences. Ce diagramme ne me semble pas satisfaisant et je pense qu'il est grandement améliorable.

Il ne me semble pas satisfaisant car j'ai le sentiment qu'il est possible d'avoir une représentation plus claire des différents scénarios avec d'autres idées. En effet, dans le diagramme de séquence si l'on exprime plusieurs "if" à l'intérieur d'une boîte "if", cela peut vite devenir illisible et la lecture de l'écoulement du temps devient mauvaise.

J'ai hésité à ne pas décrire mes travaux sur le diagramme de séquences de SysML J'ai finalement de l'inclure dans cette partie car je trouve qu'il permet tout de même de représenter des situations simples de manière assez complète.

Dans l'ensemble les diagrammes restent assez souples et peuvent inclure beaucoup d'outils (comme le diagramme d'états avec la vision Sankey). Ils ont également une méthodologie et une utilité de base qui guide le modélisateur pour obtenir une bonne vision du système. Ces diagrammes restent modifiables, et l'expert qui jugera utile d'y inclure une autre idée de modélisation trouvera sûrement le moyen de l'inclure dans les diagrammes définis dans ce chapitre.

Un autre inconvénient de ces diagrammes est le manque de «récursivité» (penser à l'aspect descendant de SADT) qu'on peut y inclure. Un des points importants que l'on souhaitait inclure était de représenter toutes les échelles dans la représentation, de pouvoir comprendre l'échelle d'un compartiment du système comme du système global et ce, juste par l'action d'un zoom dans le diagramme.

Je n'ai pas vraiment réussi à tirer pleinement profit du diagramme d'exigences. Je juge mon utilisation de ce diagramme comme trop élémentaire et je trouve qu'il n'apporte pas grand chose par rapport au diagramme de cas d'utilisation.

La meilleure façon d'inclure les fonds du système n'est également pas trivial à concevoir. J'ai, dans mes tentatives de représentation, essayé de représenter les fonds d'un système comme des acteurs, à l'intérieur du diagramme de définition des blocs, mais je ne crois pas que cela soit très satisfaisant. Peut-être doivent-ils être représentés comme un paramètre des états dans le diagramme d'état ? Peut-être figurent-ils seulement dans les diagrammes structurels ?

# Chapitre 3

## Réconciliation de données

Lors de mon stage, j'ai également pu m'intéresser aux liens entre les mathématiques appliquées et des problèmes de circulation des flux dans un système socio-technique.

En effet, lorsque l'on étudie un système au sein duquel, des flux circulent, il est souvent très intéressant de pouvoir les mesurer, malgré tout à une échelle macroscopique, il peut-être difficile de collecter toutes les données. De plus, comme les mesures effectuées sont soumises à des incertitudes, les valeurs correspondantes peuvent être fausses, dans le sens où elles ne satisfont plus les contraintes imposées par le système. Les valeurs correspondantes satisfont un système d'équations, grâce à cela, il devient possible de corriger les valeurs mesurées de l'erreur commise et d'assigner une valeur à des variables de notre système qui ne seraient pas mesurées en pratique.

Ce champ de l'analyse de données s'appelle la *réconciliation de données*. J'ai étudié deux articles [Cen16] et [CF18], le premier traitant de la réconciliation de données pour des contraintes linéaires avec des erreurs gaussiennes sur les mesures, basée sur la méthode des moindres carrés ; l'autre, plus générale, avec des contraintes et erreurs sur les mesures, quelconques, basé sur des méthodes de géométrie différentielle. Pour ce deuxième article, j'ai du étudier des notions de géométrie riemannienne pour comprendre l'article.

Mon travail a consisté à décrire mathématiquement les méthodes mathématiques utilisées dans ces articles qui n'est pas toujours très détaillée. J'en effectue une synthèse.

### 3.1 Contraintes linéaires et erreurs gaussiennes

#### 3.1.1 Cadre théorique

On dispose de  $x \in \mathbf{R}^{n_x}$ , le vecteur des  $n_x$  grandeurs mesurées,  $y \in \mathbf{R}^{n_y}$ , le vecteur des  $n_y$  grandeurs non-mesurées et  $z \in \mathbf{R}^{n_z}$  le vecteur des  $n_z$  valeurs constantes du système. On dispose d'un système  $f(y, x, z) = 0$  où  $f$  est une application linéaire de  $\mathbf{R}^{n_x+n_y+n_z}$  dans  $\mathbf{R}^{n_c}$ ,  $n_c$  étant le nombre de contraintes.

On pose note  $x$  le vecteur des valeurs réconciliées, *ie* satisfaisant les contraintes et  $\tilde{x}$ , les

vecteur des variables aléatoires des mesures. Alors on a :

$$\tilde{x} = \mu_x + \varepsilon$$

où  $\mu_x$  est l'espérance de  $\tilde{x}$ , inconnue et considérée comme la vraie valeur de la grandeur ; et  $\varepsilon \hookrightarrow \mathcal{N}(0, Q)$  où  $Q$  est la matrice de covariance de  $\tilde{x}$ , définie positive, considérée comme l'erreur de mesure de la grandeur.

Définissons le vecteur  $x$  des valeurs ajustées. On pose

$$F(x) = \left[ (x - \tilde{x})^T \right] Q^{-1} (x - \tilde{x})$$

$x$  est alors le vecteur minimisant la valeur de  $F$ , soit  $x = \arg \min F$ . La valeur de  $x$  est alors la solution d'un problème des moindres carrés pondérés. Les poids sont choisis comme l'inverse des variances, cela traduit la volonté d'accorder de l'importance à une mesure proportionnellement à sa précision.

La matrice  $Q$  est la matrice diagonale des variances des mesures, elle est estimée par l'analyste et requiert de pouvoir estimer quelle est l'erreur potentielle commise sur chacune des mesures.

On écrit  $f$  avec la formule de Taylor :

$$f(y, x, z) = J_y(\hat{y}, \hat{x}, z)(y - \hat{y}) + J_x(\hat{y}, \hat{x}, z)(x - \hat{x}) + f(\hat{y}, \hat{x}, z) = 0$$

$\hat{x}$  est l'estimation de  $x$ , initialement pris comme les mesures obtenues.  $\hat{y}$  est l'estimation des variables non-mesurées, elles doivent être choisies par l'analyste. Enfin  $J_x$  et  $J_y$  sont les matrices des dérivées partielles de  $f$  par rapport aux variables mesurées  $x$  et par rapport aux variables non-mesurées  $y$ . En notant  $\hat{f} = f(\hat{y}, \hat{x}, z)$  on obtient la réécriture matricielle suivante :

$$\begin{pmatrix} J_y & J_x & \hat{f} \end{pmatrix} \begin{pmatrix} y - \hat{y} \\ x - \hat{x} \\ 1 \end{pmatrix} = 0 \tag{3.1}$$

On va ensuite appliquer l'algorithme d'élimination de Gauss-Jordan pour rendre la matrice  $\begin{pmatrix} J_y & J_x & \hat{f} \end{pmatrix}$  échelonnée réduite. On rappelle sa définition :

**Définition 3.1.1**

Soit  $A \in \mathcal{M}_{n,p}(\mathbf{R})$ , on dit que  $A$  est **échelonnée réduite** si :

- si une ligne de la matrice est nulle, alors les lignes en dessous sont également nulles,
- la première valeur non-nulle d'une ligne non-nulle est 1,
- toutes les valeurs au dessus, d'un premier 1 d'une ligne sont des 0,
- la première valeur non-nulle d'une ligne apparaît toujours à droite de la première valeur non-nulle de la ligne précédente.

**Exemple 3.1.2** La matrice  $\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$  est une matrice échelonnée réduite tandis que  $\begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

ne l'est pas.

Une fois appliqué, l'algorithme de Gauss-Jordan, on obtient la forme échelonnée de réduite de  $(J_y \ J_x \ \widehat{f})$  qui est sous la forme

$$\begin{pmatrix} A_{cy} & A_{cx} & A_{cz} \\ 0 & A_{rx} & A_{rz} \\ 0 & 0 & A_{tz} \end{pmatrix} = (A_y \ A_x \ A_z) = \begin{pmatrix} A_c \\ A_r \\ A_t \end{pmatrix}$$

où  $A_y$  possède  $n_y$  colonnes,  $A_x$ ,  $n_x$  colonnes et  $A_z$ , 1 colonne.  $A_c$  est définie de telle sorte que  $A_{cy}$  ne contienne aucune ligne nulle et  $A_r$  est définie de telle sorte que  $A_{rx}$  ne contienne également aucune ligne nulle. On peut voir  $A_c$  comme les équations incluant des variables non-observées et  $A_r$  comme les équations n'incluant que les variables observées et les constantes.

Comme l'algorithme de Gauss-Jordan n'est composé que d'opérations sur les lignes, on conclut à l'existence d'une matrice  $B$  inversible tel que :

$$B \begin{pmatrix} A_{cy} & A_{cx} & A_{cz} \\ 0 & A_{rx} & A_{rz} \\ 0 & 0 & A_{tz} \end{pmatrix} = (J_y \ J_x \ \widehat{f})$$

Ainsi l'équation (3.1) devient

$$\begin{pmatrix} A_{cy} & A_{cx} & A_{cz} \\ 0 & A_{rx} & A_{rz} \\ 0 & 0 & A_{tz} \end{pmatrix} \begin{pmatrix} y - \widehat{y} \\ x - \widehat{x} \\ 1 \end{pmatrix} = 0$$

L'échelonnage de la matrice permet de réexprimer les contraintes en séparant le plus possibles les constantes des autres variables, puis les variables observées des variables non-observées.

Si  $A_{tz} = 0$ , alors il existe une contrainte ne dépendant que des constantes  $g(z)$  qui n'est pas égale à 0 et qui devrait l'être. Ainsi les constantes ne satisfont pas les équations, il faut donc les changer. Sinon  $A_{tz} \neq 0$  et dans ce cas, il existe des contraintes s'exprimant comme combinaisons linéaires d'autres contraintes ou/et il existe une contrainte  $g(z)$  ne dépendant que des constantes et qui vaut bien 0.

Si  $A_{rx} \neq 0$  alors on peut utiliser  $(A_{rx} \ A_{rz})$  pour la réconciliation de donnée pour obtenir l'équation :

$$A_{rx}(x - \widehat{x}) + A_{rz} = 0$$

Dans tous les cas  $A_t$  n'a plus d'importance, on peut donc l'oublier dans à partir de maintenant.

### **Théorème 3.1.3**

Si  $A_{rx} \neq 0$  alors on a l'expression suivante de  $x$  :

$$x = \tilde{x} - Q(A_{rx}^T)(A_{rx}QA_{rx}^T)^{-1}(A_{rx}(\tilde{x} - \widehat{x}) + A_{rz})$$

### **Lemme 3.1.4**

Soit  $Q \in \mathcal{S}_n(\mathbf{R})$ ,  $A \in \mathcal{M}_{n,p}(\mathbf{R})$  de rang  $p$ , la matrice  $(A^T)QA$  est inversible.

**Démonstration** (du lemme). Montrons l'existence d'une racine carrée de  $Q$  définie positive. Comme  $Q$  est symétrique définie positive, d'après le théorème spectral, il existe  $\lambda_1, \dots, \lambda_n \in \mathbf{R}_+^*$  et  $P \in \mathcal{O}_n(\mathbf{R})$  tel que,  $Q = (P^T) \text{Diag}(\lambda_1, \dots, \lambda_n) P$ . On peut poser  $R = (P^T) \text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) P$ . Ainsi  $R \in \mathcal{S}_n(\mathbf{R})$  est symétrique définie positive (car de valeurs propres strictement positives) et vérifie  $R^2 = Q$ .

Montrons que pour tout  $B \in \mathcal{M}_{n,p}(\mathbf{R})$ ,  $rg = ((B^T)B) = rg(B)$ .

C'est équivalent à montrer que  $\text{Ker}(B) = \text{Ker}((B^T)B)$ . Soit  $X \in \text{Ker}(B)$ ,

$$BX = 0 \Rightarrow (B^T)BX = 0 \Rightarrow X \in \text{Ker}((B^T)B)$$

Soit  $X \in \text{Ker}((B^T)B)$ ,

$$(B^T)BX = 0 \Rightarrow X^T B^T BX = 0 \Rightarrow \|BX\|^2 = 0 \Rightarrow BX = 0 \Rightarrow X \in \text{Ker}(B)$$

Appliquons ce résultat à  $RA$  et utilisons le fait que multiplier une matrice par une matrice inversible ne change pas le rang.

$$rg(A^T QA) = rg(A^T R^2 A) = rg(A^T R^T RA) = rg((RA)^T (RA)) = rg(RA) = rg(A) = p$$

Comme  $A^T QA \in \mathcal{M}_p(\mathbf{R})$ , cela montre que la matrice est inversible.  $\square$

**Démonstration** Il s'agit de minimiser la fonction  $F(x) = (x - \tilde{x})^T Q^{-1} (x - \tilde{x})$  sous les contraintes  $A_{rx}(x - \tilde{x}) + A_{rz} = 0$ . On introduit  $H(x, \lambda) = F(x) - \lambda^T [A_{rx}(x - \tilde{x}) + A_{rz}]$  pour tout  $\lambda \in \mathbf{R}^r$  (ce sont les multiplicateurs de Lagrange),  $r$  étant le nombre de contraintes (le nombre de lignes de  $A_r$ ). Pour ce problème d'optimisation, on résout ces deux équations :

$$\frac{\partial H}{\partial x}(x, \lambda) = 0 \tag{3.2}$$

$$\frac{\partial H}{\partial \lambda}(x, \lambda) = 0 \tag{3.3}$$

On peut réécrire  $H$  comme :

$$H(x, \lambda) = \sum_{1 \leq i, j \leq n_x} (x - \tilde{x})_i (x - \tilde{x})_j (Q^{-1})_{i,j} - \sum_{i=1}^r \sum_{j=1}^{n_x} \lambda_i (A_{rx})_{i,j} (x - \tilde{x})_j - \sum_{i=1}^r \lambda_i (A_{rz})_i$$

Or soit  $1 \leq k \leq n_x$ , on a (en utilisant le fait que  $Q^{-1}$  est symétrique) :

$$\frac{\partial H}{\partial x_k}(x, \lambda) = 2 \sum_{i=1}^{n_x} (Q^{-1})_{k,i} (x - \tilde{x})_i - \sum_{i=1}^r \lambda_i (A_{rx})_{i,k} = \left( 2Q^{-1}(x - \tilde{x}) - A_{rx}^T \lambda \right)_k = 0$$

D'où  $\frac{\partial H}{\partial x}(x, \lambda) = 2Q^{-1}(x - \tilde{x}) - A_{rx}^T \lambda = 0$ . On obtient ainsi que

$$x - \tilde{x} = -QA_{rx}^T \frac{\lambda}{2} \quad (3.4)$$

De plus l'équation (3.3) donne que :

$$\begin{aligned} A_{rx}(x - \hat{x}) + A_{rz} &= A_{rx}(x - \tilde{x}) + A_{rx}(\tilde{x} - \hat{x}) + A_{rz} \\ &= -A_{rx}QA_{rx}^T \frac{\lambda}{2} + A_{rx}(\tilde{x} - \hat{x}) + A_{rz} \\ &= 0 \end{aligned}$$

D'après le lemme précédent, comme  $rg(A_{rx}) = r$  (la matrice étant échelonnée réduite et ne contenant aucune ligne nulle) d'où  $rg(A_{rx}^T) = r$ ,  $A_{rx}QA_{rx}^T$  est inversible et

$$\frac{\lambda}{2} = \left(A_{rx}QA_{rx}^T\right)^{-1} \left(A_{rx}(\tilde{x} - \hat{x}) + A_{rz}\right)$$

En réinjectant cette équation dans l'équation (3.4), on obtient

$$x = \tilde{x} - QA_{rx}^T \left(A_{rx}QA_{rx}^T\right)^{-1} \left(A_{rx}(\tilde{x} - \hat{x}) + A_{rz}\right)$$

Cela conclut la démonstration. □

**Remarque 3.1.5** On vérifie aisément que cette valeur  $x$  vérifie bien les contraintes

$$A_{rx}(x - \hat{x}) + A_{rz} = 0$$

.

Maintenant que l'on a les valeurs réconciliées des variables observées, essayons d'estimer les valeurs non-observées. On reprend l'équation :

$$\begin{pmatrix} A_{cy} & A_{cx} & A_{cz} \\ 0 & A_{rx} & A_{rz} \end{pmatrix} \begin{pmatrix} y - \hat{y} \\ x - \hat{x} \\ 1 \end{pmatrix} = 0$$

Si  $A_{cy} = I_{n_y}$  (seule matrice échelonnée réduite inversible) alors les variables  $y$  sont observables et on a  $y = \hat{y} - A_{cx}(x - \hat{x}) - A_{cz}$ . Sinon,  $A_{cy} \neq I_{n_y}$ , dans ce cas il est nécessaire d'enlever des entrées dans  $A_{cy}$ . Toutes les lignes de  $A_c$  contenant au moins deux entrées non-nulles doivent être supprimées ainsi que toutes les colonnes contenant une valeur non-nulle à l'une des lignes supprimées. Ainsi la nouvelle matrice  $A_{cy}^*$  est l'identité, les variables  $y$  correspondant aux colonnes supprimées doivent également l'être dans  $y^*$  et  $\hat{y}^*$ . On obtient finalement l'expression suivante pour  $y^*$

$$y^* = \hat{y}^* - A_{cx}^*(x - \hat{x}) - A_{cz}^* \quad (3.5)$$



où  $A_{cx}^*$  représente la matrice  $A_{cx}$  une fois les modifications sur les lignes de  $A_c$  effectuées. Si les nouvelles estimations  $x$  et  $y^*$  sont en dessous d'un certain palier de convergence alors nous devons réitérer le processus avec les nouvelles valeurs de  $J_y$ , de  $J_x$  et de  $\hat{f}$  avec  $\hat{x}$  prend la valeur  $x$  et  $\hat{y}$  prend la valeur  $y$  (qui contient les anciennes valeurs pour les entrées qui n'ont pas changé et  $y^*$  pour celles qui ont changé).

Lorsqu'enfin le processus a "convergé", on peut conclure en calculant les matrices de covariance de  $x$  et de  $y^*$ . Le théorème 3.1.3, en passant à la matrice de covariance, donne :

$$Q_x = Q - QA_{rx}^T \left( A_{rx} Q A_{rx}^T \right)^{-1} \left( A_{rx} Q \right)$$

et par (3.5),

$$Q_y^* = A_{cy}^* Q_x A_{cy}^T$$

où  $Q_x$  est la matrice de covariance du vecteur  $x$  et  $Q_y$  celle du vecteur  $y$ .

**Démonstration** On applique le principe de propagation d'erreur si  $X$  admet une matrice de covariance  $Q$ , alors la matrice de covariance de  $AX$  où  $A$  est une matrice est  $AQA^T$ . Le deuxième cas est évident, montrons le premier

$$Q_x = Q - QA_{rx}^T \left( A_{rx} Q A_{rx}^T \right)^{-1} \left( A_{rx} Q \right) A_{rx}^T \left[ \left( A_{rx} Q A_{rx}^T \right)^{-1} \right]^T \left( A_{rx} \right)^T Q$$

Cela établit bien le résultat. □

Cela conclut cette partie de présentation théorique de la méthode.

## 3.2 Contraintes non-linéaires et erreur non-gaussienne

On retrouve dans [CF18] une méthode plus générale pour la réconciliation de données. Dans le premier cas, on utilisait des outils très simples (que de l'algèbre linéaire) mais moyennant des hypothèses assez fortes. En effet, les contraintes se devaient<sup>1</sup> d'être linéaires tandis que nos erreurs de mesure étaient gaussiennes.

Il existe une multitude de situations où les contraintes sont non-linéaires (économie, traitement du signal et plus généralement dans beaucoup de phénomènes physiques).

De plus supposer une erreur gaussienne est une hypothèse vraiment forte, car elle suppose que nos grandeurs sont non-bornées et peuvent prendre des valeurs négatives (ce qui n'est pas le cas par exemple pour des flux de masse.)

Pour généraliser cette méthode et alléger ces hypothèses, des méthodes différentielles vont être employées.

---

1. Il était possible d'inclure de la non-linéarité dans la première méthode, mais cette dernière est pensée pour être utilisé sur des contraintes linéaires en premier lieu.

### 3.2.1 Cadre et hypothèses

Soit  $n \in \mathbf{N}^*$ , on se fixe un vecteur  $(x, y) \in \mathbf{R}^n$  où  $x \in \mathbf{R}^{n_x}$  contient les variables observées et  $y \in \mathbf{R}^{n_y}$  les variables non-mesurées, on a donc  $n = n_x + n_y$ . Soit  $n_c \in \mathbf{N}$ , le nombre de contraintes indépendantes régissant nos variables. On suppose  $n_y \leq n_c \leq n$ . Ce système est résumé par l'équation suivante :

$$G(x, y) = 0 \text{ où } G : \mathbf{R}^n \longrightarrow \mathbf{R}^{n_c}$$

On effectue les hypothèses suivantes :

1. Les variables appartiennent à un sous-espace  $D \subset \mathbf{R}^n$
2. Les observations  $x$  constituent un vecteur aléatoire de dimension  $n_x$  de densité continue  $f d\lambda_{n_x}$
3. Le système d'équations définit une sous-variété  $S$  de  $\mathbf{R}^n$  de dimension  $n_w = n - n_c$ .
4. Il est possible de résoudre le système avec seulement  $n_w$  variables observées indépendantes et où les autres variables s'écrivent comme des fonctions  $\mathcal{C}^1$  de  $w$ .

La densité à priori des observations doit être estimée par un expert ou un analyste, elle reflète notre vision de la distribution de  $x$ .

### 3.2.2 Rappels de géométrie différentielle et riemannienne

J'ai dû me familiariser avec des concepts de géométrie riemannienne pour pouvoir effectuer les preuves rigoureuses des affirmations de l'article. J'ai choisi de présenter le résultat des mes recherches bibliographiques dans l'appendice A.

### 3.2.3 Explication de la méthode

On pose  $w \subset x$  les  $n_w = n - n_c$  variables observées indépendantes permettant de résoudre le système de manière unique. D'après l'hypothèse, on peut donc exprimer  $u$ , le sous-vecteur de  $x$  de taille  $n_u \leq n_x$  contenant les variables observées sauf celles contenues dans  $w$ , et  $y$  comme des fonctions différentiables de  $w$  (à noter que  $u$  est de taille  $n_u = n_x - (n - n_c) = -n_y + n_c = n_c - n_y$ .) D'où

$$\exists h, k \in \mathcal{C}^1 \text{ tels que } u = h(w) \text{ et } y = k(w)$$

La loi jointe des observations est donnée par une densité  $f(x)$  par rapport à la mesure de Lebesgue sur  $\mathbf{R}^{n_x}$ ,  $\lambda_{n_x}(x)$  (elle ne dépend pas des variables non-observées  $y$ ). Réconcilier les données c'est équivaler à conditionner la distribution par rapport à l'évènement  $\{u = h(w)\}$  puis de considérer la distribution obtenue. Les variables non-observées sont mises de côté par rapport aux variables observées, on peut alors juste considérer  $S$  comme une sous-variété de dimension  $n_w$  incluse dans  $\mathbf{R}^{n_x}$  par le plongement  $\varphi(w) = (h(w), w)$  pour tout  $w \in W$  (l'espace des valeurs possibles pour les variables  $W$ , qu'on supposera être un domaine de  $\mathbf{R}^{n_w}$ ).

**Théorème 3.2.1**

La loi  $W$  réconcilié (càd prenant en compte les contraintes) admet une densité par rapport à la mesure de Lebesgue  $\lambda_{n_w}$ .

**Remarque 3.2.2** La distribution à priori restreinte à  $S$  (aux contraintes) n'admet pas de densité par rapport à la mesure de Lebesgue  $\lambda_{n_x}$ ,  $S$  étant une variété de dimension  $n_w$  donc de mesure nulle d'après le théorème A.0.6, on dit que cette distribution est singulière par rapport à  $\lambda_{n_x}$ .

**Lemme 3.2.3**

Soit  $H$  la jacobienne de  $h$ ,

$S$  admet comme métrique riemannienne la forme bilinéaire symétrique définie positive donnée par la matrice  $G(y) = I_{n_w} + \left( H(\varphi^{-1}(y))^T \times H(\varphi^{-1}(y)) \right)$  pour tout  $y \in S$ .

**Démonstration** On considère la variété  $\varphi(W)$ . Comme  $(\varphi(W), \varphi^{-1})$  est une carte locale de tous les points, il est possible de l'utiliser pour calculer la métrique riemannienne  $g$  de  $\varphi(W)$  en tous les points (la métrique ne dépend pas de la carte choisie). Soit  $y \in \varphi(W)$ , la métrique  $g$  est définie par, d'après le théorème A.0.11 :

$$\begin{aligned}
 (g(y))_{i,j} &= \left\langle d(\varphi^{-1})^{-1}(\varphi^{-1}(y)).e_i, d(\varphi^{-1})^{-1}(\varphi^{-1}(y)).e_j \right\rangle_{\mathbf{R}^{n_x}} \\
 &= \left\langle d\varphi(\varphi^{-1}(y)).e_i, d\varphi(\varphi^{-1}(y)).e_j \right\rangle_{\mathbf{R}^{n_x}} \\
 &= \left\langle dh(\varphi^{-1}(y)).e_i, dh(\varphi^{-1}(y)).e_j \right\rangle_{\mathbf{R}^{n_u}} + \left\langle d\text{id}_W(\varphi^{-1}(y)).e_i, d\text{id}_W(\varphi^{-1}(y)).e_j \right\rangle_{\mathbf{R}^{n_w}} \\
 &= \left\langle \partial_{w_i} h(\varphi^{-1}(y)), \partial_{w_j} h(\varphi^{-1}(y)) \right\rangle_{\mathbf{R}^{n_u}} + \left\langle \partial_{w_i} \text{id}_W(\varphi^{-1}(y)), \partial_{w_j} (\text{id}_W(\varphi^{-1}(y))) \right\rangle_{\mathbf{R}^{n_w}} \\
 &= \left( H(\varphi^{-1}(y))^T \times H(\varphi^{-1}(y)) \right)_{i,j} + (I_{n_w})_{i,j}
 \end{aligned}$$

D'où la métrique riemannienne de  $S$  est bien celle attendue.

**Théorème 3.2.4 (Distribution à postériori)**

La distribution à postériori des observations  $w$  est donnée par la densité suivante :

$$f_S(w) d\lambda_{n_w}(w) = \frac{f(h(w), w) \sqrt{|I_{n_w} + H(w)^T \times H(w)|} d\lambda_{n_w}(w)}{\int_W f(h(w), w) \sqrt{|I_{n_w} + H(w)^T \times H(w)|} d\lambda_{n_w}(w)}$$

**Démonstration** En effet, comme on vient de trouver la métrique riemannienne de  $S$ , on obtient d'après le théorème A.0.12 un volume élémentaire,  $dS(w)$ , sur  $S$ . Comme  $(\varphi(W), \varphi^{-1})$  permet de recouvrir tout  $S$  et que  $f|_{\varphi(W)}$  est mesurable sur  $(M, \mathcal{B}(M))$  (étant mesurable car continue), on a

$$\int_{\varphi(w)} f dS = \int_W f(h(w), w) \sqrt{|I_{n_w} + H(w)^T \times H(w)|} d\lambda_{n_w}(w)$$

Ainsi, après normalisation on obtient bien le résultat voulu.  $\square$

**Obtention des autres valeurs réconciliées par MCMC** Seule la loi de  $w$  a été trouvée et non celle de  $u$  et  $y$ . De plus, si on souhaite simuler selon  $w$ , l'intégrale au dénominateur dans la densité de  $w$  peut-être difficile à calculer lorsque la dimension devient grande. Pour simuler selon la loi, de  $u$ ,  $y$  ou même de  $w$  une méthode par chaîne de Markov est très pratique. En effet si l'on parvient à simuler des valeur de  $w$  grâce à l'algorithme de Metropolis-Hastings, on aura pas besoin de calculer l'intégrale et on pourra également obtenir  $u = h(w)$  et  $y = k(w)$ .

La probabilité d'accepter de passer d'un état  $w_i$  à un état  $\dot{w}$  dans l'algorithme de M-H ici utilisé est

$$\alpha(w_i, \dot{w}) = \min \left( 1, \frac{f_S(\dot{w})p(w_i)}{f_S(w_i)p(\dot{w})} \right)$$

où  $p$  est la distribution de proposition, indépendante du point de départ qui serait les distributions les plus adaptées au contexte de la réconciliation de données [CF18]. Cela n'est pas vraiment discuté dans l'article, mais une distribution de proposition  $p$  adéquate doit être choisie (en fait, il suffit juste que  $p$  soit irréductible pour garantir le résultat, seulement la vitesse de convergence de la chaîne de Markov vers sa distribution stationnaire en dépend fortement, c'est donc un choix à réfléchir).

# Conclusion

## Récapitulatif

Mes contributions personnelles ont nécessité une quantité non-négligeable de travail préliminaire, à me familiariser avec les concepts de l'analyse systémique. Je n'aurai pas pu me lancer directement dans les représentations de SST sans cet état de l'art. Mes lectures sur l'analyse de vulnérabilité m'a fait comprendre de nombreux critères pour qualifier la fragilité d'un système. J'ai pu me familiariser avec les principes de l'ingénierie système comme SysML ou SADT.

Dans mes représentations de SST, que j'ai effectué en adaptant le langage SysML, beaucoup de choses et d'aspects différents ont pu être représentées. Ces diagrammes sont assez souples d'utilisation et peuvent être adaptables à beaucoup de situations. Des choses sont encore à améliorer et ce travail de modélisation dresse de nouvelles perspectives de travail.

En effet, les aspects chronologiques et probabilistes des diagrammes doivent être perfectionnés car seul le diagramme de séquences les fournit. D'autres diagrammes sont encore à améliorer mais le diagramme de séquence est selon moi celui qui mériterait le plus d'attention si un autre travail devait poursuivre ce que j'ai décrit dans ce mémoire. Enfin il reste à appliquer ces diagrammes à des situations complexes et connaître avec quel niveau de précision il est possible de décrire un SST. Cette étape permettrait sûrement de se rendre compte d'autres faiblesses de ces diagrammes.

## Remerciements

Je tenais à remercier très chaleureusement Guillaume et Jean-Yves, mes tuteurs qui ont été d'une grande implication envers moi durant ce stage. Leurs conseils, leur ouverture d'esprit et leur vision des choses m'a beaucoup apporté dans mes réflexions en rapport avec le stage mais également de manière générale sur ma conception de la science. Je les remercie également d'avoir bien voulu me relire et de m'avoir accompagné dans la correction.

Ce stage m'a permis de découvrir d'autres horizons scientifiques et d'essayer un autre domaine, un peu éloigné des maths. Je reste convaincu que la vision des mathématiques aide dans ce genre de réflexion et que quelques-unes, des gymnastiques habituelles de raisonnement, appartenant aux mathématiques, ont été les miennes durant ce stage.

J'ai également pu travailler sur des problématiques où je me sentais engagé. J'ai pu considérer le problème de l'urgence climatique de manière plus précise et y réfléchir sous un autre angle. J'ai choisi de présenter les informations, que j'ai pu collecter durant mon stage, très largement dans

l'introduction. Je trouve que cela illustre bien l'utilité du travail produit par l'équipe Steep et je suis heureux de pouvoir inclure ce genre de contenu dans un travail personnel comme celui-ci.

Personnellement, même si tout n'a pas été facile avec le COVID-19, donc l'intégralité du stage en visio-conférence, et le fait de travailler dans un domaine complètement nouveau (ce qui m'a mis en difficulté) ; je garde un très heureux souvenir et je suis très satisfait d'avoir pu effectuer ce stage dans l'équipe Steep. Je me réjouis également du fait qu'il soit possible pour des étudiants en mathématiques comme moi de pouvoir approcher ce genre de problématique scientifique. Je souhaiterais féliciter pour cela, les encadrants de mon master pour la liberté et la confiance qu'ils offrent aux étudiants.

# Annexe A

## Rappels de géométrie différentielle et riemannienne

### Définition A.0.1

Soit  $X$  un espace topologique,  $X$  est une **variété topologique** si  $X$  est séparé, à base dénombrable et si pour tout  $x \in X$ , il existe  $U$  ouvert de  $X$  tel que  $U$  soit homéomorphe à un ouvert de  $\mathbb{R}^n$  pour  $n \in \mathbf{N}^*$ . On appelle **dimension** de  $M$ , l'entier  $n$ .

Soit  $X$  une variété topologique, on appelle **carte** de  $X$ , tout couple  $(U, \phi)$ , tel que  $U$  soit un ouvert de  $X$  et  $\phi$  un homéomorphisme  $\phi : U \rightarrow \phi(U) \subset \mathbb{R}^n$ . Si  $x \in U$ , on dit alors que  $(U, \phi)$  est une carte contenant  $x$ . On appelle alors **atlas** de  $X$ , un ensemble de cartes  $\{(U_\alpha, \phi_\alpha), \alpha \in I\}$  recouvrant  $X$  (ie tel  $\bigcup_{\alpha \in I} U_\alpha = M$ ). L'atlas est dit de classe  $\mathcal{C}^k$  si pour toutes cartes  $(U_\alpha, \phi_\alpha)$  et  $(U_\beta, \phi_\beta)$  dans l'atlas tels que  $U_\alpha \cap U_\beta \neq \emptyset$

$$\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(U_\alpha \cap U_\beta) \rightarrow \phi_\alpha(U_\alpha \cap U_\beta)$$

est un  $\mathcal{C}^k$ -difféomorphisme. Deux atlas  $\mathcal{A}_1, \mathcal{A}_2$  sont dit **compatibles** si  $\mathcal{A}_1 \cup \mathcal{A}_2$  est encore un atlas (on vérifie que c'est bien une relation d'équivalence). On appelle **variété différentielle de classe  $\mathcal{C}^k$** , tout couple  $(X, [\mathcal{A}])$  où  $X$  est une variété topologique et  $[\mathcal{A}]$  une classe d'équivalence d'atlas de classe  $\mathcal{C}^k$  compatibles avec un atlas  $\mathcal{A}$  de classe  $\mathcal{C}^k$ .

### Définition A.0.2

Soit  $M$  une sous partie de  $\mathbf{R}^n$ , on dit que  $M$  est une **sous-variété** de dimension  $k$  et de classe  $\mathcal{C}^p$  si pour tout  $x \in M$ , il existe  $V \in \mathcal{V}_{\mathbf{R}^n}(x)$  et un  $\mathcal{C}^p$ -difféomorphisme  $\phi : V \rightarrow \mathbf{R}^n$  tel que

$$\phi(V \cap M) = \phi(V) \cap (\mathbf{R}^k \times \{0\}^{n-k})$$

**Remarque A.0.3** Une sous-variété de  $\mathbf{R}^n$  de classe  $\mathcal{C}^p$ ,  $M$ , est bien une variété différentielle de classe  $\mathcal{C}^p$  où l'atlas est obtenue en prenant tous difféomorphismes locaux munis de leur ensembles de définition. La compatibilité de ces cartes est triviale.

Pour la réciproque, on pourra mentionner le théorème de Whitney qui stipule que toute variété différentielle de dimension  $m$  peut être plongée dans  $\mathbf{R}^{2m}$  et être ainsi considérée comme une sous-variété.

**Définition A.0.4**

Soit  $(M, \mathcal{A}_M), (N, \mathcal{A}_N)$  deux variétés différentielles et  $f : M \rightarrow N$ , on dit que  $f$  est de classe  $\mathcal{C}^k$  si  $f$  est continue et si pour tout cartes  $(U_\alpha, \phi_\alpha) \in \mathcal{A}_M$  et  $(U_\beta, \phi_\beta) \in \mathcal{A}_N$

$$\phi_\beta \circ f \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap f^{-1}(U_\beta)) \rightarrow \phi_\beta(U_\beta)$$

est de classe  $\mathcal{C}^k$ .

**Exemple A.0.5** Tout ouvert de  $\mathbf{R}^n$  est une variété différentielle de classe  $\mathcal{C}^\infty$  de dimension respective  $n$  et  $\dim_{\mathbf{R}}(F)$ , muni des classes d'équivalences respectives

**Théorème A.0.6** ([Laf97] p.38) Soit  $M$  une sous-variété de  $\mathbf{R}^n$  de dimension  $d < n$ .  $M$  est de mesure nulle pour la mesure de Lebesgue  $\lambda_n$ .

On va tout d'abord montrer le lemme suivant :

**Lemme A.0.7** Soit  $U \subset \mathbf{R}^n$  un ouvert et  $f : U \rightarrow \mathbf{R}^n$  de classe  $\mathcal{C}^1$ . L'image par  $f$  d'un ensemble de mesure nulle est de mesure nulle.

**Démonstration** Soit  $A$  un ensemble de mesure nulle. Montrons que pour toute  $B \subset U$  boule fermée,  $f(A \cap B)$  est de mesure nulle. Comme  $f$  est de classe  $\mathcal{C}^1$ ,  $x \mapsto df(x)$  est continue. Ainsi  $K := \max_{x \in B} \|df(x)\| < +\infty$ . D'après l'inégalité des accroissement finis,  $f$  est  $K$ -lipschitzienne sur  $B$  pour  $\|\cdot\|$ . On en déduit que si  $C_\delta$  est un  $n$ -cube inclus dans  $B$  tel que  $\lambda(C) \leq \delta$  alors  $f(C)$  est un ensemble de mesure au plus  $K^n \delta$  ie  $\lambda_n(f(C)) \leq K^n \lambda_n(C)$ . Comme  $A \cap B \subset A$ ,  $A \cap B$  est de mesure nulle. Ainsi il est possible de recouvrir  $A \cap B$  par une infinité dénombrable de cubes  $C = \bigcup_{k \in \mathbf{N}} C_k$  tel que  $\sum_{k=0}^{+\infty} \lambda_n(C_k) \leq \varepsilon$  pour tout  $\varepsilon > 0$  ( $C$  n'est pas forcément inclus dans  $B$  mais il est inclus dans une boule fermée plus grande sur laquelle  $f$  est encore lipschitzienne et on pourra donc réappliquer le raisonnement suivant avec une autre constante de Lipschitz). Ainsi on a :

$$\lambda_n(f(A \cap B)) \leq \lambda_n\left(f\left(\bigcup_{k=0}^{+\infty} C_k\right)\right) \leq \sum_{k=0}^{+\infty} \lambda_n(f(C_k)) \leq K^n \sum_{k=0}^{+\infty} \lambda_n(C_k) \leq K^n \varepsilon$$

Comme cette inégalité est valable pour tout  $\varepsilon > 0$ , on en déduit que  $f(A \cap B)$  est de mesure nulle. On peut donc conclure en notant que  $\mathbf{R}^n = \bigcup_{n \in \mathbf{N}} \overline{B}(0, n)$  d'où  $A = \bigcup_{n \in \mathbf{N}} \overline{B}(0, n) \cap A$ . Ainsi  $A$  est une union dénombrable d'ensemble de mesure nulle et donc un ensemble de mesure nulle. Ainsi  $f$  envoie bien tout ensemble de mesure nulle sur un ensemble de mesure nulle.  $\square$

**Définition A.0.8** Un espace topologique  $X$  est dit **de Lindelöf** si pour tout recouvrement d'ouverts de  $X$ , il existe un sous-recouvrement dénombrable.

Tout espace topologique à base dénombrable est de Lindelöf.



**Démonstration (du théorème)** Pour tout  $x \in M$ , il existe un ouvert  $U \subset \mathbf{R}^n$  contenant  $x$  et  $\varphi_U : U \rightarrow \varphi_U(U)$  un difféomorphisme  $\mathcal{C}^1$  tel que  $\varphi_U(U \cap M) = \varphi_U(U) \cap (\mathbf{R}^d \times \{0\}^{n-d})$  (qui est alors évidemment un ensemble de mesure nulle). Ainsi en passant à  $\varphi_U^{-1}$ , on obtient :

$$U \cap M = \varphi_U^{-1} \left( \varphi_U(U) \cap (\mathbf{R}^d \times \{0\}^{n-d}) \right)$$

En appliquant le lemme précédent, on obtient que  $U \cap M$  est de mesure nulle. Comme  $M$  est à base dénombrable, c'est un espace de Lindelöf. Ainsi comme on possède un recouvrement de  $M$  par chaque ouvert de cartes, on peut en extraire un sous-recouvrement dénombrables  $M \subset \bigcup_{k \in \mathbf{N}} U_k$ , d'où :

$$M = \bigcup_{k \in \mathbf{N}} (M \cap U_k) = \bigcup_{k \in \mathbf{N}} \varphi_{U_k}^{-1} \left( \varphi_{U_k}(U_k) \cap (\mathbf{R}^d \times \{0\}^{n-d}) \right)$$

$M$  est donc une union dénombrable d'ensembles de mesure nulle et donc de mesure nulle.  $\square$

### Définition A.0.9

Soit  $M$  une sous-variété de  $\mathbf{R}^n$ , une **métrique riemannienne** sur  $M$  est la donnée pour tout  $p \in M$ , d'un produit scalaire sur  $T_p M$ ,  $g(p)$ , dépendant de manière lisse de  $p$ .

### Remarque A.0.10

• De manière lisse dans la définition signifie que pour tout  $(U, \varphi)$  une carte en  $p \in M$ , signifie que  $p \mapsto g(p) \left( d\varphi^{-1}(\varphi(p)).e_i, d\varphi^{-1}(\varphi(p)).e_j \right)$  est régulière ( $\mathcal{C}^k$  pour  $k \in \mathbf{N} \cup \{\infty\}$ ) pour tout  $i, j \in \{1, \dots, d\}$  où  $d$  désigne la dimension de  $M$  et  $(e_i)_{i \in \{1, \dots, d\}}$  est la base canonique de  $\mathbf{R}^d$ . Cette définition a bien du sens car on peut se rappeler du fait que si  $p \in M$  est un point de la variété et que  $(U, \varphi)$  est une carte contenant  $p$ , alors l'espace tangent  $T_p M$  peut s'écrire :

$$T_p M = d\varphi(p)^{-1} [\mathbf{R}^d \times \{0\}^{n-d}] = d\varphi^{-1}(\varphi(p)) [\mathbf{R}^d \times \{0\}^{n-d}] \quad (\text{A.1})$$

Ainsi,  $\left( d\varphi^{-1}(\varphi(p)).e_i \right)_{i \in \{1, \dots, d\}}$  forme une base de  $T_p M$ , comme image d'une base par un isomorphisme. Ainsi comme  $g(p)$  est une forme bilinéaire, il suffit de montrer que l'image d'une base de  $T_p M$  par  $g(p)$  est une fonction continue de  $p$ .

• Souvent lorsqu'une variété  $M$  est plongée dans  $\mathbf{R}^n$ , on munit  $M$  de la métrique riemannienne euclidienne, le produit scalaire usuel de  $\mathbf{R}^n$ . Cette métrique peut alors être calculée, elle est décrite par le théorème suivant.

### Théorème A.0.11

Soit  $M$  une sous-variété de  $\mathbf{R}^n$  de dimension  $d \leq n$ , on peut munir  $M$  de la métrique en  $p \in M$ ,  $g(p)$  suivante :

$$\left( g(p) \right)_{i,j} = \left\langle d\psi^{-1}(\psi(p)).e_i, d\psi^{-1}(\psi(p)).e_j \right\rangle$$

pour tout  $i, j \in \{1, \dots, d\}$  et où  $\psi$  correspond au difféomorphisme d'une carte locale  $(U, \psi)$  contenant  $p$ . La matrice  $g(p)$  est indépendante de la carte choisie.

**Démonstration** On a déjà mentionné que  $\left(d\psi^{-1}(\psi(p)) \cdot e_i\right)_{i \in \{1, \dots, d\}}$  était une base de  $T_p M$ .

Ainsi si l'on admet que  $\left(g(p)\right)_{i,j}$  ne dépend pas de la carte contenant  $p$  choisie, alors cela définit bien une métrique riemannienne sur  $M$ . En effet, le produit scalaire étant celui de  $\mathbf{R}^n$ , la régularité de  $g$  est celle de  $\psi$ .  $\square$

**Théorème A.0.12 (Volume élémentaire)**

Soit  $(M, g)$  une variété riemannienne de dimension  $n$ ,  
 $g$  induit une mesure  $dM(x)$  sur  $(M, \mathcal{B}(M))$  où  $\mathcal{B}(M)$  désigne la tribu borélienne de  $M$  appelée **volume élémentaire**  $dM(p)$  définie pour toute fonction mesurable  $f : U \rightarrow \mathbf{R}^n$ , pour toute carte  $(U, \psi)$

$$\int_U f dM = \int_{\psi(U)} f \circ \psi^{-1} \sqrt{\det G \circ \psi^{-1}} d\lambda_n$$

où  $G(x)$  est la matrice de  $g(x)$  en tant que forme bilinéaire.

# Bibliographie

- [Cen16] CENCIC, O. : Nonlinear data reconciliation in material flow analysis with software STAN. (2016). <https://www.sciencedirect.com/science/article/pii/S246820391630053X>
- [CF18] CENCIC, O.; FRÜHWIRTH, Rudolf : Data reconciliation of nonnormal observations with nonlinear constraints. (2018). <https://www.tandfonline.com/doi/pdf/10.1080/02664763.2017.1421916>
- [Del17] DELAÎTRE, M. : Échelles spatiales et approches méthodologiques pour l'analyse de la vulnérabilité. (2017). <http://journals.openedition.org/eps/7044>
- [Foe09] FOERSTER, E. : *Vulnérabilité : état de l'art sur les concepts et méthodologies d'évaluation. Rapport final.* <http://infoterre.brgm.fr/rapports/RP-57471-FR.pdf>. Version : 2009
- [Fü07] FÜSSEL, H-M. : Vulnerability : A generally applicable conceptual framework for climate change research. (2007). <https://www.d.umn.edu/~pfarrell/Natural%20Hazards/Readings/vulnerability%20climate%20change.pdf>
- [Gia14] GIAMPIETRO, M. : *An Innovative Accounting Framework for the Food-Energy-Water Nexus.* Food and Agriculture Organization, 2014 <http://www.fao.org/3/i3468e/i3468e.pdf>
- [GS18] GERBER, J.F; SCHEIDEL, Arnim : In Search of Substantive Economics : Comparing Today's Two Major Socio-metabolic Approaches to the Economy—MEFA and MuSIASEM. (2018)
- [Laf97] LAFONTAINE, J. : *Introduction aux variétés différentielles.* Presses Universitaires de Grenoble (PUG), 1997
- [LG13] L. GENDRE, J-M. V. : *Systems Modeling Language SysML*, 2013
- [MG08] M. GIAMPIETRO, J. Ramos-Martin K. M. K. Mayumi : Multi-scale integrated analysis of societal and ecosystem metabolism (MuSIASEM) : Theoretical concepts and basic rationale. (2008)
- [Tur03] TURNER, B.L. : A framework for vulnerability analysis in sustainability science. (2003). [https://www.researchgate.net/publication/10720959\\_A\\_framework\\_for\\_vulnerability\\_analysis\\_in\\_sustainability\\_science](https://www.researchgate.net/publication/10720959_A_framework_for_vulnerability_analysis_in_sustainability_science)
- [Wan16] WANG, Z. : Lecture 3 : The Riemannian Measure. (2016). <http://staff.ustc.edu.cn/~wangzuoq/Courses/16S-RiemGeom/Notes/Lec03.pdf>