

168753

T.C. GALATASARAY ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

BULANIK REGRESYON VE BULANIK OPTİMİZASYONUN
TEKNOLOJİ SEÇİMİ PROBLEMLERİNE UYGULAMALARI
(LES APPLICATIONS DE REGRESSION FLOUE ET D'OPTIMISATION FLOUE
AUX PROBLEMES DE SELECTION DE LA TECHNOLOGIE)

YÜKSEK LİSANS TEZİ

End. Müh. Zeynep YILMAZ

Tezin Enstitüye Verildiği Tarih : 20 Mayıs 2005

Tezin Savunulduğu Tarih : 02 Haziran 2005

Tez Danışmanı : Prof. Dr. E. Ertuğrul KARSAK

E. Karşak 28/06/2005

Diğer Jüri Üyeleri : Prof. Dr. H. Ziya ULUKAN

H. Ziya Ulukan 28/06/2005

Yrd. Doç. Dr. C. Ruhi TUNCER

R. Tuncer 28/06/2005

HAZİRAN 2005

Préface

L'analyse de régression classique est une méthode ancienne, largement utilisée dans la plupart des domaines scientifiques. Pendant les deux dernières décennies, la méthode de régression classique s'est évaluée donnant naissance à une nouvelle méthode: la régression floue. Comme la théorie des sous-ensembles flous est une théorie mathématique qui prend en compte l'imprécision et les incertitudes qui règnent sur les événements naturels de la vie, la régression floue permet de modéliser les systèmes influencés par les jugements humains ou ceux qui ne sont pas bien définis.

Ce travail est consacré à l'utilisation de la méthode de régression linéaire floue dans les deux problèmes de sélection de la technologie importants de l'industrie: le problème de sélection de robot industriel et le problème de sélection de système flexible de fabrication. Des modèles de décision basé sur la régression linéaire floue sont développés et les solutions sont obtenues à l'aide des techniques variées d'optimisation floue.

Tout d'abord, je tiens à remercier à Prof. Dr. E. Ertuğrul Karsak, le directeur de cette thèse, qui m'a permis d'enrichir mes idées et de les élaborer en m'encourageant au cours de ce long travail. J'exprime toute ma gratitude à lui, pour le temps qu'il m'a consacré, pour ses critiques et pour sa patience.

Je tiens aussi à exprimer ma reconnaissance à ma famille et à mes amis pour leur soutien et leur aide.

Zeynep YILMAZ

Mai, 2005

Table des matières

Liste des figures	v
Liste des tableaux	vi
Résumé	vii
Özet	x
1. Introduction	1
2. La régression floue	3
2.1. Généralités sur l'analyse de régression	3
2.2. La régression statistique	4
2.3. La régression floue	5
2.3.1. La théorie des sous-ensembles flous	5
2.3.1.1. Le concept d'incertitude	5
2.3.1.2. D'un ensemble classique à un ensemble flou	6
2.3.1.3. Définitions de base	7
2.3.2. La régression floue: Généralités et survol de la littérature	10
2.3.3. La régression linéaire floue: Le modèle de Tanaka	16
2.4. Les propriétés de la régression linéaire floue	21
2.4.1. La régression linéaire floue versus la régression linéaire statistique	21
2.4.2. Vue générale sur les modèles de régression linéaire floue	22
3. Le problème de sélection de robot industriel	31
3.1. La performance des robots	32
3.2. Sélection de robot	35
3.2.1. Modèles statistiques	36
3.3. Application de la régression linéaire floue à la sélection de robot industriel	37
3.3.1. Méthodologie	38
3.3.2. Résultats et conclusion	43
3.4. Modèle de décision basé sur la régression linéaire floue pour la sélection de robot industriel	48
3.4.1. Préliminaires	48

3.4.2. Méthodologie	49
3.4.3. Résultats de l'application	49
4. Evaluation des alternatives de technologie à l'aide de la méthode de régression linéaire floue	52
4.1. Introduction aux systèmes flexibles de fabrication	52
4.2. Les modèles d'évaluation des systèmes flexibles de fabrication	53
4.3. Les critères utilisés	54
4.4. La méthodologie proposée	56
4.4.1. Utilisation de la régression linéaire floue	56
4.4.2. Sélection de SFF par la programmation booléenne floue	57
4.4.2.1. Méthodes de classement des sous-ensembles flous	58
4.4.2.1.1. Méthode de Chu et Tsao	59
4.4.2.2. Modèle de programmation booléenne floue	60
4.4.3. Sélection de SFF par l'optimisation de l'intervalle de la fonction objectif	61
4.5. Application de la méthodologie	62
5. Conclusion	67
Bibliographie	69
Appendices	78
Appendice A	78
Appendice B	80
Appendice C	83
Appendice D	84
Appendice E	87
Biographie	89

Liste des figures

Figure 2.1	: Un nombre triangulaire flou	9
Figure 2.2	: Un coefficient flou triangulaire symétrique	18
Figure 2.3	: La représentation de l'algorithme de la régression linéaire floue	20
Figure 3.1	: Répétabilité et exactitude du robot	34
Figure 3.2	: Caractéristiques d'ingénierie des robots et performance des systèmes de production	36



Liste des tableaux

Tableau 2.1 : Survol de la littérature sur la régression floue	11
Tableau 3.1 : Caractéristiques utilisées pour l'évaluation des robots	33
Tableau 3.2 : Les valeurs de 4 critères pour 27 robots	40
Tableau 3.3 : Les valeurs utilisées dans les modèles	41
Tableau 3.4 : Résultats du modèle de Tanaka	44
Tableau 3.5 : Résultats du modèle de Peters	45
Tableau 3.6 : Calcul du coefficient de la corrélation du rang de Spearman pour le modèle de Tanaka	46
Tableau 3.7 : Calcul du coefficient de la corrélation du rang de Spearman pour le modèle de Peters	47
Tableau 3.8 : Nouveaux résultats du modèle de Peters	50
Tableau 4.1 : Critères considérés pour la sélection de système de fabrication	55
Tableau 4.2 : L'ensemble de données utilisé pour la régression floue (cas1)	62
Tableau 4.3 : L'ensemble de données utilisé pour la régression floue (cas2)	63
Tableau 4.4 : Les valeurs estimées du débit	64
Tableau 4.5 : Les valeurs de e_i	64
Tableau 4.6 : Les valeurs obtenues par la méthode de Chu et Tsao	65
Tableau 4.7 : L'espace occupé par chaque SFF et leurs coûts	65

Résumé

Les incertitudes présentes dans les événements quotidiens ne permettent pas l'utilisation des techniques classiques de modélisation. La complexité et l'incertitude de la vie vient de l'imprécision du raisonnement humain. Les systèmes qui sont influencés par la pensée humaine deviennent en général plus difficiles à comprendre et à interpréter. Comme les méthodes classiques restent insuffisantes dans le cas de modélisation des systèmes complexes et mal définis, un nouveau concept a été introduit qui contenait les domaines de l'incertitude ou de l'imprécision. La théorie des sous-ensembles flous est une théorie mathématique qui prend en compte de l'imprécision et des incertitudes qui règnent sur les événements naturels de la vie. Cette théorie est capable d'interpréter des informations imprécises et d'agir. De nos jours les applications de cette théorie sont primordiales dans un domaine assez vaste de l'ingénierie et des sciences appliquées.

L'analyse de régression est considérée comme une technique indispensable dans la plupart des domaines scientifiques. Elle est utilisée pour modéliser les relations entre les variables. La technique est basée sur le calcul des coefficients (les paramètres du modèle) qui décrivent à quelle mesure les variables dites indépendantes sont liées à une variable qui dépend d'elles, nommée variable dépendante. L'analyse de régression permet de décrire les relations entre les variables par estimation des paramètres du système et de prédire les valeurs de la variable dépendante en utilisant les valeurs des variables indépendantes.

L'analyse de régression classique est une méthode ancienne, largement utilisée dans presque chaque matière. Dans les deux dernières décennies, la méthode de régression classique s'est évaluée donnant naissance à une nouvelle méthode: la régression floue. Grâce à cela, la régression a ajouté à son domaine d'application les systèmes flous. Elle est maintenant utilisée pour décrire les événements imprécis en même temps que les événements précis. Afin de faire une distinction entre les deux méthodes de

régression, la régression classique basée sur la théorie classique ensembliste est appelée aussi la régression statistique.

La régression classique fait des hypothèses strictes sur les propriétés statistiques du modèle qui ne sont pas facilement être justifiées si l'ensemble de données n'est pas suffisamment large. La violation des ces règles peut affecter la validité et la performance de la régression statistique. La régression floue est plus effective que la régression linéaire statistique quand les hypothèses ne peuvent pas être employées proprement, si par exemple les jugements humains ou des processus ambigus sont présents.

La régression floue a été proposée en 1982 par Tanaka et ses collègues. Depuis cette date la régression floue a évolué par la création de nombreux modèles par différents auteurs. Dans ce travail les modèles de régression linéaire floue sont utilisés. Le but de régression linéaire flou est de déterminer les paramètres flous estimés à condition que le degré d'appartenance des valeurs de la variable dépendante à ses valeurs floues estimées soit au moins égal à un degré cible de confiance, qui est choisi par le décideur. Le but de la méthode est de déterminer les coefficients flous afin que la valeurs estimées de la variable dépendante soient la moins floue possible.

Ce travail est consacré à l'utilisation de la méthode de régression linéaire floue dans les deux problèmes importants de l'industrie: le problème de sélection de robot industriel et le problème de sélection de système flexible de fabrication.

La sélection de robot industriel constitue la source des travaux publiés depuis un certain temps. Il existe différentes catégories de modèles pour trouver une solution au problème donné. Afin de montrer la validité d'utilisation de régression linéaire floue pour la sélection de robot, les résultats obtenus par la régression linéaire floue sont comparés avec ceux obtenus par un travail déjà fait en utilisant la régression statistique robuste. Leur méthodologie est appliquée à la régression linéaire floue et la corrélation des résultats est déterminée par le test de corrélation du rang de Spearman. La valeur du coefficient de corrélation calculée montre une relation positive entre le classement

des robots suivant les deux méthodes citées, et les résultats obtenus sont semblables. Les différentes techniques de régression linéaire floue sont aussi utilisées dans un contexte d'un modèle de décision présenté en détails. Cette étape a pour but de montrer que la régression linéaire floue peut être utilisée pour la sélection de robot industriel au lieu de la régression statistique, aux cas où les paramètres sont flous, ou l'incertitude règne sur le système étudié, ou encore l'ensemble de données n'est pas suffisant pour pouvoir appliquer la régression statistique classique.

Dans l'étape suivante qui représente l'utilisation plus générale de la régression linéaire floue, le problème de sélection de système flexible de fabrication est traité. Un nouveau modèle de sélection basé sur la régression linéaire floue est proposé. Ce modèle est établi en se servant de deux méthodes distinctes d'optimisation, la programmation booléenne floue et la programmation multiobjectif dans le cas des fonctions objectif sous forme d'intervalles. Comme le modèle nécessite les valeurs floues de classement des alternatives, les méthodes de classement des nombres flous sont aussi étudiées. Les résultats de deux méthodes sont analysés pour en tirer des conclusions.

Enfin, la régression floue permet de modéliser les cas où les relations entre les paramètres du système ne sont pas bien définis et les variables sont données sous forme des nombres flous. C'est grâce à cette capacité que la méthode de régression floue est une alternative viable pour la résolution des problèmes de sélection de la technologie.

Özet

Günlük yaşantımızda karşılaştığımız olayların çoğunun özünde varolan belirsizlikler, modelleme aşamasında kesin tekniklerin kullanılmasına imkan vermemektedir. İnsan düşünce ve sezgisinin rol oynadığı sistemler, verilerin genellikle kesin olarak elde edilememesi, değişkenler arası ilişkilerin belirlenememesi gibi nedenlerle daha karmaşık bir yapıya bürünmektedir. Karşılaşılan bu zorlukların üstesinden gelmek için geliştirilen bulanık küme teorisi, insan düşünce yapısının bulanık özelliğinin etkili olduğu sistemleri modelleme becerisine sahiptir.

İlk başlarda kullanım alanı sınırlı kalmış olan bu teori daha sonralarda bilim dünyasının hemen her alanında yapılan çalışmalarda kullanılmış ve başarılı sonuçlar alınmıştır. Günümüzde bulanık küme teorisi, kesin kuralların olmadığı ya da kesin sınırlarla belirlenmemiş olayların konu edildiği çok çeşitli problemlerin çözümünde, sayısız teknikle beraber kullanılmakta ve bu sayede oluşturulan modeller gerçeği daha iyi yansıtmaktadır.

Değişkenler arasındaki ilişkileri modellemede kullanılan klasik regresyon analizi, çok sayıda bilim dalı için temel araçlardan biridir. Bu teknikle, bağımlı ve bağımsız değişkenler arasındaki ilişkinin fonksiyonel olarak verilmesi sağlanarak sistem değişkenleri arasındaki ilişkinin incelenmesine ve sistemin olası durumları için tahmin yapılmasına imkan verilmektedir.

Klasik regresyon analizi modelin istatistiksel özellikleri hakkında bazı varsayımlar gerektirmektedir. Fakat bu varsayımların sağlanması, her koşulda kolay olmamaktadır. Sistem yapısındaki belirsizliğe bağlı olarak verilerin tamamının ya da bir kısmının kesin sayılar olarak elde edilememesi veya sistem yapısının değişkenler arasında kesin ilişkiler tanımlanmasına imkan vermemesi gibi durumlarda klasik regresyon analizinin uygulanması mümkün değildir. Sözü geçen durumların da içinde olduğu regresyon analizinin gerektirdiği şartların bazılarının sağlanamadığı birçok durumda, modelin

performansının düşmesi ve kötü sonuçlar vermesi sebebiyle klasik regresyon analizinin uygulanması önerilmemektedir.

Klasik regresyon analizinin gerektirdiği şartların sağlanamadığı ve belirsizliğin hakim olduğu durumlarda bulanık regresyon analizi etkili bir araç haline gelmektedir. Bulanık regresyonun, değişkenlerin bulanık sayılar olarak ifade edilmesi ya da sistem parametreleri arasındaki ilişkilerin net olarak tanımlanamaması halinde kullanılması önerilmektedir. Klasik regresyon analizini bulanık regresyon analizinden ayırmak için, klasik regresyon analizinden, barındırdığı istatistiksel varsayımlardan yola çıkarak literatürde istatistiksel regresyon analizi olarak bahsedilmektedir. Bu çalışmada da kullanılan istatistiksel regresyon terimi klasik küme teorisine dayalı regresyon tekniğini işaret etmektedir.

İlk olarak 1982 yılında Tanaka ve diğerleri tarafından ortaya atılan bulanık regresyon analizi bulanık ya da deterministik veri kümeleri için bulanık parametreler hesaplamaktadır. Bulanık regresyon analizinin, değişkenlerin bulanık sayılar halinde ifade edilmesi ya da sistem parametreleri arasındaki ilişkilerin net olarak tanımlanamaması halinde kullanılması önerilmektedir. Bulanık regresyon analizi belirsizlik içeren problemlerin çözümü için farklı yazarlar tarafından çeşitli regresyon modelleri kullanılarak geliştirilmiş ve bunlara göre çeşitli isimler almıştır.

İstatistiksel regresyon analizi, regresyon fark terimlerinin gözlem hatalarından kaynaklanan, sıfır ortalamalı bağımsız rastgele değişkenler olduğunu varsaymaktadır. Oysa bulanık regresyon modelinde, istatistiksel regresyonun aksine, ortaya çıkan sapmaların modelin kendi yapısındaki belirsizliğe bağlı olduğu düşünülmektedir.

Bu çalışmada kullanılan bulanık regresyon modelleri, bulanık doğrusal regresyon modelleri olup, literatürde uygulamalarına yaygın olarak rastlanmakta ve bu modeller uygulama kolaylığı sebebiyle birçok araştırmacı tarafından tercih edilmektedir. Çalışmanın ilerleyen bölümlerinde farklı bulanık doğrusal regresyon modelleri incelenerek özellikleri belirtilmiş, bunun yanı sıra şu ana kadar bulanık regresyon

analizinin kullanıldığı yayımlanmış makalelerin incelendiği bir literatür çalışması yapılmıştır.

Bulanık doğrusal regresyon analizi, bağımlı değişkenlerin gözlemlenen her değerinin, tahmin edilen bağımlı değişken bulanık kümesindeki üyelik derecesinin, karar vericinin belirlediği bir eşik değerinin üzerinde olması şartıyla bulanık regresyon katsayıları hesaplamaktadır. Bulanık regresyon analizinde belirlenen parametrelerin bulanık olmalarından dolayı, bu parametreleri kullanarak hesaplanan ve tahmin edilen bağımlı değişkenin değerleri de bulanık olacaktır. Bulanık doğrusal regresyon probleminde istenen belirtilen bu kısıt altında, tahmin edilen bağımlı değişkenin değerlerinin bulanıklığının en düşük seviyeye indirilmesidir. Bulanık doğrusal regresyon problemi bir doğrusal programlama modeline dönüştürülerek çözülebilmekte ve bu kullanıcı açısından büyük kolaylık sağlamaktadır.

Bu çalışmanın uygulama kısmının ilk bölümünde, endüstride oldukça sık karşılaşılan ve literatürde çok sayıda uygulaması bulunan endüstriyel robot seçimi probleminde bulanık doğrusal regresyon analizinin alternatif bir karar destek yöntemi olarak kullanılması irdelenmektedir.

Mühendislik bilimlerinde ve bilgisayar teknolojilerinde gözlemlenen hızlı değişimler sonucunda, endüstriyel robotlar ileri üretim sistemlerinin çoğunda yer almaya başlamıştır. Robotlar, tekrarlanan ya da karmaşık işleri hatasız yaparak, kalitenin iyileştirilmesinin yanında maliyetin düşürülmesinde de rol oynamaktadırlar. Üreticiler, özellikle insan gücünün yeterli olmadığı ya da insan sağlığını tehdit eden işlerin yerine getirilmesinde robotlardan faydalanmaktadırlar. Endüstriyel robotların genellikle pahalı olmaları ve çok sayıda özelliklerinin bulunması, robot alternatiflerinin ihtiyaçlara göre dikkatle incelenmesini gerektiren bir seçim sürecine yol açmaktadır. Robot seçimi problemlerinin güçlüğü, robot özelliklerinin özünde birbirleriyle çelişmesi ve aralarındaki ilişkilerin net olarak ifade edilememesi gibi nedenlerle artmaktadır. Bu sebeple seçim aşamasında sayısal tekniklerin kullanılması fayda sağlamaktadır.

Endüstriyel robotların sayı ve çeşitliliği, kullanımının yaygınlaşması ve kullanım alanlarının genişlemesi sonucunda belirgin ölçüde artış göstermiştir. Günümüzde potansiyel bir robot kullanıcısı farklı performans ve maliyetteki çok sayıda alternatifi değerlendirmek zorunda kalmaktadır. Bu durum, karar vericilere robot seçim aşamasında yardımcı olabilecek modeller geliştirilmesi için önemli çalışmaların yapılmasına ortam hazırlamıştır. Robot seçimi problemi son yirmi yıl içinde aktif bir araştırma ortamı olarak belirmiştir.

Bu çalışmanın uygulama kısmının ilk bölümünde endüstriyel robot seçiminde kullanılacak bulanık doğrusal regresyon tabanlı bir karar modeli verilmektedir. Daha önce farklı yazarlar tarafından da kullanılmış olan yirmi yedi robot ve dört özellikten oluşan veri kümesi bu çalışmada da kullanılarak, bulanık doğrusal regresyon analizi uygulandığında elde edilen sonuçlar, literatürde aynı veri kümesine istatistiksel regresyon uygulanmasının sonuçlarıyla karşılaştırılarak, bulanık doğrusal regresyonun robot seçimi sürecinde alternatif bir karar destek yöntemi olarak kullanılması irdelenmiştir.

Robot özelliklerinden tekrarlayabilirlik, taşıma kapasitesi, hız ve maliyetin değerlendirme ölçütü olarak dikkate alındığı modelde, maliyet karar verici için bütçe kısıtını oluştururken, taşıma kapasitesi ve hız regresyon analizinin bağımsız değişkenlerini, tekrarlayabilirlik ise bağımlı değişkenini meydana getirmektedir. Gözlemlenen ile tahmin edilen değer arasındaki fark değerlerinin negatif çıktığı durumlar gerçek değer tahmin edilenden daha küçük olduğuna işaret etmektedir. Robotlar için tekrarlayabilirlik özelliğinin en küçüklenmesinin istenmesi düşünüldüğünde mutlak değeri en yüksek negatif farka sahip robot, performansı beklenenden en yüksek olan robot alternatifi konumuna gelmektedir.

Bu şekilde yapılan sıralama ile istatistiksel regresyon sonucu yapılan sıralama arasında pozitif ilişki bulunduğu Spearman sıra korelasyon testi uygulanarak gösterilmiş, alınan sonuçların birbiriyle örtüştüğü görülmüştür. Kullanılan değişik bulanık doğrusal regresyon modelleri ile sonucun iyileştirilmesine çalışılmış ve farklı yazarların sonuçlarıyla karşılaştırılmıştır.

Bu sayede endüstriyel robot seçimi problemine uygulanan bulanık doğrusal regresyonun, incelenen verilerin bulanık olduğu ya da teknoloji seçimi problemlerinde de karşımıza çıkan sistem parametreleri arasındaki ilişkilerin net olarak tanımlanamadığı durumlarda istatistiksel regresyona alternatif olarak kullanılabileceği kanısına varılmıştır.

Çalışmanın sonraki bölümünde bulanık doğrusal regresyonun uygulanmasına elverişli bir problem olan esnek imalat sistemi seçimi problemi ele alınmıştır. Modern imalat sistemleri, işletmelerin küreselleşmenin de etkisiyle sürekli değişen pazar şartlarına uyumunu kolaylaştırırken, kalitenin artırılması ve maliyetin düşürülmesinde de etkili olmaktadır. Modern imalat sistemlerine yapılan yatırımların büyük ölçekli olması ve seçim aşamasında değerlendirilmesi gereken çok sayıda birbiriyle çelişebilen özelliklerinin bulunması nedeniyle bu süreçte karar vericilerin kullanabileceği birçok model geliştirilmiştir.

On adet esnek imalat sisteminin beş özelliği için oluşturulan veri kümesi yardımıyla önerilen bulanık doğrusal regresyon tabanlı karar modelinin uygulanabilirliği incelenmiştir. Bulanık regresyon yöntemi, esnek imalat sistemi seçiminde de karşımıza çıkan sistem parametreleri arasındaki ilişkilerin net olarak tanımlanamadığı durumlarda istatistiksel regresyona alternatif olarak kullanılabilmektedir. Esnek imalat sistemi seçim sürecinde dikkate alınan değerlendirme ölçütleri, çıktı miktarı, üretim hacmi esnekliği, ortalama esnekliği, ara ürün stoğu ve çalışma dışı süre olup, çıktı miktarı regresyon analizinin bağımlı değişkeni olarak belirlenmiştir. Çıktı miktarını diğer özelliklerle ilişkilendiren regresyon denkleminin oluşturulmasından sonra bulanık regresyon parametreleri hesaplanmaktadır.

Birinci aşamada, oluşturulan veri kümesinde bağımlı değişkenin değerleri üçgensel simetrik bulanık sayılar halinde verilirken diğer değişkenler kesin sayılar halinde ifade edilmiştir. İkinci aşamada ise çıktı miktarının yanı sıra bağımsız bir değişken olan çalışma dışı süre de bulanık olarak verilmiş ve farklı regresyon modellerinin uygulanmasına imkan yaratılmıştır.

Gözlemlenen değer ile regresyon sonucu hesaplanan tahmini değerler arasındaki bulanık farkın büyüklüğü gerçek değer tahmin edilenden ne kadar fazla olduğunu, bir başka deyişle söz konusu alternatifin beklenen üzerindeki performansının derecesini göstermektedir. Hesaplanan fark değerinin en büyüklenmesi amacının yanında kullanım alanı ve maliyet kısıtlarının dikkate alınması, bulanık 0-1 programlama modeli yardımıyla gerçekleştirilmektedir. Bulanık 0-1 programlama modelinin yanı sıra aralık cinsinden ifade edilen amaç fonksiyonlarının optimizasyonuna dayanan farklı bir teknik de kullanılarak alınan sonuçlar karşılaştırılmıştır.

Sonuç olarak bulanık regresyon yönteminin, sistem parametreleri arasındaki kesin olmayan ilişkilerin bulunduğu ve değişkenlerin bulanık sayılar halinde ifade edildiği durumları modelleyebilme özelliği, teknoloji seçimi problemlerinin çözümünde istatistiksel regresyona alternatif, etkin bir araç olarak belirmesini sağlamaktadır.

Literatürde üzerine çok sayıda uygulama yapılmış iki konu, endüstriyel robot seçimi ve esnek imalat sistemi seçimi problemlerinin farklı bulanık doğrusal regresyon modelleri esas alınarak yapılandırılmış karar destek modelleri kullanılarak çözümlenebildiği gösterilmiş ve bu aşamada bulanık sayıların sıralanması, bulanık 0-1 programlama ve aralık cinsinden amaç fonksiyonların optimizasyonu gibi farklı tekniklerden yararlanılmıştır.

Bulanık regresyon analizi son yıllarda çok çeşitli alanlarda uygulamalarına rastlanan oldukça hızlı gelişen bir tekniktir. Belirsizliğin var olduğu her türlü alanda kullanımı başarılı sonuçlar vermekte ve istatistiksel regresyonun uygulanamadığı problemleri çözüme kavuşturmaktadır.

1. Introduction

La description mathématique des relations présentes entre les variables joue un rôle important pour la modélisation des systèmes. La connaissance d'une équation de cette forme facilite la compréhension des effets de changement des valeurs des variables et permet de prédire les états nouveaux du phénomène étudié. L'analyse de régression est une technique conçue pour effectuer ce type de travail cité, c'est à dire elle a pour but de modéliser les relations entre les variables. Elle calcule les coefficients de régression qui décrivent à quelle mesure les variables dites indépendantes sont liées à une variable qui dépend d'elles, nommée variable dépendante.

L'analyse de régression classique ne peut être utilisée que si les hypothèses sont vérifiées. Cette technique est valable pour les systèmes strictement définis où les variables prennent des valeurs exactes. L'analyse de régression classique reste insuffisante quand on est face à des relations imprécises. L'imprécision et l'incertitude créées en général par le raisonnement humain constituent la source de la théorie des sous-ensembles flous, une théorie mathématique qui est capable d'interpréter des informations imprécises et d'agir. La méthode de régression floue prend ses racines dans la théorie des sous-ensembles flous, et permet de modéliser les systèmes possédant des relations ou des données imprécises.

La deuxième étape de ce travail comporte les informations générales sur l'analyse de régression classique et une partie consacrée à la théorie des sous-ensembles flous afin de mieux comprendre la régression floue. Dans cette étape les modèles variés de régression linéaire floue sont présentés et une étude détaillée sur la littérature de régression floue est donnée.

Les étapes suivantes traitent l'utilisation des modèles de régression linéaire floue dans les problèmes de sélection de la technologie. Les problèmes de sélection de la technologie forment un grand ensemble de sous-problèmes rencontrés souvent dans

l'industrie. Le problème de sélection de robot industriel et le problème de sélection de système flexible de fabrication font partie de ce groupe. L'augmentation perpétuelle du nombre des technologies existantes rendent de plus en plus difficile le processus de prendre une décision définitive de la technologie qui va être utilisée. A cause de la variété et la complexité des technologies industrielles, il est assez compliqué de choisir l'alternative la plus conforme aux attentes d'une société.

Dans la troisième étape, afin de montrer la validité d'utilisation de régression linéaire floue pour la sélection de robot industriel, les résultats obtenus par la régression floue sont comparés avec ceux de la régression robuste. La sélection de robot industriel constitue la source des travaux publiés depuis un certain temps. Il existe différentes catégories de modèles pour trouver une solution au problème donné. Une méthodologie déjà pratiquée en utilisant la régression statistique classique, est appliquée à la régression linéaire floue et la corrélation des résultats est déterminée par le test de corrélation du rang de Spearman.

Les résultats étant satisfaisants pour la sélection de robot, dans la quatrième partie le problème de sélection de système flexible de fabrication est traité. L'adoption des systèmes flexibles de fabrication (SFF) qui font partie des technologies modernes, attire l'attention des chercheurs depuis une vingtaine d'années. Le gain de flexibilité, l'augmentation de la qualité, la diminution des coûts sont compris parmi les avantages d'intégrer un système flexible de fabrication. Par contre la sélection de ce type de technologie, rend obligatoire l'analyse de plusieurs caractéristiques qui augmente la complexité des modèles proposés.

Dans ce travail un modèle de décision basé sur la régression linéaire floue est proposée pour la sélection de système flexible de fabrication et la résolution est faite par différentes méthodes d'optimisation floue: la programmation booléenne floue et la programmation multiobjectif dans le cas des fonctions objectif sous forme d'intervalles. La dernière partie du travail aboutit à des conclusions et propose des idées nouvelles pour l'utilisation de la régression floue.

2. La régression floue

2.1. Généralités sur l'analyse de régression

L'analyse de régression est considérée comme une technique indispensable dans la plupart des domaines scientifiques [1]. Elle appartient à un grand ensemble constitué de méthodes de l'analyse des données. Les données sont utilisées soigneusement par les analystes afin de tirer des conclusions qui servent à enrichir les connaissances des disciplines variées de science.

L'analyse de régression est utilisée pour modéliser les relations entre les variables. La technique est basée sur le calcul des coefficients (les paramètres du modèle) qui décrivent à quelle mesure les variables dites indépendantes sont liées à une variable qui dépend d'elles, nommée variable dépendante. L'analyse de régression se distingue des autres types d'analyses statistiques avec son but qui est d'exprimer la variable dépendante comme une fonction des variables indépendantes [1].

Les deux cas majeurs où l'analyse de régression est fortement pratiquée sont les suivants [2]:

- la description des relations entre les variables par estimation des paramètres du modèle, et
- la prédiction des valeurs de la variable dépendante en utilisant les valeurs des variables indépendantes.

Par conséquent, la performance d'une méthode de régression est mesurée par le niveau de ces deux capacités: la capacité descriptive et la capacité prédictive.

L'analyse de régression statistique est une méthode ancienne, largement utilisée dans presque chaque matière. Dans les deux dernières décennies, la méthode de régression classique s'est évaluée donnant naissance à une nouvelle méthode: la régression floue.

Grâce à cela, la régression a ajouté les systèmes flous à son domaine d'application. Elle est maintenant utilisée pour décrire les événements imprécis en même temps que les événements précis.

La régression statistique et la régression floue ont été développées suivant différentes perspectives; nommées la théorie probabiliste et la théorie possibiliste respectivement. Comme leurs méthodes d'estimation des paramètres et par conséquent leurs résultats sont différents, l'approche utilisée est spécifique au problème et dépend de la situation étudiée [2].

2.2. La régression statistique

Le terme "régression statistique" employé tout au long de ce travail, signifie désormais la régression linéaire statistique. C'est à dire le modèle de régression de base comportant N variables indépendantes et dont la fonction de régression est linéaire, est étudié.

Le modèle de la régression statistique classique utilise une fonction linéaire afin d'exprimer la relation entre la variable dépendante y et les variables indépendantes x_1, \dots, x_N [3]:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_N x_{iN} + e_i \quad (2.1)$$

où β_j ($j = 0, 1, \dots, N$) représentent les paramètres (on peut les nommer aussi comme les coefficients de régression) et e_i ($i = 1, 2, \dots, M$) où M est le nombre de données pour la variable dépendante) représentent les termes d'erreur.

Le modèle suppose que les termes d'erreur qui sont des variables aléatoires suivent une loi de moyenne nulle et d'une variance constante. La détermination des coefficients β_j permet de calculer les valeurs estimées de y_i^* . Les valeurs d'écart e_i égales à la

différence entre y_i et y_i^* ($e_i = y_i - y_i^*$) sont généralement supposées dues aux erreurs de mesure [2] [4].

La résolution du modèle classique de la régression statistique se repose sur la méthode des moindres carrés. La méthode des moindres carrés détermine le vecteur optimum des coefficients de régression qui minimise la somme de e_i^2 [4]. Si on suppose que les valeurs de e_i suivent la loi normale, les valeurs de la variable dépendante peuvent être prédites utilisant les valeurs données des variables indépendantes et les intervalles de confiance peuvent être établis pour les paramètres de régression [2].

2.3. La régression floue

2.3.1. La théorie des sous-ensembles flous

La théorie des sous-ensembles flous est une théorie mathématique formée par Zadeh [5] afin de prendre en compte l'imprécision et les incertitudes qui règnent sur les événements naturels de la vie. Zadeh a écrit dans la préface d'un livre, qu'il a eu premièrement l'idée du concept des sous-ensembles flous au printemps de l'année 1964 quand il travaillait avec Bellman sur la classification et l'analyse des systèmes. Les méthodes classiques restaient insuffisantes dans le cas de modélisation des systèmes complexes et mal définis, ainsi ils cherchaient un nouveau concept qui pourrait contenir les domaines d'incertitude et d'imprécision. De ce point de départ Zadeh a décidé de généraliser la théorie des ensembles classiques et la notion de la fonction caractéristique. C'est ainsi que l'idée d'appartenance graduelle et ensuite les concepts de sous-ensemble flous et de degré d'appartenance naquirent [6]. La théorie des sous-ensemble flous a été annoncée pour la première fois dans l'article bouleversant de Zadeh [5].

2.3.1.1. Le concept d'incertitude

La pensée humaine est en général loine d'être précise. La vie a une nature complexe et floue. Les classes d'objet dans le monde réel n'ont pas des critères précis d'appartenance. La classe des objets larges ou la classe des gens en bonne santé ne

peuvent pas être définies strictement. Le même type d'ambiguïté se règne dans le cas d'un nombre comme 10 qui est en relation avec la classe des nombres réels supérieurs à 1 [5] [7].

Pour les mathématiques ordinaires, les sous-ensembles d'un ensemble sont définis clairement. Mais quand nous parlons des personnes en bonne santé dans l'ensemble des personnes, c'est assez difficile de décider si une personne appartient à ce sous-ensemble ou non. Même si une réponse est donnée, il existe une perte d'information car il n'y a pas de moyen de calculer le degré de santé de cette personne.

Si nous avons une proposition comme "John est jeune", il n'est pas toujours possible de dire si c'est vrai ou faux. Quand nous connaissons l'âge de John, disons x , la véracité de la proposition est plus juste. Mais cela dépend de notre point de vue sur la jeunesse. Si la proposition est "John a en dessous de 22 ans" et nous connaissons l'âge de John, nous pouvons donner une réponse. Les âges possibles se trouvent dans l'intervalle $[0, \infty)$. Soit A le sous-ensemble $\{x: x \in [0, \infty): x < 22\}$. A l'aide de cela nous pouvons déterminer si l'âge de John est dans A ou non. Mais la notion de "jeunesse" ne peut pas être définie comme un sous-ensemble de $[0, \infty)$. C'est juste ici que Zadeh introduit la notion de sous-ensemble flou [7].

Les classes imprécisément définies jouent un rôle important dans le raisonnement humain particulièrement dans les domaines de communication de l'information et la reconnaissance des formes [5]. La technologie qui avance rapidement a créé un rêve extraordinaire: des machines qui imitent le raisonnement humain et la théorie des ensembles flous a attiré l'attention de nombreux scientifiques. De nos jours les applications de cette théorie sont primordiales dans un domaine assez vaste de l'ingénierie et des sciences appliquées [7].

2.3.1.2. D'un ensemble classique à un ensemble flou

Avant de donner les définitions de base sur la théorie des sous-ensembles flous, il est important de faire une distinction entre les ensembles classiques et les ensembles flous.

Cette partie qui est une introduction aux propriétés mathématiques, afin de faciliter la compréhension générale concerne les différences de ces deux théories.

Un ensemble classique est un ensemble des objets satisfaisant des propriétés précises. C'est à dire les limites d'un ensemble classique sont définis strictement et un objet appartient complètement ou n'appartient pas du tout à cet ensemble. La théorie des sous-ensembles flous est une extension de la théorie des ensembles classiques. Cette théorie plus générale, englobe la théorie des ensembles classiques qui désormais considérée comme un cas particulier. Cette théorie est capable d'interpréter des informations imprécises et d'agir.

Tandis qu'un ensemble classique est un ensemble des objets satisfaisant des propriétés précises, un sous-ensemble flou est un ensemble des objets satisfaisant des propriétés imprécises. Un ensemble flou permet de graduer l'appartenance d'un élément à une classe c'est à dire la théorie des sous-ensembles flous permet de mesurer une gradation dans l'appartenance à un ensemble.

2.3.1.3. Définitions de base

Soit X est un espace de points avec un élément générique noté par x [5].

Un ensemble classique (non-flou) est défini comme un groupe d'éléments ou objets $x \in X$ qui peut être fini, comptable ou surcomptable. Chaque élément ne peut être qu'appartenir ou non à un ensemble A , $A \subseteq X$ [8].

Un sous-ensemble classique A d'un ensemble X est déterminé par une fonction caractéristique $\mu_A(x)$ qui est définie par [7]:

$$\mu_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases} \quad (2.2)$$

Un ensemble flou \tilde{A} dans X est caractérisé par une fonction d'appartenance $\mu_{\tilde{A}}(x)$ qui associe chaque point de X à un nombre réel de l'intervalle $[0, 1]$, avec une valeur de $\mu_{\tilde{A}}(x)$ qui représente le degré d'appartenance de x dans A [5].

Un sous-ensemble flou d'un ensemble X est une fonction $X \rightarrow [0,1]$. Ces fonctions dont les images sont contenues dans un ensemble de 2 éléments $\{0, 1\}$ correspondent aux sous-ensembles classiques, ainsi les sous-ensembles classiques sont les cas spéciaux des sous-ensembles flous.

La différence majeure qui existe entre les deux notions d'ensemble est la suivante: la fonction caractéristique est unique pour un ensemble classique, mais il y a une infinité de fonction d'appartenance pour un ensemble flou. Les opérations possibles sur les sous-ensembles flous sont identiques aux ensembles classiques mais les lois de non-contradiction et du tiers exclus restent en dehors de cette généralisation.

Les sous-ensembles flous sont nommés simplement comme *les ensembles flous* dans la littérature. Mais comme, la France reste en dehors de cet habitude, la théorie des ensembles flous y est encore nommée la théorie des sous-ensembles flous.

Le sous-ensemble flou permet de graduer l'appartenance d'un élément à un ensemble. Pour un concept flou donné, on peut considérer des fonctions d'appartenance diverses. Le choix de la fonction d'appartenance est subjectif et dépend du problème étudié [7].

Une α -coupe (alpha-coupe) d'un sous ensemble flou \tilde{A} est un sous-ensemble classique A_α extrait du sous ensemble flou \tilde{A} , définie en fonction d'un seuil $\alpha \in [0,1]$ fixé [8]:

$$A_\alpha = \{ x \in X \mid \mu_{\tilde{A}}(x) \geq \alpha \} \quad (2.3)$$

Un ensemble flou \tilde{A} est convexe si [8]

$$\mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geq \min\{\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)\}, x_1, x_2 \in X, \lambda \in [0,1] \quad (2.4)$$

Si $\sup_x \mu_{\tilde{A}}(x) = 1$, l'ensemble flou est dit normal [8].

Un nombre flou est un ensemble dont la fonction d'appartenance satisfait les propriétés de "normalité" (la fonction égale à 1 pour au moins un x) et convexité.

Un nombre flou triangulaire est un nombre flou spécial représenté par (a,b,c), a étant la borne inférieure, c étant la borne supérieure et b la valeur moyenne.

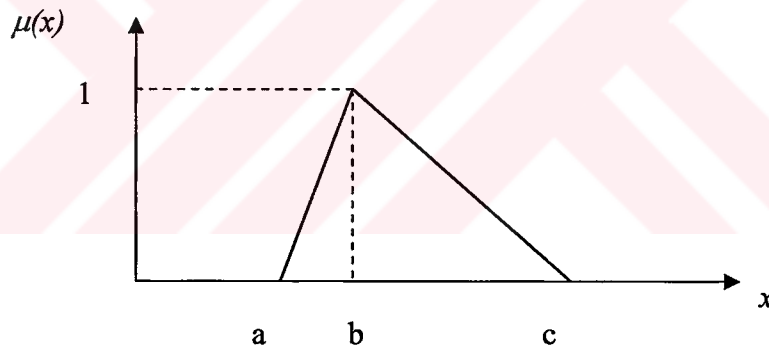


Figure 2.1: Un nombre triangulaire flou.

Si A_1 et A_2 sont des nombres flous triangulaires $A_1 \oplus A_2$, $A_1 \ominus A_2$, $-A_1$ sont des nombres flous triangulaires. Soient deux nombres flous triangulaires représentés par $A_1=(a_1,b_1,c_1)$, $A_2=(a_2,b_2,c_2)$. Les opérations sur ces nombres flous sont définies par les formules (2.5), (2.6), et (2.7).

Addition: \oplus représente l'addition des nombres flous,

$$A_1 \oplus A_2 = (a_1 + a_2, b_1 + b_2, c_1 + c_2) \quad (2.5)$$

Soustraction: \ominus représente la soustraction des nombres flous,

$$A_1 \ominus A_2 = (a_1 - c_2, b_1 - b_2, c_1 - a_2) \quad (2.6)$$

Multiplication: \otimes représente la multiplication des nombres flous et k est une constante scalaire:

$$k \otimes A_1 = \begin{cases} (ka_1, kb_1, kc_1), k > 0 \\ (kc_1, kb_1, ka_1), k < 0 \end{cases} \quad (2.7)$$

2.3.2. La régression floue : Généralités et survol de la littérature

L'analyse de régression floue où un système flou linéaire a été utilisé comme le modèle de régression, a été proposée premièrement par Tanaka, Uejima et Asai [4]. Par le fait que les fonctions d'appartenance soient considérées des distributions de possibilité, cet approche est nommée aussi l'analyse de régression possibiliste[9].

Les valeurs des variables indépendantes qui servent à calculer les coefficients et les valeurs estimées de la variable dépendante sont les données d'entrée (input) de l'analyse et les valeurs de la variable dépendante sont considérées comme les données de sortie (output). Il existe 3 cas pour analyser les données d'entrée et de sortie (input-output): *les données d'entrée et de sortie sont non-floues (crisp input-output data), *les données d'entrée sont non-floues et les données de sortie sont floues (crisp input-fuzzy output data), *les deux sont floues (fuzzy input-output data). Par conséquent, les modèles de régression floue sont assez variés et cela rend difficile le fait de grouper les différents types de l'analyse de régression floue en catégories bien définies. Néanmoins les techniques de régression sont classifiées en 2 catégories distinctes: les approches possibilistes et les approches de moindre-carrées [9].

Dans le cadre de ce travail afin d'observer l'évolution de la méthode de régression floue et d'analyser les domaines d'application, une étude détaillée sur la régression floue a été faite. A cette stade, la liste des publications sur la régression floue jusqu'en 1996 élaborée par Wen et Lee [10] était très utile. En utilisant ce travail et en faisant des recherches détaillées sur ce sujet, une liste contenant des articles publiés jusqu'au Mai 2005 sur la régression floue est élaborée. Le tableau 2.1 montre cette liste contenant les noms des auteurs, le type de modèle et l'étude de cas étudié dans l'article, le genre de l'article (théorie ou application) et son numéro de référence dans la liste générale des références.

Tableau 2.1 Survol de la littérature sur la régression floue

Auteur(s)	Modèle	Etude de Cas	Théorie ou Application	Réf.
Tanaka, Uejima, Asai (1982)	Problème de Min	Mécanisme de prix	Théorie et application	[4]
Yager (1982)	Variables linguistiques / Arithmétique floue	Exemples numériques	Théorie	[11]
Bandemer (1985)	Problème de Max (appartenance)	Exemples numériques	Théorie	[12]
Heshmaty, Kandel (1985)	Problème de Min	Prévision des ventes	Application	[13]
Gharpuray, Tanaka, Fan, Lai (1986)	Problème de Min	Hydrolyse de cellulose	Application	[14]
Jajuga (1986)	Moindres carrées	Exemples numériques	Théorie	[15]
Celmins (1987)	Moindres carrées	Problème de perforation	Théorie et application	[16]
Celmins (1987)	Moindres carrées	Prévision des ventes	Théorie et application	[17]
Tanaka (1987)	Problèmes de Min et de Max	Exemples numériques	Théorie	[18]
Chen (1988)	Problème de Max	Indice de réfraction	Théorie et application	[19]
Diamond (1988)	Moindres carrées	Exemples numériques	Théorie	[20]
Tanaka, Watada (1988)	Problèmes de Min et de Max	Taux de change de Yen	Théorie et application	[21]
Sakawa, Yano (1989)	Problème de Min	Prévision des ventes	Application	[22]

Tanaka, Hayashi, Watada (1989)	Conjonction des problèmes de Min et de Max	Exemples numériques	Théorie	[23]
Bardossy (1990)	Problèmes de Min modifiés	Mécanisme de prix	Théorie et application	[24]
Bardossy, Bogardi, Duckstein (1990)	Problèmes de Min modifiés	Hydrologie	Application	[25]
Hayashi, Tanaka (1990)	Problème de Min	Quantité de production / Prédiction de la température de l'eau	Application	[26]
Nather, Albrecht (1990)	BLUE (moindres carrées)	Exemples numériques	Théorie	[27]
Oh, Kim, Lee (1990)	Problème de Max (appartenance)	Prévision financière	Théorie et application	[28]
Savic, Pedrycz (1991)	Problème de Min VS Moindres carrées	Exemples numériques	Application	[29]
Tanaka, Ishibuchi (1991)	Conjonction des problèmes de Min et de Max	Exemples numériques	Théorie	[30]
Ishibuchi, Tanaka (1992)	Réseaux neuraux	Evaluation de la qualité	Théorie et application	[31]
Jozsef (1992)	Problème de Min	-	Théorie	[32]
Sakawa, Yano (1992)	Problème multiobjectif	Exemples numériques	Théorie	[33]
Wang, Ha (1992)	Modified Min problems	Exemples numériques	Théorie	[34]
Bardossy, Bogardi, Duckstein (1993)	Problème de Min	Dose-effet	Théorie et application	[35]
Ishibuchi, Tanaka, Okada (1993)	Réseaux neuraux	Exemples numériques	Théorie	[36]
Lu, Xu (1993)	Problème de Max (appartenance)	-	Théorie	[37]
Moskowitz, Kim (1993)	Problème de Min	Taux de change de Yen	Théorie et application	[38]
Chang, Lee (1994a)	Problème de Min	Exemples numériques	Théorie	[39]
Chang, Lee (1994b)	Moindres carrées	Exemples numériques	Théorie	[40]
Peters (1994)	Problème de Max (appartenance)	Indicateurs économiques	Théorie et application	[41]
Redden, Woodall (1994)	Problème de Min	Mécanisme de prix	Théorie et application	[3]

Tanaka, Ishibuchi, Yoskikawa(1995)	Problème de Min	Exemples numériques	Théorie	[42]
Chang, Lee, Konz (1996)	Problème de Min	Terminal de vidéo	Application	[43]
Chang, Lee (1996)	Moindres carrées pondérées	Exemples numériques	Théorie	[44]
Kim, Moskowitz, Koksalan (1996)	Problème de Min	Exemples numériques	Théorie	[2]
Redden, Woodall (1996)	Problème de Min	Exemples numériques	Théorie	[45]
Chang (1997)	Problème de Min	Prévision de saisonnalité	Application	[46]
Kim, Bishu (1998)	Problème de Min (appartenance)	Exemples numériques	Théorie	[47]
Chen (1999)	Régression floue quadratique	Exemples numériques	Théorie	[48]
Cheng, Lee (1999)	Régression floue avec réseau flou adapté	Exemples numériques	Théorie	[49]
Wen, Lee (1999)	Problème de Min	Développement d'une fonction de coût	Application	[10]
Yen, Ghoshray, Roig (1999)	Problème de Min	Exemples numériques	Théorie	[50]
Buckley, Feuring (2000)	Algorithmes évolutionnaires	Exemples numériques	Théorie	[51]
Dunyak, Wunsch (2000)	Régression floue utilisant les réseaux neuraux	Problème du contrôle de qualité industriel	Théorie	[52]
D'Urso, Gastaldi (2000)	Moindres carrées	Exemples numériques	Théorie	[53]
Kim, Moskowitz, Dhingra, Evans (2000)	Problème de Min	DFQ	Application	[54]
Özelkan, Duckstein (2000)	Régression multiobjectif floue	Exemples numériques	Théorie	[55]
Wang, Tsaur (2000)	Moindres carrées (modifié)	Decision problem	Théorie et application	[56]
Wang, Tsaur (2000)	Problème de Min	Exemples numériques	Théorie	[57]
Chang (2001)	Moindres carrées (hybride)	Exemples numériques	Théorie	[58]
Chang, Ayyub (2001)	Problème de Min Moindres carrées Régression à intervalles	Exemples numériques	Application	[59]
Chen (2001)	Problème de Min modifié	Exemples numériques	Théorie	[60]

Cheng, Lee (2001)	Régression floue + radial basis network	Exemples numériques	Théorie	[61]
Ishibuchi, Nii (2001)	Régression floue utilisant les réseaux neuraux	Exemples numériques	Théorie	[62]
Lee, Chen (2001)	Problème de Min (fuzzy input)	Prévision de "Manpower"	Théorie et application	[63]
Özelkan, Duckstein (2001)	Régression multiobjectif floue Moindres carrées	Modèles de pluviosité et ruissellement	Théorie et application	[64]
Tseng, Tzeng, Yu, Yuan (2001)	Problème de Min	Modèle d'ARIMA pour les bourses étrangères	Théorie et application	[65]
Yu, Tzeng, Li (2001)	Régression floue par morceaux	Exemples numériques	Théorie	[66]
Chang (2002)	Régression floue pondérée multi-étapes	Modélisation du cycle de vie d'un produit	Théorie et application	[67]
Cron, Dubuisson (2002)	Régression à intervalles	Reconnaissance des formes	Théorie et application	[68]
Kahraman (2002)	Problème de Min	Prévision des ventes	Application	[69]
Kao, Chyu (2002)	Régression floue à 2 étapes	Exemples numériques	Théorie	[70]
Tran, Duckstein (2002)	Régression multiobjectif floue	Exemples numériques	Théorie	[71]
Tran, Ridgley, Duckstein, Sutherland (2002)	Régression multiobjectif floue	Erosion du sol	Application	[72]
Tseng, Tzeng (2002)	Problème de Min	Modèle d'ARIMA pour la prévision	Théorie et application	[73]
Wu, Tseng (2002)	Moindres carrées	Analyse de cycle de business	Théorie et application	[74]
Wünsche, Nather (2002)	Moindres carrées	-	Théorie	[75]
Yang, Lin (2002)	Moindres carrées	Exemples numériques	Théorie	[76]
D'Urso (2003)	Moindres carrées	Exemples numériques	Théorie	[77]
Hong, Hwang (2003)	La machine de vecteur support	Exemples numériques	Théorie	[78]
Ip, Kwong, Wong (2003)	Problème de Min	Modélisation d'un processus	Application	[79]
Kao, Chyu (2003)	Moindres carrées	Exemples numériques	Théorie	[80]

Soliman, Alammari, El-Hawary (2003)	Problème de Min	Evaluation de la fréquence et de l'harmonie	Application	[81]
Tsaur (2003)	Problème de Min	Modèle de prévision	Application	[82]
Yang, Liu (2003)	Moindres carrées	Exemples numériques	Théorie	[83]
Al-Kandari, Soliman, El-Hawary (2004)	Problème de Min	Prévision de la quantité de charge	Application	[84]
Chaves, Tsukatani, Kojiri (2004)	Régression multiobjectif flou	Opération sur un réservoir	Application	[85]
Hong, Hwang (2004)	Technique de régularisation	Exemples numériques	Théorie	[86]
Hong, Hwang, Ahn (2004)	Estimation robuste	Exemples numériques	Théorie	[87]
Li, Liao, Yu (2004)	Méthode Tabu	Exemples numériques	Théorie et application	[88]
Nasrabadi, Nasrabadi (2004)	Programmation mathématique	Exemples numériques	Théorie et application	[89]
Sanchez, Gomez (2004)	Régression multiobjectif floue	Estimation des taux de change	Théorie et application	[90]
Uddameri, Kuchanur (2004)	Moindres carrées	Prédiction de $\log K_{oc}$	Application	[91]
Chang (2005)	Problème de Min	Stratégie de renouvellement	Théorie et application	[92]
Hojati, Bector, Smimou (2005)	Problème de Min	Exemples numériques	Théorie	[93]
Nasrabadi, Nasrabadi, Nasrabady (In Press)	Régression multiobjectif floue	Exemples numériques	Théorie et application	[94]
Tseng, Lin (2005)	Programmation quadratique	Prévision de banqueroute	Théorie et application	[95]

Dans ce travail nous nous intéressons seulement à des approches possibilistes qui sont basées sur les modèles possibilistes. Les modèles étudiés sont des modèles linéaires de régression floue qui sont très répandus dans la littérature de régression floue.

2.3.3. La régression linéaire floue : Le modèle de Tanaka

Le modèle de la régression linéaire classique utilise une fonction linéaire afin d'exprimer la relation entre la variable dépendante y et les variables indépendantes x_1, \dots, x_N [3]:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_N x_{iN} + e_i \quad (2.8)$$

où β_j ($j = 0, 1, \dots, N$) représentent les coefficients de régression et e_i ($i = 1, \dots, M$ où M est le nombre de données pour la variable dépendante) représentent les termes d'écart entre les valeurs observées y_i de la variable dépendante et celles des valeurs estimées y_i^* . La régression linéaire floue utilise aussi la même type de fonction avec une différence primordiale qui constitue la base de la nouvelle méthode: les coefficients de l'analyse de régression floue sont sous forme des nombres flous. Par conséquent les valeurs estimées de y_i^* calculées par la suite d'obtention des paramètres \tilde{A}_j sont aussi des nombres flous.

Tanaka et ses collègues [4] dans leur article fondateur de cette méthode, décrivent une fonction de régression linéaire floue comme (2.9).

$$y_i^* = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \tilde{A}_2 x_{i2} + \dots + \tilde{A}_N x_{iN} \quad (2.9)$$

où x_{ij} désigne ($i = 1, 2, \dots, M$ et $j = 1, 2, \dots, N$) la valeur de la variable indépendante pour la i -ème observation, et y_i^* désigne la valeur de la variable dépendante de l'observation i calculée à partir de l'équation (2.9).

Le choix de la fonction d'appartenance des paramètres flous dépend du décideur. Afin de simplifier la formulation et alléger les notations, les paramètres de ce modèle sont des nombres flous symétriques $\tilde{A}_j = (\alpha_j, c_j)$, qui peuvent être définies par:

$$\mu_{\tilde{A}_j}(a_j) = L((a_j - \alpha_j) / c_j) \quad (2.10)$$

où α_j représente la moyenne, c_j représente la dispersion des paramètres flous, et $\mu_{\tilde{A}_j}(a_j)$ représente la fonction d'appartenance de a_j au nombre flou \tilde{A}_j . La fonction de référence satisfait les propriétés suivantes: * $L(x) = L(-x)$, * $L(0)=1$, * $L(x)$ est strictement décroissante sur $[0, +\infty)$ [21].

Le but de régression linéaire flou est de déterminer les paramètres flous estimés $\tilde{A}_j = (\alpha_j, c_j)$ à condition que le degré d'appartenance de y_i (la i -ième valeur observée de la variable dépendante) de sa valeur floue estimée $y_i^* = \tilde{A}X_i$ soit au moins H , qui est choisi par le décideur [4].

Le modèle de Tanaka étudié dans cette partie est destiné à des valeurs de variables indépendantes et de variables dépendantes non-floues. Notons que les valeurs estimées

y_i^* ont la forme de $(\sum_{j=0}^N \alpha_j x_{ij}, \sum_{j=0}^N c_j |x_{ij}|)$. La condition citée ci-dessus peut être

représentée par deux contraintes d'inégalités:

$$\sum_{j=0}^N \alpha_j x_{ij} + |L^{-1}(H)| \sum_{j=0}^N c_j |x_{ij}| \geq y_i, \quad i = 1, 2, \dots, M \quad (2.11)$$

$$\sum_{j=0}^N \alpha_j x_{ij} - |L^{-1}(H)| \sum_{j=0}^N c_j |x_{ij}| \leq y_i, \quad i = 1, 2, \dots, M \quad (2.12)$$

Le modèle de Tanaka suppose que $L(x) = \max(0, 1 - |x|)$. Ainsi les coefficients sont des nombres triangulaires symétriques représentés par (2.13) et aussi géométriquement dans la figure 2.1 et $|L^{-1}(H)| = 1 - H$.

$$\mu_{\tilde{A}_j}(a_j) = \begin{cases} 1 - \frac{|\alpha_j - a_j|}{c_j}, & \alpha_j - c_j \leq a_j \leq \alpha_j + c_j \\ 0, & \text{sinon} \end{cases} \quad (2.13)$$

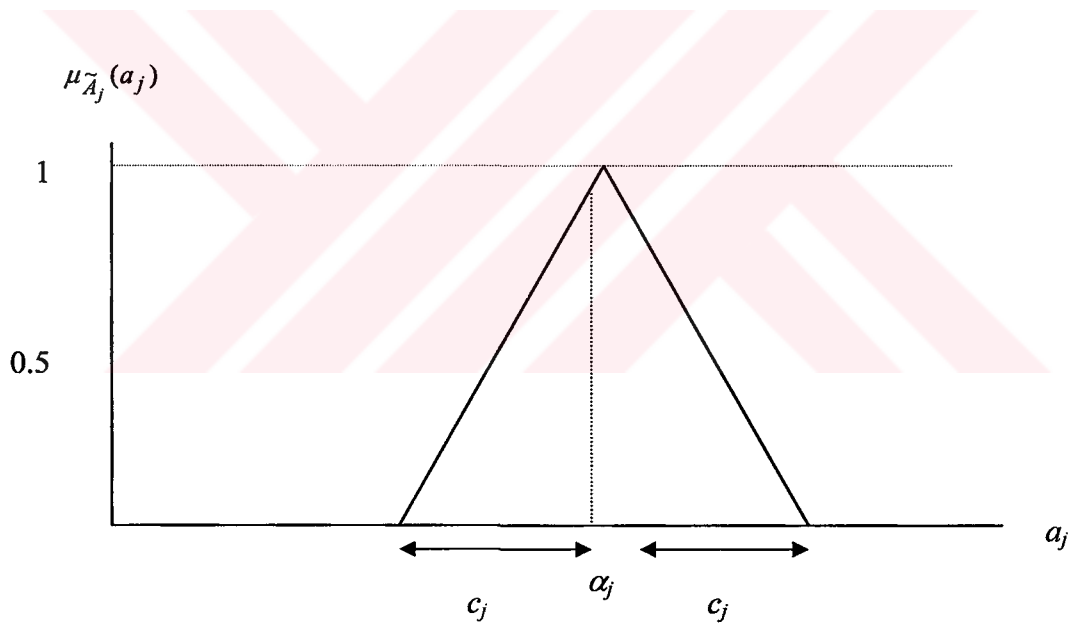


Figure 2.2: Un coefficient flou triangulaire symétrique.

Le but de la méthode de Tanaka est de déterminer les coefficients flous afin que la valeur estimée y_i^* de la variable dépendante soit la moins floue possible, i.e. ait la plus petite valeur de dispersion c_j en satisfaisant un degré cible de confiance H . Le terme H est considéré comme une mesure d'adéquation ("goodness-of-fit" en anglais) ou une mesure de compatibilité entre l'ensemble de données et le modèle [59]. Une valeur de H est égal à 0 indique que le modèle établi est complètement compatible avec

l'ensemble de données, au contraire une valeur de H égale à 1 indique l'incompatibilité totale .

L'imprécision totale de l'ensemble de données étudié est calculée par:

$$\sum_{j=0}^N (c_j \sum_{i=1}^M |x_{ij}|) \quad (2.14)$$

Le modèle qui minimise l'imprécision totale sous les contraintes (2.11) et (2.12) peut être formulé comme un problème de programmation linéaire indiqué ci-dessous et ainsi il peut être résolu facilement.

$$\text{Min } \sum_{j=0}^N (c_j \sum_{i=1}^M |x_{ij}|) \quad (2.15)$$

$$\sum_{j=0}^N \alpha_j x_{ij} + (1-H) \sum_{j=0}^N c_j |x_{ij}| \geq y_i, \quad i = 1, 2, \dots, M$$

$$\sum_{j=0}^N \alpha_j x_{ij} - (1-H) \sum_{j=0}^N c_j |x_{ij}| \leq y_i, \quad i = 1, 2, \dots, M$$

$$c_j \geq 0, \quad j = 1, 2, \dots, N$$

$$x_{i0} := 1, H \in [0,1]$$

La détermination des paramètres flous par l'algorithme de la régression linéaire flou dépend de la valeur H qui est choisie subjectivement. Pour donner une interprétation mathématique de H , on peut dire que chaque observation y_i est contenu dans l'intervalle de support de la valeur estimée y_i^* correspondante avec un degré d'appartenance $\geq H$.

Il faut aussi signaler que très rarement, la valeur H est définie comme la plus petite valeur de h_i , h_i étant les valeurs choisies pour chaque observation [38].

Une importante caractéristique de la régression linéaire floue est que les valeurs moyennes des coefficients n'apparaissent pas dans la fonction objectif, seule l'augmentation des dispersions pénalise le critère choisi. Ceci peut expliquer le fait qu'un point de donnée input ajouté, qui satisfait la valeur d'appartenance du modèle déterminé sans ce point particulier, à l'ensemble de données ne résulte aucun changement dans le modèle. C'est à dire le modèle est insensible aux nouveaux points qui satisfont les contraintes du modèle déjà établi. La valeur de H devient importante à cause de cette caractéristique citée, il faut noter que les valeurs très petites de H peut rendre le modèle très robuste mais en même temps le modèle ne peut pas être assez spécifique [29]. La figure 2.3 représente l'algorithme de la régression linéaire floue.

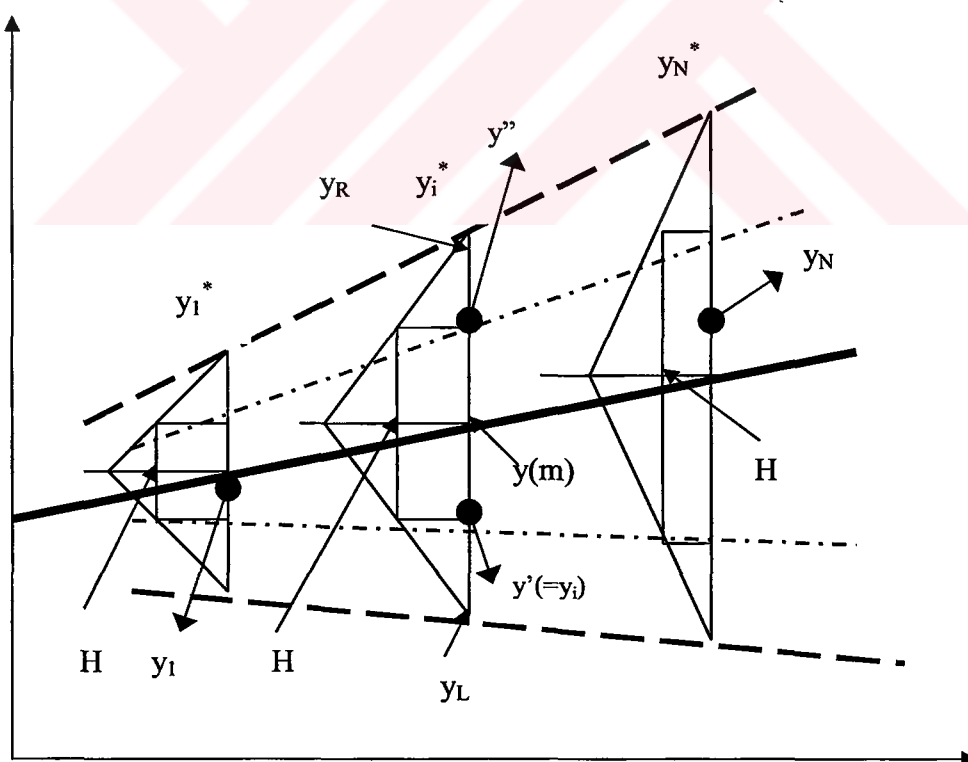


Figure 2.3: La représentation de l'algorithme de la régression linéaire floue [2].

Les valeurs moyennes des paramètres flous restent inchangées pour différents degrés de confiance tandis que le modèle, il devient plus flou avec un degré de confiance c'est à dire une valeur H plus grande [59]. Tanaka et Watada [21] proposent une méthode pour la détermination de la valeur H selon la suffisance de l'ensemble de données étudié. La valeur H est égale à 0 lorsque l'ensemble de données est suffisamment large et elle augmente au fur et à mesure que l'ensemble de données devient petit quand on le compare avec la taille idéale. Le critère de sélection de la valeur H est généralement subjectif et difficile à justifier dans des situations réelles. Il existe deux approches (une approche analytique et une approche de recherche) développées par Moskowitz et Kim [38] pour affecter une valeur H plus rationnellement. Il faut noter que les travaux publiés sur ce sujet n'utilisent quasiment pas ces approches qui demandent un effort supplémentaire.

2.4. Les propriétés de la régression linéaire floue

2.4.1. La régression linéaire floue versus la régression linéaire statistique

La régression statistique peut être utilisée dans un environnement non-flou où les relations entre les variables sont parfaitement définies [2]. Le modèle de la régression statistique est applicable seulement si les données sont réparties selon une loi statistique et aussi si les données et les relations ne sont pas floues [3]. Si le phénomène étudié ne varie pas suivant des règles stochastiques mais s'il est aussi imprécis, il est plus naturel d'établir une équation floue exprimant les relations d'un ensemble de données qui peut être flou ou non-flou. C'est à dire un phénomène flou doit être modélé par une équation floue [9]. La méthode de la régression linéaire floue proposée par Tanaka, Uejima et Asai [4] a pour but de modéliser les phénomènes imprécis tandis que la régression statistique, elle a été conçue pour les systèmes bien définis.

Le modèle suppose que les termes d'erreur e_i qui sont des variables aléatoires, suivent une loi de moyenne nulle et d'une variance constante. Les valeurs d'écart e_i égales à la différence entre la valeurs observées et les valeurs estimées sont généralement

supposées dues aux erreurs de mesure [2] [4]. On contraire, les valeurs d'écart e_i sont dues à la nature imprécise du système dans l'analyse de régression floue.

La régression classique fait des hypothèses strictes sur les propriétés statistiques du modèle qui ne sont pas facilement être justifiées si l'ensemble de données n'est pas suffisamment large. La violation des ces règles peut affecter la validité et la performance de la régression statistique. La régression linéaire floue est plus effective que la régression linéaire statistique quand les hypothèses ne peuvent pas être employées proprement, si par exemple les jugements humains ou des processus ambigus sont présents [2].

L'analyse de régression est capable de: 1) décrire les relations entre les variables par estimation des paramètres du modèle, et 2) prédire les valeurs de la variable dépendante en utilisant les valeurs des variables indépendantes. La régression statistique linéaire est généralement supérieure à la régression linéaire floue pour le pouvoir prédictif mais pour la capacité descriptive la réponse n'est pas facile à trouver. Les facteurs associés à l'ensemble de données, le choix de la méthode sont importants dans ce cas-là. La régression floue devient une technique très effective quand l'ensemble de données est insuffisant pour supporter l'analyse statistique ou quand il y a des relations incertaines entre les variables [2].

Comme la méthode d'estimation des paramètres est différente pour les deux approches considérées, le décideur doit décider quelle approche faut-il utiliser dans une situation spécifique [2]. En général, les problèmes de la régression statistique doivent être ouvert à l'expérience, c'est à dire on peut faire suffisamment d'expériences afin d'arriver à une relation statistique valide [29]. La régression floue est appropriée aux situations où les valeurs d'une partie ou de la totalité des variables sont floues [2].

2.4.2. Vue générale sur les modèles de régression linéaire floue

Dans cette partie nous allons essayer de donner les propriétés de quelques méthodes de régression linéaire floue y compris particulièrement la méthode de Tanaka. Un

formulaire contenant les modèles importants de la littérature dont certains seront utilisés dans les parties suivantes du travail est présenté. Les modèles sont classés selon la nature des données. Le premier groupe contient les modèles établis pour les variables dépendantes et indépendantes non-floues (crisp input - crisp output), le deuxième groupe renferme les modèles pour la variable dépendante floue et les variables indépendantes non-floues (crisp input - fuzzy output) et le dernier groupe possède des modèles pour les variables dépendantes et indépendantes floues (fuzzy input - fuzzy output).

- variables dépendantes et indépendantes non-floues

Le premier modèle proposé par Tanaka, Uejima et Asai [4] a été conçu avec des variables dépendantes floues. Le modèle, qui serait ultérieurement développé suivant deux méthodes, a été basé sur le critère de minimisation de la somme des dispersions. La version pour les données non-floues de ce premier type peut être aisément obtenue en supprimant les dispersions des variables dépendantes. Le modèle de Tanaka ainsi formulé est encore couramment utilisé grâce à sa facilité d'application.

Tanaka 1a [4](variables dépendantes et indépendantes non-floues)

$$\text{Minimiser } c_0 + c_1 + \dots + c_N \quad (2.16)$$

Sous les contraintes

$$\sum_{j=0}^N \alpha_j x_{ij} + \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \geq y_i, \quad i = 1, 2, \dots, M$$

$$\sum_{j=0}^N \alpha_j x_{ij} - \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \leq y_i, \quad i = 1, 2, \dots, M$$

$$c_j \geq 0, \quad j = 0, 1, \dots, N$$

Tanaka et ses collègues ont essayé d'améliorer les défauts de leur modèle au fur et à mesure que de nombreux critiques apparaissaient sur ce sujet. L'une des critiques les plus importantes était que la solution était dépendante de l'échelle des valeurs des variables indépendantes. Les coefficients du modèle ainsi que leurs valeurs de dispersion dépendent de l'unité de mesure des variables indépendantes [3] [93]. Afin de rectifier ce problème, le critère de minimisation de la somme des dispersions a été changé par le critère de minimisation de la somme pondérée des dispersions.

Tanaka 2a [21] (variables dépendantes et indépendantes non-floues)

$$\text{Minimiser } \sum_{j=0}^N (c_j \sum_{i=1}^M |x_{ij}|) \quad (2.17)$$

Sous les contraintes

$$\sum_{j=0}^N \alpha_j x_{ij} + \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \geq y_i, \quad i = 1, 2, \dots, M$$

$$\sum_{j=0}^N \alpha_j x_{ij} - \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \leq y_i, \quad i = 1, 2, \dots, M$$

$$c_j \geq 0, \quad j = 0, 1, \dots, N$$

Le modèle de Tanaka a été aussi fortement critiqué à cause de sa faiblesse envers les données extrêmes ou "outliers". Peters [41] avec son modèle dont l'intervalle estimé est influencé par toutes les données existantes est une viable alternative à l'approche de Tanaka. Le modèle de Peters est robuste en présence des données extrêmes.

Peters [41] (variables dépendantes et indépendantes non-floues)

$$\text{Maximiser } \bar{\lambda} = 1/M \sum_{i=1}^M \lambda_i \quad (2.18)$$

Sous les contraintes

$$(1 - \bar{\lambda})p_0 - \sum_{i=1}^M \sum_{j=0}^N c_j |x_{ij}| \geq -d_0$$

$$(1 - \lambda_i)p_i + \sum_{j=0}^N \alpha_j x_{ij} + \sum_{j=0}^N c_j |x_{ij}| \geq y_i, \quad i=1, 2, \dots, M$$

$$(1 - \lambda_i)p_i - \sum_{j=0}^N \alpha_j x_{ij} + \sum_{j=0}^N c_j |x_{ij}| \geq -y_i, \quad i=1, 2, \dots, M$$

$$\lambda_i \leq 1, \quad |L^{-1}(H)| := 1, \quad c_j \geq 0, \quad i=1, 2, \dots, M, \quad j=0, 1, \dots, N$$

Ici, d_0 représente la valeur désirée de la fonction objectif. Comme nous préférons obtenir un modèle non-flou (minimisation de l'incertitude) d_0 peut-être choisi égale à 0. Les valeurs de p_0 et de p_i doivent être choisies selon le contexte général du problème (une valeur élevée de p_0 et des faibles valeurs de p_i signifient un faible besoin de minimisation des dispersions). Les détails du modèle de Peters vont être données ultérieurement dans un contexte de comparaison avec le modèle de Tanaka.

• variables dépendantes floues, variables indépendantes non-floues

Le modèle pour les variables dépendantes floues de Tanaka a été développé suivant deux méthodes et s'est transformé en deux problèmes nommés Max et Min. Pour les variables dépendantes floues \bar{y}_i désigne la valeur moyenne et e_i désigne la valeur de dispersion de y_i , $y_i = (\bar{y}_i, e_i)$.

Tanaka 1b [4] (variables dépendantes floues, variables indépendantes non-floues)

Minimiser $c_0 + c_1 + \dots + c_N$ (2.19)

Sous les contraintes

$$\sum_{j=0}^N \alpha_j x_{ij} + \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \geq \bar{y}_i + (1-H)e_i, \quad i=1, 2, \dots, M$$

$$\sum_{j=0}^N \alpha_j x_{ij} - \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \leq \bar{y}_i - (1-H)e_i, \quad i=1, 2, \dots, M$$

$$c_j \geq 0, \quad j = 0, 1, \dots, N$$

Tanaka 2b [21] (variables dépendantes floues, variables indépendantes non-floues)

Problème de Max

$$\text{Maximiser } \sum_{j=0}^N (c_j \sum_{i=1}^M |x_{ij}|) \quad (2.20)$$

Sous les contraintes

$$\sum_{j=0}^N \alpha_j x_{ij} + \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \leq \bar{y}_i + (1-H)e_i, \quad i=1, 2, \dots, M$$

$$\sum_{j=0}^N \alpha_j x_{ij} - \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \geq \bar{y}_i - (1-H)e_i, \quad i=1, 2, \dots, M$$

$$c_j \geq 0, \quad j = 0, 1, \dots, N$$

Tanaka 2c [21] (variables dépendantes floues, variables indépendantes non-floues)

Problème de Min

$$\text{Minimiser } \sum_{j=0}^N (c_j \sum_{i=1}^M |x_{ij}|) \quad (2.21)$$

Sous les contraintes

$$\sum_{j=0}^N \alpha_j x_{ij} + \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \geq \bar{y}_i + (1-H)e_i, \quad i=1,2,\dots,M$$

$$\sum_{j=0}^N \alpha_j x_{ij} - \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \leq \bar{y}_i - (1-H)e_i, \quad i=1,2,\dots,M$$

$$c_j \geq 0, \quad j=0,1,\dots,N$$

Hojati, Bector et Smimou [93] proposent des modèles de régression linéaire floue qui sont simples à résoudre et qui ont pour but de corriger les défauts du modèle de Tanaka.

Les deux cas suivants ont été examinés par Hojati, Bector et Smimou [93]:

Cas 1: variables indépendantes non-floues, variable dépendantes floues

Cas 2: variables indépendantes et dépendantes floues

Leurs modèles donnent des solutions meilleures pour le premier cas, quand on les compare avec celles de Tanaka et pour le second cas ils donnent une nouvelle approche à un problème qui n'a pas été traité suffisamment dans la littérature. d_{iu} et d_{il} étant les variables de déviation supérieures et inférieures, les coefficients de régression sont déterminés de façon où la déviation totale entre les points supérieurs des des intervalles H -coupe estimés et observés $|d_{iu}^+ - d_{iu}^-|$ et entre les points inférieurs des des intervalles H -coupe estimés et observés $|d_{iL}^+ - d_{iL}^-|$ est minimisé.

HBS 1 [93] (variables indépendantes non-floues, variables dépendantes floues)

$$\text{Minimiser } \sum_{i=1}^M (d^+_{iU} + d^-_{iU} + d^+_{iL} + d^-_{iL}) \quad (2.22)$$

Sous les contraintes

$$\sum_{j=0}^N (\alpha_j + (1-H)c_j)x_{ij} + d^+_{iU} - d^-_{iU} = \bar{y}_i + (1-H)e_i, \quad i=1, 2, \dots, M$$

$$\sum_{j=0}^N (\alpha_j - (1-H)c_j)x_{ij} + d^+_{iL} - d^-_{iL} = \bar{y}_i - (1-H)e_i, \quad i=1, 2, \dots, M$$

α_j : sans signe imposé, $c_j \geq 0$, $j = 0, 1, \dots, N$, $d^+_{iL}, d^-_{iL}, d^+_{iU}, d^-_{iU} \geq 0$, $i=1, 2, \dots, M$

• variables dépendantes et indépendantes floues

Les variables indépendantes aussi sont sous forme des nombres flous, $x_{ij} = (\bar{x}_{ij}, f_{ij})$.

Le modèle de Hojati, Bector et Smimou [93] pour le second cas est donné pour une seule variable indépendante afin d'alléger l'expression.

Sakawa et Yano [33] proposent un modèle multiobjectif afin de résoudre ce type de régression linéaire floue. Ils considèrent le conflit entre le degré de confiance et l'ampleur des valeurs de dispersion des coefficients. Les valeurs des variables dépendantes et indépendantes sont sous forme des nombres flous comme ceux-ci: $x_{ij} = (\bar{x}_{ij}, d_{ij})$, $y_i = (\bar{y}_i, e_i)$. L et R désignent la borne inférieure et la borne supérieure d'un nombre flou donné sous forme d'intervalle $[L, R]$.

Sakawa et Yano [33] divisent les paramètres en 3 catégories J_1 , J_2 et J_3 satisfaisant les propriétés données au modèle. Cette hypothèse n'est pas très restrictive car dans les situations réelles le signe des paramètres est en général connu en avance.

HBS 2 [93] (variables dépendantes et indépendantes floues)

$$\text{Minimiser } \sum_{i=1}^M (d^+_{iU} + d^-_{iU} + d^+_{iL} + d^-_{iL} + d^+_{iR} + d^-_{iR} + d^+_{iL} + d^-_{iL}) \quad (2.23)$$

Sous les contraintes

$$\sum_{j=0}^1 (\alpha_j + (1-H)c_j)(\bar{x}_{ij} - (1-H)f_{ij}) + d^+_{iU} - d^-_{iU} = \bar{y}_i + (1-H)e_i, \quad i=1, 2, \dots, M$$

$$\sum_{j=0}^1 (\alpha_j + (1-H)c_j)(\bar{x}_{ij} + (1-H)f_{ij}) + d^+_{iR} - d^-_{iR} = \bar{y}_i + (1-H)e_i, \quad i=1, 2, \dots, M$$

$$\sum_{j=0}^1 (\alpha_j + (1-H)c_j)(\bar{x}_{ij} - (1-H)f_{ij}) + d^+_{iL} - d^-_{iL} = \bar{y}_i - (1-H)e_i, \quad i=1, 2, \dots, M$$

$$\sum_{j=0}^1 (\alpha_j + (1-H)c_j)(\bar{x}_{ij} + (1-H)f_{ij}) + d^+_{iR} - d^-_{iR} = \bar{y}_i - (1-H)e_i, \quad i=1, 2, \dots, M$$

α_j : sans signe imposé, $c_j \geq 0$, $j=0,1$ $x_{ij} = (\bar{x}_{ij}, f_{ij})$

$$d^+_{iL}, d^-_{iL}, d^+_{iU}, d^-_{iU}, d^+_{iR}, d^-_{iR}, d^+_{iL}, d^-_{iL} \geq 0, \quad i=1, 2, \dots, M$$

Sakawa et Yano [33] (variables dépendantes et indépendantes floues)

$$\begin{aligned} Y_{i\alpha}^L &= \sum_{j \in J_1} (\bar{a}_j - L^{-1}(\alpha)c_j)(\bar{x}_{ij} - L^{-1}(\alpha)d_{ij}) \\ &+ \sum_{j \in J_2 \cup J_3} (\bar{a}_j - L^{-1}(\alpha)c_j)(\bar{x}_{ij} + L^{-1}(\alpha)d_{ij}) \end{aligned} \quad (2.24)$$

$$\begin{aligned}
Y_{i\alpha}^R &= \sum_{j \in J_3} (\bar{a}_j + L^{-1}(\alpha)c_j)(\bar{x}_{ij} - L^{-1}(\alpha)d_{ij}) \\
&+ \sum_{i \in J_1 \cup J_2} (\bar{a}_j + L^{-1}(\alpha)c_j)(\bar{x}_{ij} + L^{-1}(\alpha)d_{ij})
\end{aligned} \tag{2.25}$$

$$\text{Min } Z = \sum_{i=1}^k (Y_{i0}^R - Y_{i0}^L) \tag{2.26}$$

$$\bar{y}_i + L^{-1}(\varepsilon)e_i \geq Y_{i\varepsilon}^L, \quad i = 1, 2, \dots, M$$

$$\bar{y}_i - L^{-1}(\varepsilon)e_i \leq Y_{i\varepsilon}^R, \quad i = 1, 2, \dots, M$$

$$\bar{a}_j - L^{-1}(\varepsilon)c_j \geq 0, \quad j \in J_1$$

$$\bar{a}_j - L^{-1}(\varepsilon)c_j \leq 0, \quad \bar{a}_j + L^{-1}(\varepsilon)c_j \geq 0, \quad j \in J_2$$

$$\bar{a}_j + L^{-1}(\varepsilon)c_j \leq 0, \quad j \in J_3$$

$$c_j \geq 0, \quad j = 0, 1, \dots, N$$

3. Le problème de sélection de robot industriel

Le mot “robot” est utilisé pour la première fois en 1920 dans un oeuvre intitulé *Rossum’s Universal Robots* par Karel Capek, l’écrivain de science-fiction tchèque. Ce mot courant de la vie industrielle, synonyme du mot “travailleur” en tchèque, exerce toutes les responsabilités assumées par sa signification littéraire. En 1942, Asimov, écrivain d’une série d’histoires concernant les robots qui serait publiée dans les années suivantes, a développé “les trois lois de robotique” suivantes [96]:

- Un robot ne doit blesser personne ni, par manque d’activité, permettre qu’une personne ne fasse mal.
- Un robot doit toujours obéir à des ordres données par des gens sauf dans des circonstances telles que ces ordres soient en conflit avec la loi précédente.
- Un robot doit protéger sa propre existence sauf dans des circonstances telles qu’ils soient en conflit avec les deux lois précédentes.

L’apparition des robots sur la scène industrielle, a occasionné la naissance d’une nouvelle période de production. Alors que la robotique industrielle est née en 1963, il a fallu attendre les années 80 pour les travaux de normalisation de la mesure de performances des robots existants. Pendant les deux dernières décennies, d’une part, le nombre de robot industriel a augmenté brusquement et d’autre part, l’élévation du coût de fabrication d’un robot industriel s’est resté dans les intervalles acceptables. Grâce à ces deux événements observés, les robots sont utilisés de manière de plus en plus répandue dans la vie du travail.

Au départ, le rôle des robots industriels n’était limité qu’à l’exécution des tâches répétitives à la place des opérateurs humaines. Ils sont maintenant utilisés dans beaucoup d’applications industrielles, comme soudure, finissage, chargement de machine, manipulation de matériaux etc.

Actuellement, un utilisateur de robot potentiel est confronté à un grand nombre d'options. La décision de quel robot à choisir est fait plus complexe parce que l'exécution de robot est indiquée par un grand nombre de paramètres pour lesquels il n'y a, jusqu'ici, aucune norme à l'échelle industrielle [97]. Récemment, un effort considérable a été concentré sur le développement des modèles afin de faciliter le processus de sélection de robot pour les décideurs. Des modèles de problème de sélection de robot sont classifiés en 5 catégories [98] :

1. Modèles de décision de multi-critères;
2. Modèles d'optimisation de performance des systèmes de production;
3. Modèles assistés par ordinateur;
4. Modèles statistiques; et
5. Les autres approches.

Les approches au problème de sélection de robot sont évaluées en termes de [98] :

1. La considération des attributs de robot (Quels attributs de robot sont considérés pour effectuer le choix ?)
2. La nécessité de données (Quels paramètres doivent être estimés? et Quelle est la difficulté de l'estimation ?)
3. Polyvalence (Le modèle peut-il être appliqué à différentes applications robotiques ou est-il limité à qu'une seule application spécifique?)
4. Les hypothèses du modèle

3.1. La performance des robots

Khouja et Offodile [98] ont groupé les caractéristiques qui sont déjà utilisées pour l'évaluation des robots comme indique le tableau 3.1.

1. Les caractéristiques d'ingénierie: Elles déterminent l'habileté des robots à accomplir des tâches. Elles concernent la capacité de charge, la vitesse, le répétabilité et l'exactitude.

2. Les caractéristiques reliées au fournisseur: Elles déterminent l'attraction des fournisseurs de robot.
3. Les caractéristiques de coût: Elles déterminent les coûts totaux d'installer et de faire travailler des robots.

Tableau 3.1: Caractéristiques utilisées pour l'évaluation des robots [98]

Caractéristiques utilisées pour l'évaluation des robots		
Les caractéristiques d'ingénierie	Les caractéristiques reliées au fournisseur	Les caractéristiques de coût
Capacité de charge Répétabilité Vitesse Méthode de programmation Atteinte verticale Atteinte horizontale Taille de la mémoire Accélération Décélération Degrés de liberté Fiabilité Habilité au diagnostic Langage de programmation Contrôle du logiciel Corrigé de l'erreur	Marque Disponibilité de formation Qualité de formation Documentation Aide pour l'installation Disponibilité des pièces de rechange Délai d'installation Services de pré-vente Disponibilité de rendre service Garantie	Coûts des robots Coût d'installation Coût d'outillage Consommation d'énergie Coût de travail Coût de maintenance

Les principales caractéristiques d'un robot industriel, ainsi que les conditions dans lesquelles les mesures doivent se dérouler, sont définies par les normes ISO 9230. Même si ces caractéristiques peuvent être actuellement vérifiées par des essais décrits dans les normes, l'expérience nous montre que les essais restent parfois insuffisants pour la caractérisation totale d'un robot par rapport à une application. En vérité, pour une application donnée le décideur doit faire face à des trajectoires, des vitesses et des charges spécifiques.

L'exactitude et la répétabilité peuvent être citées parmi les caractéristiques d'ingénierie les plus primordiales mais les plus confuses. La répétabilité est définie comme

l'habileté du robot à retourner au même point. Afin de mieux comprendre le sens de répétabilité et celui d'exactitude supposons que les points d'arrêt du robot forment un ensemble de n points avec le centre en c , en essayant d'exercer la tâche de toucher au point p . Dans ce cas décrit, l'exactitude qui est une mesure de proximité entre les tâches du robot et le point de commande, est égale à la distance entre c et p comme indiquée sur la figure 3.1 [98].

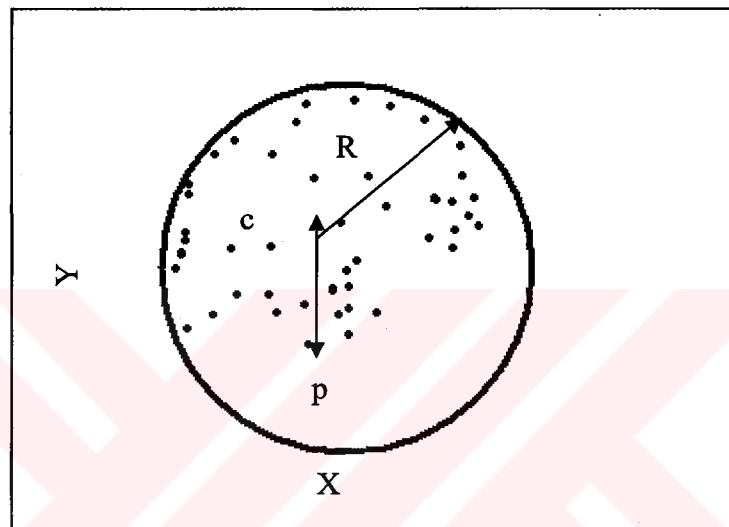


Figure 3.1: Répétabilité et exactitude du robot [97].

La répétabilité peut être définie comme la variation de l'endroit actuel du robot quand le manipulateur le déplace périodiquement vers à un même endroit commandé, tandis que l'exactitude est la différence entre l'endroit actuel et l'endroit commandé. Il est plus facile de corriger une exactitude dégradée qu'une répétabilité en même situation, et par conséquent, généralement on assume que la répétabilité est une caractéristique plus critique [99].

Les fabricants de robots ont une tendance à donner des valeurs spécifiques pour la vitesse maximale du robot, la capacité de charge maximale, et la répétabilité minimale. Cela crée une série de difficultés pour les utilisateurs car ces limites sont souvent mutuellement exclusif. De plus, la relation fonctionnelle parmi ces paramètres est souvent inconnue et rarement, disponible à l'utilisateur comme relation analytique [100].

3.2. Sélection de robot

Khouja et Offodile [98] ont formulé le problème de sélection de robot de façon générale comme suit:

$$\text{Maximiser } F(X) \quad (3.1)$$

$$j \in J$$

sous les contraintes

$$C_j \leq b,$$

$$L_i \leq x_{ij} \leq U_i \quad \forall i \in I.$$

Où

$j \in J$: ensemble des robots commercialement disponibles,

$i \in I$: ensemble des caractéristiques d'un robot,

X est le vecteur de toutes les caractéristiques d'un robot,

X_i est la variable représentant la i -ième caractéristique,

x_{ij} est la valeur de X_i pour le robot j ,

L_i : la valeur minimale requise de X_i déterminée par les tâches qui seront exécutées,

U_i : la valeur maximale permise de X_i déterminée par les tâches qui seront exécutées,

C_j : le coût du robot j ,

b : le budget consacré au robot, et

$f(x)$ est une fonction qui évalue la mesure d'efficacité de la sélection en fonction des caractéristiques du robot.

Même si les contraintes du problème permettent d'éliminer un nombre important de robots, la sélection reste encore difficile parmi les robots admissibles. Comme indiqué sur la figure 3.2, les caractéristiques d'ingénierie ont un effet critique sur la performance des systèmes de production [98], mais elles rendent la sélection plus dure à cause de la dépendance de ces caractéristiques entre elles: les charges lourdes ou la décélération et l'accélération élevées d'un robot de grande vitesse sont à la cause d'une exactitude et d'une répétabilité insuffisantes.

Dans ce travail nous concentrons notre attention sur les modèles statistiques pour le choix de robot. Ce type de modèles du choix de robot est décrit dans les détails dans la partie suivante.

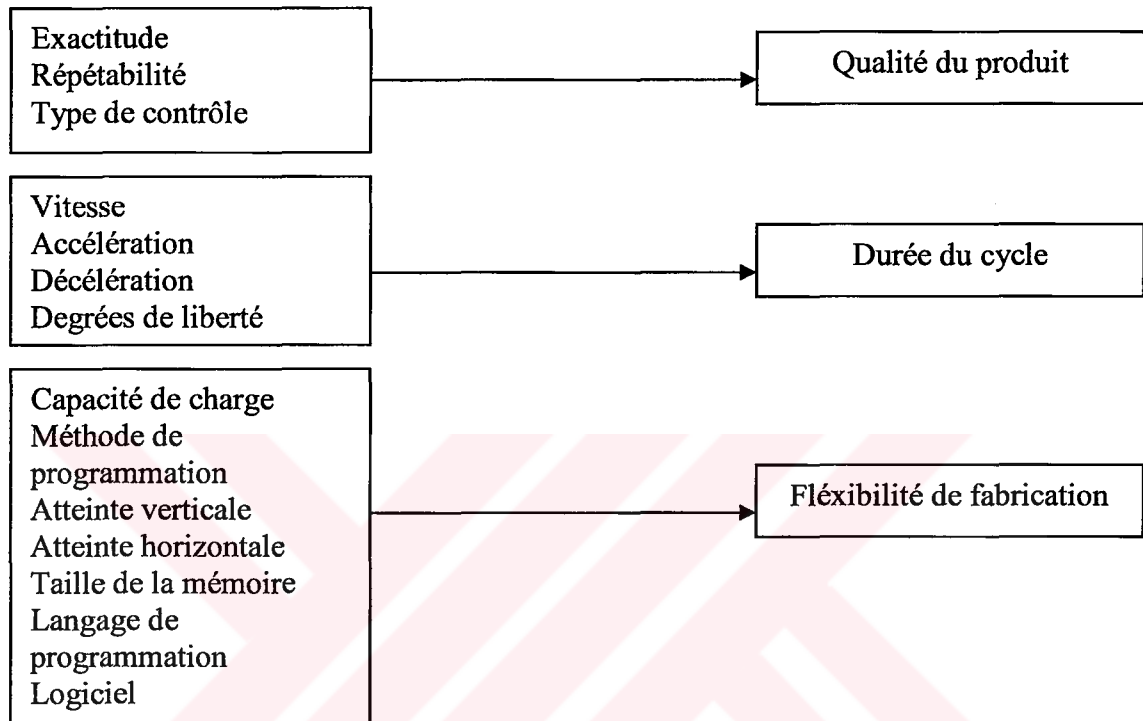


Figure 3.2 : Caractéristiques d'ingénierie des robots et performance des systèmes de production [98].

3.2.1. Modèles statistiques

Le processus de sélection des modèles statistiques est basé sur les caractéristiques d'ingénierie des robots. La performance du robot résultant de ce type de modèle peut manifester des variations en comparaison avec la performance attendue théoriquement, car les seules données requises du problème (les spécifications des fabricants) sont mutuellement exclusif. Les modèles de cette catégorie utilisent des techniques statistiques à multivariées.

Imany et Schlesinger [100] ont essayé de mettre en place une technique formelle pour le processus de sélection de robot industriel. Un modèle basé sur la programmation

linéaire d'objectif et la régression ordinaire de moindres carrés sont appliqués de façon indépendante sur un ensemble de données de 27 robots. La répétabilité est choisie comme la variable dépendante due à son effet critique sur la qualité du produit. Ainsi, l'équation de régression adaptée relie la répétabilité à la capacité de charge et à la vitesse. Le robot qui assure les contraintes et qui possède la plus petite erreur (la valeur de l'erreur " e_i " est définie comme la différence entre y_i , la valeur observée de répétabilité du robot i , et y_i^* , la valeur estimée de répétabilité du robot i) c'est à dire le robot le plus proche à l'hyper-plan de la régression, est sélectionné. A l'aide des résultats, le travail montre aussi que le modèle proposé est plus robuste que la régression ordinaire de moindres carrés en présence des données extrêmes ("outliers").

Booth, Khouja et Hu [101] ont proposé un modèle en utilisant une analyse robuste de la distance "Mahalanobis" et une analyse de composants principaux. A la différence de la plupart des modèles développés pour la sélection de robot industriel, leur modèle tient en compte le fait que la performance du robot choisi est impossible à obtenir en réalité, ce que signale les fabricants des robots en général. Les robots sélectionnés par ce modèle deviennent candidats pour l'usine afin de vérifier les spécifications du fabricant. L'ensemble de données utilisé diffère de l'ensemble qui a été précédemment utilisé dans le travail de Imany et Schlesinger [100] avec un seul robot hypothétique inclus dans l'application.

Le travail publié par Khouja et Booth [102] en 1991, utilise une régression de M-estimateur parce qu'elle est robuste en présence des valeurs extrêmes et elle est efficace pour minimiser la variance. Les détails de leur méthode sont expliqués dans la partie ultérieure faisant comparaison avec la régression linéaire floue.

3.3. Application de la régression linéaire floue à la sélection de robot industriel

La sélection de robot industriel constitue la source des travaux publiés depuis un certain temps. Il existe différentes catégories de modèles pour trouver une solution au problème donné. Nous concentrons sur les modèles statistiques et particulièrement celui proposé par Khouja et Booth [102].

Afin de montrer la validité d'utilisation de régression linéaire floue pour la sélection de robot, les résultats obtenus par la régression floue sont comparés avec ceux de la régression robuste de M-estimateur utilisée par Khouja et Booth [102]. Leur méthodologie est appliquée à la régression linéaire floue et la corrélation des résultats est déterminée par le test de corrélation du rang de Spearman.

3.3.1. Méthodologie

Un ensemble de données comportant les 4 caractéristiques –le coût, la capacité de charge, la vitesse et la répétabilité– de 27 robots est utilisé pour l'application. La répétabilité est choisie comme la variable dépendante à cause de son effet primordiale sur la qualité. Le critère de coût constitue la contrainte budgétaire pour le décideur. Ainsi l'équation de régression qui met en évidence la relation de la répétabilité avec la capacité de charge et la vitesse est de la forme suivante:

$$y_i^* = a_0 + a_1 x_{i1} + a_2 x_{i2} \quad (3.2)$$

où y_i^* la valeur de répétabilité estimée, x_{i1} la valeur de la capacité de charge et x_{i2} la valeur de la vitesse du robot i .

L'équation de régression trouvée permet de calculer les valeurs de répétabilité estimées pour les robots candidats en utilisant les valeurs observées des variables indépendantes c'est à dire la capacité de charge et la vitesse. Il est naturel qu'une différence existe entre les valeurs de répétabilité estimées et celles d'observées. Cet écart est représenté par e_i pour le robot i .

$$e_i = y_i - y_i^* \quad (3.3)$$

Les valeurs strictement négatives de e_i ($e_i < 0$), impliquent que la répétabilité estimée est supérieure à la répétabilité observée ($y_i < y_i^*$). Le i -ième robot est plus performant

que l'on prévoit et la grandeur de e_i négative indique son échelle de performance. Plus la valeur de $|e_i|$ devient grande, plus le robot est un robot extrême (ou "outlier") et par conséquent, le robot est plus ou moins performant dépendant dans quelle partie de l'hyperplan de régression il se trouve [102]. Il ne faut pas oublier que notre but est la minimisation de la répétabilité, et c'est ainsi que le robot possédant la valeur de e_i minimale et qui assure les contraintes budgétaires devient candidat pour le test et la sélection.

Khouja et Booth [102] en appliquant la régression robuste de M-estimateur, afin de vaincre le problème de la variance non-constante, ont pris le logarithme népérien de la capacité de charge et celle de la répétabilité. Pour faire une comparaison exacte des résultats nous avons décidé d'utiliser aussi l'ensemble de données modifié par eux. Puis, le modèle de régression linéaire floue de Tanaka a été appliqué. La valeur de H a été choisie égale à "0.5".

Tableau 3.2: Les valeurs des 4 critères pour 27 robots [100]

Robot	Coût (\$ 10.000)	Capacité de Charge (kilogramme)	Vitesse (m/s)	Répétabilité (millimètre)
1	7.20	60.0	1.35	0.150
2	4.80	6.0	1.1	0.05
3	5.00	45.0	1.27	1.270
4	7.20	1.5	0.66	0.025
5	9.60	50.0	0.05	0.250
6	1.07	1.0	0.30	0.100
7	1.76	5.0	1.00	0.100
8	3.20	15.0	1.00	0.100
9	6.72	10.0	1.10	0.200
10	2.40	6.0	1.00	0.050
11	2.88	30.0	0.90	0.500
12	6.90	13.6	0.15	1.000
13	3.20	10.0	1.20	0.050
14	4.00	30.0	1.20	0.050
15	3.68	47.0	1.00	1.000
16	6.88	80.0	1.00	1.000
17	8.00	15.0	2.00	2.000
18	6.30	10.0	1.00	0.200
19	0.94	10.0	0.30	0.050
20	0.16	1.5	0.80	2.000
21	2.81	27.0	1.70	2.000
22	3.80	0.9	1.00	0.050
23	1.25	2.5	0.50	0.100
24	1.37	2.5	0.50	0.100
25	3.63	10.0	1.00	0.200
26	5.30	70.0	1.25	1.270
27	4.00	205.0	0.75	2.030

Tableau 3.3: Les valeurs utilisées dans les modèles

Robot	Capacité de Charge (ln)	Vitesse	Répétabilité (ln)
1	4.094345	1.35	-1.89712
2	1.791759	1.1	-2.99573
3	3.806662	1.27	0.239017
4	0.405465	0.66	-3.68888
5	3.912023	0.05	-1.38629
6	0	0.30	-2.30259
7	1.609438	1.00	-2.30259
8	2.70805	1.00	-2.30259
9	2.302585	1.10	-1.60944
10	1.791759	1.00	-2.99573
11	3.401197	0.90	-0.69315
12	2.61007	0.15	0
13	2.302585	1.20	-2.99573
14	3.401197	1.20	-2.99573
15	3.850148	1.00	0
16	4.382027	1.00	0
17	2.70805	2.00	0.693147
18	2.302585	1.00	-1.60944
19	2.302585	0.30	-2.99573
20	0.405465	0.80	0.693147
21	3.295837	1.70	0.693147
22	-0.10536	1.00	-2.99573
23	0.916291	0.50	-2.30259
24	0.916291	0.50	-2.30259
25	2.302585	1.00	-1.60944
26	4.248495	1.25	0.239017
27	5.32301	0.75	0.708036

Modèle de Tanaka utilisé:

$$\text{Minimiser } \sum_{j=0}^N (c_j \sum_{i=1}^M |x_{ij}|) \quad (3.4)$$

Sous les contraintes

$$\sum_{j=0}^N \alpha_j x_{ij} + \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \geq y_i, \quad i = 1, 2, \dots, M$$

$$\sum_{j=0}^N \alpha_j x_{ij} - \left| L^{-1}(H) \right| \sum_{j=0}^N c_j |x_{ij}| \leq y_i, \quad i = 1, 2, \dots, M$$

$$c_j \geq 0, \quad j = 0, 1, \dots, N$$

Le modèle proposé par Tanaka a été critiqué vivement à cause de sa faiblesse envers les valeurs extrêmes. Ces valeurs présentes dans un ensemble de données détériore la performance du modèle et change la forme de la droite de régression. Le modèle oblige les données à être dans l'intervalle estimé; c'est à dire les limites de l'intervalle sont déterminées par la valeur la plus extrême. Ainsi on observe des déviations pour les variables dépendantes estimées, ce qui n'est pas du tout désirable.

Modèle de Peters

$$\text{Maximiser } \bar{\lambda} = 1/M \sum_{i=1}^M \lambda_i \quad (3.5)$$

Sous les contraintes

$$(1 - \bar{\lambda}) p_0 - \sum_{i=1}^M \sum_{j=0}^N c_j |x_{ij}| \geq -d_0$$

$$(1 - \lambda_i) p_i + \sum_{j=0}^N \alpha_j x_{ij} + \sum_{j=0}^N c_j |x_{ij}| \geq y_i, \quad i = 1, 2, \dots, M$$

$$(1 - \lambda_i) p_i - \sum_{j=0}^N \alpha_j x_{ij} + \sum_{j=0}^N c_j |x_{ij}| \geq -y_i, \quad i = 1, 2, \dots, M$$

$$\lambda_i \leq 1, \quad \left| L^{-1}(H) \right| := 1 \quad c_j \geq 0, \quad j = 0, 1, \dots, N$$

où p_i est la largeur de l'intervalle de tolérance de y_i . Les paramètres p_o et p_i doivent être choisis selon le contexte général du problème. Tandis qu'une valeur élevée de p_o et des faibles valeurs de p_i signifient un faible besoin de minimisation des écarts, une petite valeur affectée à p_o et des grandes valeurs de p_i forment un intervalle étroit.

La variable λ représente le degré d'appartenance de la solution à l'ensemble des "bonnes solutions". Le paramètre d_0 représente la valeur désirée de la fonction objectif. Comme nous préférons obtenir un modèle non-flou (minimisation de l'incertitude) d_0 peut-être choisi égale à 0, cette situation est représentée par $\lambda = 1$. Le modèle a été modifié en utilisant la moyenne arithmétique comme l'opérateur d'agrégation pour λ . La compensation entre les données est ainsi réalisée, chacune influence la fonction de régression avec un poids de $1/M$ [41].

Peters [41], avec son modèle dont l'intervalle estimé est influencé par toutes les données existantes est une viable alternative à l'approche de Tanaka. Cette approche est robuste en présence des données extrêmes.

3.3.2. Résultats et conclusion

Les deux modèles sont appliqués. Les équations des droites de régression trouvées ci-dessous sont obtenus en utilisant le logiciel GAMS. Les modèles écrits en code GAMS se trouvent dans l'appendice A et B.

Modèle de Tanaka

$$\ln(\text{répétabilité}) = (-2.346, 4.211) + 1.149 * \text{vitesse} + 0.024 * \ln(\text{capacité de charge}) \quad (3.6)$$

Modèle de Peters

$$\ln(\text{répétabilité}) = (-1.942, 1.418) + (0.1, 0.268) * \text{vitesse} + 0.18 * \ln(\text{capacité de charge}) \quad (3.7)$$

Les valeurs de y_i^* sont calculées uniquement en utilisant la moyenne des paramètres flous trouvés. Les résultats sont présentés dans les tableaux 3.4 et 3.5.

Tableau 3.4: Résultats du modèle de Tanaka

Rang du robot	Numéro du robot	Répétabilité estimée (ln) (Tanaka) y_i^*	$e_i = y_i - y_i^*$
1	R4	-3.68888	-2.11095
2	R14	-2.99573	-2.11016
3	R13	-2.99573	-2.08379
4	R2	-2.99573	-1.95663
5	R10	-2.99573	-1.84173
6	R22	-2.99573	-1.7962
7	R1	-1.89712	-1.20053
8	R8	-2.30259	-1.17058
9	R7	-2.30259	-1.14421
10	R19	-2.99573	-1.04969
11	R9	-1.60944	-0.5826
12	R23	-2.30259	-0.55308
13	R24	-2.30259	-0.55308
14	R18	-1.60944	-0.4677
15	R25	-1.60944	-0.4677
16	R6	-2.30259	-0.30129
17	R11	-0.69315	0.537124
18	R17	0.693147	0.676154
19	R5	-1.38629	0.808367
20	R21	0.693147	1.006747
21	R3	0.239017	1.034427
22	R26	0.239017	1.046803
23	R16	0	1.091831
24	R15	0	1.104596
25	R27	0.708036	2.064534
26	R20	0.693147	2.110216
27	R12	0	2.111008

Tableau 3.5: Résultats du modèle de Peters

Rang du robot	Numéro du robot	Répétabilité estimée (ln) (Peters) y_i	$e_i = y_i - y_i^*$
1	R4	-3,6888795	-1,8858632
2	R14	-2,9957323	-1,7859478
3	R13	-2,9957323	-1,5881976
4	R19	-2,9957323	-1,4981976
5	R2	-2,9957323	-1,486249
6	R10	-2,9957323	-1,476249
7	R22	-2,9957323	-1,1347674
8	R8	-2,3025851	-0,9480341
9	R1	-1,89712	-0,827102
10	R7	-2,3025851	-0,7502839
11	R23	-2,3025851	-0,5755174
12	R24	-2,3025851	-0,5755174
13	R6	-2,3025851	-0,3905851
14	R9	-1,6094379	-0,1919032
15	R18	-1,6094379	-0,1819032
16	R25	-1,6094379	-0,1819032
17	R5	-1,3862944	-0,1534585
18	R11	-0,6931472	0,5466373
19	R16	0	1,0532352
20	R15	0	1,1489734
21	R26	0,2390169	1,2912878
22	R3	0,2390169	1,3688177
23	R12	0	1,4571874
24	R27	0,7080358	1,616894
25	R21	0,6931472	1,8718965
26	R17	0,6931472	1,9476981
27	R20	0,6931472	2,4821635

Supposons qu'un directeur doit choisir un robot d'une capacité de charge de 7 kg., qui fonctionne avec une vitesse de 0.6 m/s et qui possède une répétabilité de 0.10 mm. Un budget de \$ 37.000 est consacré à l'achat de ce robot. Dans cette situation Khouja et Booth [102] choisissent le 13 ième robot , qui est aussi le robot choisi par les modèles de régression floue.

Afin de montrer la validité d'utilisation de la régression linéaire floue, les coefficients de corrélation du rang de Spearman des deux modèles sont calculés.

Tableau 3.6: Calcul du coefficient de la corrélation du rang de Spearman pour le modèle de Tanaka

Numéro du robot	Khouja & Booth	Régression linéaire floue (Tanaka)	Différence (d)	d ²
1	2	7	-5	25
2	4	4	0	0
3	22	21	1	1
4	8	1	7	49
5	10	19	-9	81
6	24	16	8	64
7	9	9	0	0
8	6	8	-2	4
9	11	11	0	0
10	5	5	0	0
11	16	17	-1	1
12	26	27	-1	1
13	3	3	0	0
14	1	2	-1	1
15	21	24	-3	9
16	15	23	-8	64
17	25	18	7	49
18	13	14	-1	1
19	7	10	-3	9
20	27	26	1	1
21	23	20	3	9
22	14	6	8	64
23	19	12	7	49
24	20	13	7	49
25	12	15	-3	9
26	17	22	-5	25
27	18	25	-7	49
				614
r_s = 0.8125764				

Tableau 3.7: Calcul du coefficient de la corrélation du rang de Spearman pour le modèle de Peters

Numéro du robot	Khouja & Booth	Régression linéaire floue (Peters)	Différence (d)	d ²
1	2	9	-7	49
2	4	5	-1	1
3	22	22	0	0
4	8	1	7	49
5	10	17	-7	49
6	24	13	11	121
7	9	10	-1	1
8	6	8	-2	4
9	11	14	-3	9
10	5	6	-1	1
11	16	18	-2	4
12	26	23	3	9
13	3	3	0	0
14	1	2	-1	1
15	21	20	1	1
16	15	19	-4	16
17	25	26	-1	1
18	13	15	-2	4
19	7	4	3	9
20	27	27	0	0
21	23	25	-2	4
22	14	7	7	49
23	19	11	8	64
24	20	12	8	64
25	12	16	-4	16
26	17	21	-4	16
27	18	24	-6	36
				578
r_s = 0,8235653				

La formule du coefficient de corrélation du rang de Spearman est la suivante:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (3.8)$$

où n représente le nombre de données.

D'après les coefficients calculées, l'hypothèse nulle qui défend l'indépendance des observations est rejetée (Les valeurs limites de la distribution du coefficient de corrélation du rang de Spearman se trouvent dans l'appendice C). Les résultats montrent que le modèle de Peters est plus robuste quand il s'agit d'un ensemble de données contenant des valeurs extrêmes comme nous possédons dans cet exemple-là. La régression floue linéaire peut être utilisée pour la sélection de robot industriel à la place de la régression statistique, aux cas où les paramètres sont floues, ou l'incertitude règne sur le système étudié, ou encore l'ensemble de données n'est pas suffisant pour pouvoir appliquer la régression statistique classique.

3.4. Modèle de décision basé sur la régression linéaire floue pour la sélection de robot industriel

3.4.1. Préliminaires

La partie précédente du travail a fait preuve d'une suffisance de capacité pour la régression linéaire floue à propos de son utilisation à la place de la régression statistique dans le cas d'un problème de sélection de robot industriel. Le travail a été effectué en utilisant le même ensemble de données utilisé par Khouja et Booth [102] afin d'obtenir des résultats comparables. Khouja et Booth [102], dans leur travail, face au problème de la variance non-constante ont pris le logarithme népérien de la capacité de charge et celle de la répétabilité. La régression floue, étant une méthode non-statistique, ne s'occupe pas de la variance des termes. Par conséquent, l'ensemble de données peut être utilisé dans sa forme originale, sans avoir été modifié par la fonction logarithmique.

Khouja et Booth [102] considèrent le critère de coût comme une contrainte budgétaire pour le décideur et ne le font pas intervenir dans l'équation de régression. Pour élargir le nombre de critères qui influencent le résultat du modèle, le coût peut être aisément inclus dans l'analyse. Le modèle proposé dans cette partie prend en compte ces modifications primordiales.

3.4.2. Méthodologie

L'ensemble de données comportant les 4 caractéristiques –le coût, la capacité de charge, la vitesse et la répétabilité de 27 robots est encore utilisé pour l'application. La répétabilité est choisie comme la variable dépendante à cause de son effet primordiale sur la qualité. Une équation de régression est établie de façon où elle met en évidence la relation de la répétabilité avec la vitesse, la capacité de charge et le coût.

$$y_i^* = a_0 + a_1 x_{i1} + a_2 x_{i2} + a_3 x_{i3} \quad (3.9)$$

où y_i^* désigne la valeur de répétabilité estimée, x_{i1} la valeur de la vitesse, x_{i2} la valeur de la capacité de charge et x_{i3} le coût du robot i . La méthodologie appliquée est identique à celle représentée dans la partie précédente. Le même ensemble de données et la même équation de régression a été déjà utilisés par Imany et Schlesinger [100] avec une méthodologie différente.

Le modèle de Peters est choisi comme le modèle de régression linéaire floue pour cette partie de travail. Le caractère robuste de ce modèle qu'on a vérifié dans la partie précédente est la facteur dominante qui le met en premier rang parmi les autres modèles de régression floue. Notre modèle de décision basé sur la régression linéaire floue choisit le robot qui a la plus grande valeur de $|e_i|$ (c'est une donnée extrême) entre les robots possédant une valeur d'écart $e_i = y_i - y_i^*$ négative. Le modèle de Peters satisfait les besoins de la méthodologie: il est robuste en présence des données extrêmes et permet de les identifier.

3.4.3. Résultats de l'application

L'équation de régression qui concrétise la relation de la répétabilité avec la vitesse, la capacité de charge et le coût, obtenue grâce au modèle de Peters (où $d_0 = 0$, $p_o = 1000$ et $p_i = 1$) est la suivante:

$$\text{Répétabilité} = (0.550, 0.526) + (0.315, 0.241) * \text{vitesse} + 0.003 * \text{capacité de charge} - 0.029 * \text{coût} \quad (3.10)$$

Les valeurs de y_i^* sont calculées à partir de (3.10) en utilisant uniquement la moyenne des paramètres flous trouvés. Les résultats sont présentés dans le tableau 3.8. Le modèle écrit en code GAMS se trouve dans l'appendice D.

Tableau 3.8: Nouveaux résultats du modèle de Peters

Rang du robot	Numéro du Robot	Répétabilité estimée y_i^*	$e_i = y_i - y_i^*$
1	R14	0.94645	-0.852
2	R13	0.7753	-0.8152
3	R1	0.94005	-0.79645
4	R10	0.5536	-0.7634
5	R7	0.43735	-0.72896
6	R2	0.61647	-0.7253
7	R8	0.82896	-0.7172
8	R22	0.8172	-0.7075
9	R19	0.73162	-0.59724
10	R25	0.8134	-0.58973
11	R23	0.83998	-0.57875
12	R24	0.43795	-0.57527
13	R9	0.8652	-0.53162
14	R4	0.902	-0.5286
15	R6	0.89928	-0.51647
16	R18	0.90548	-0.5123
17	R11	0.993	-0.33998
18	R5	0.7123	-0.18735
19	R16	0.64724	0.09452
20	R15	0.80186	0.10072
21	R26	1.08501	0.26995
22	R3	0.7575	0.32995
23	R12	0.67875	0.56205
24	R27	0.67527	0.74475
25	R21	0.78973	0.91499
26	R17	1.00005	1.007
27	R20	1.28525	1.19814

Supposons qu'un décideur exige un robot avec une répétabilité inférieure à 0.2 mm et son budget est limité à \$ 37.000. D'après les résultats présentés dans le tableau 3.8, le robot numéro13 (R13), qui occupe le deuxième rang parmi les 27 robots satisfait les exigences déterminés. Par conséquent le modèle basé sur la régression linéaire floue de Peters choisit le robot 13 comme le meilleur candidat. La méthode de Imany et Schlesinger [100] choisit le robot 25 si on considère les mêmes circonstances. D'après les données du tableau 3.2, on peut facilement remarquer que le meilleur robot déterminé dans ce travail coûte moins cher et possède une répétabilité plus petite et une vitesse plus grande que ce qui est choisi par Imany et Schlesinger [100]. Si on ajoute à ces remarques l'égalité des capacités de charge, cette analyse aboutit à une conclusion évidente: le meilleur robot déterminé dans ce travail est plus performant que le robot déterminé par l'analyse d'Imany et Schlesinger [100].



4. Evaluation des alternatives de technologie à l'aide de la méthode de régression linéaire floue

4.1. Introduction aux systèmes flexibles de fabrication

L'augmentation perpétuelle du nombre des technologies existantes rendent de plus en plus difficile le processus de prendre une décision définitive de la technologie qui va être utilisée. La sélection de technologie est la plus importante décision afin de contrôler le milieu concurrentiel. A cause de la variété et la complexité des technologies industrielles, il est assez compliqué de choisir l'alternative la plus conforme aux attentes d'une compagnie. Face à des problèmes de critères multiples au sujet de sélection et d'affectation des nouvelles technologies, de nombreux modèles ont été développés dans la littérature.

L'adoption des systèmes flexibles de fabrication (SFF) qui font partie des technologies avancées, attire l'attention des chercheurs depuis une vingtaine d'années. Le gain de flexibilité, l'augmentation de la qualité, la diminution des coûts et des durées de cycle sont compris parmi les avantages d'intégrer un système flexible de fabrication qui sont prouvés expérimentalement. Malgré ces nombreux avantages, les systèmes de technologie avancées, en général, n'ont pas pu se diffuser totalement dans le monde industriel. L'adoption et la justification de la littérature des systèmes flexibles de fabrication s'est focalisée en fonction pourquoi ces systèmes n'ont pas été largement mis en application [103]. Les raisons du manque d'adoption de ces systèmes sont reliées aux facteurs techniques, économiques, organisationnelles ou politiques [104]. De plus, la sélection de ce type de technologie, rend obligatoire l'analyse de plusieurs caractéristiques qui augmente la complexité des modèles.

4.2. Les modèles d'évaluation des systèmes flexibles de fabrication

Les modèles d'évaluation des systèmes flexibles de fabrication et les autres technologies stratégiques de fabrication se forment un large éventail constitué des approches mathématiques et de modélisation des systèmes [104]. Durant les années 90, un grand nombre d'articles ont été publiés sur la sélection des technologies avancées. Ces travaux peuvent être groupés en 3 catégories essentielles: l'analyse économique, les techniques stratégiques, et les méthodes analytiques [105].

L'analyse économique est constitué des techniques classiques d'évaluation de l'économie d'ingénierie. On peut en citer comme exemples le taux de rentabilité interne (internal rate of return: IRR) et la valeur actualisée nette (net present value: NPV) qui ont déjà été fréquemment utilisés dans la littérature. Le défaut primordial des techniques de l'analyse économique, c'est qu'elles ne prennent pas en compte les bénéfices stratégiques et non-économiques [105].

La prise en compte des objectifs généraux de la société a causé l'intégration des mesures stratégiques et tactiques dans l'analyse. Les méthodes stratégiques d'évaluation nécessitent moins de données techniques et sont utilisées souvent à côté des méthodes économiques ou analytiques [105].

Les approches analytiques nécessitent en général beaucoup plus de données et elles sont considérées comme difficiles par rapport à l'analyse économique. L'incertitude du futur et la présence des objectifs multiples peuvent être considérés dans ces méthodes. L'analyse de la valeur, la modélisation mathématique sont des applications essentielles de l'approche analytique [105].

L'utilisation des techniques modernes et les approches nouvelles de modélisation a créé une augmentation du nombre de modèles et a agrandi leurs régions d'application. La programmation dynamique, la théorie de décision, la programmation linéaire, l'analyse de simulation, le processus hiérarchique analytique, la méthode d'enveloppement des données sont les exemples des méthodologies récemment utilisées [104]. Shang et

Sueyoshi [106] utilise le processus hiérarchique analytique, la simulation et la méthode d'enveloppement des données pour la sélection d'un SFF, Sarkis [104] établit un modèle basé sur la méthode d'enveloppement des données et Karsak et Kuzgunkaya [105] propose un modèle flou à objectifs multiples. L'article de Proctor et de Canada [107] donne des renseignements détaillés sur les méthodes d'évaluation des investissements de la fabrication. Le travail élaboré par Son [108] est une véritable source bibliographique qui contient 274 articles déjà publiés sur ce sujet.

4.3. Les critères utilisées

Comme nous l'avons déjà signalé, le processus de sélection des technologies comporte l'analyse d'un grand nombre de facteurs. Une liste limitée peut être donnée afin de montrer la variété des critères utilisées dans la littérature: qualité, débit, travail en cours, temps d'écoulement, flexibilité de la gamme de fabrication, temps d'arrêt, coût total, taux de production, risque, etc. Il existe plusieurs travaux publiés détaillés faits sur les critères possibles pour la sélection des systèmes flexibles de fabrication. Les critères doivent être déterminées par la société et doivent refléter les stratégies adoptées.

Les critères qui vont être utilisées pour la technique de sélection d'un système de fabrication à l'aide de régression linéaire floue sont déterminées à la suite de l'analyse des plusieurs travaux [103, 104, 106 et 107]. Les critères choisies sont les suivantes:

- Débit ("Throughput" en anglais): Dans le monde industriel, le débit d'un système est le taux de transformation des produits en argent, contrairement aux sorties (ou "ouput" en anglais) qui peuvent aussi être stockés dans un dépôt.
- Flexibilité de capacité ("Volume flexibility" en anglais): La flexibilité de capacité représente la capacité disponible d'un système pour la production des produits pendant un intervalle de temps donné [104].

- Flexibilité de la gamme de fabrication (“Routing flexibility” en anglais): Le nombre d’opérations par un groupe de machines dans un système flexible de fabrication [104].
- Travail en cours (“Work in process” –WIP- en anglais): Le travail en cours représente les stocks en cours des produits d’un système qui fonctionne.
- Temps d’arrêt (“Downtime” en anglais): La durée totale des arrêts d’un système productif. Les temps d’arrêts sont dus en général aux périodes de panne des machines ou les temps consacrés à la calibration pour la production d’un nouveau produit. Pour la méthodologie utilisée, on considère comme critère, l’amélioration possible des temps d’arrêt ce qui est plus visible et compréhensible pour les décideurs.

L’unité de mesure des critères citées ci-dessus et les limites inférieures et supérieures sont données dans le tableau suivant.

Tableau 4.1: Critères considérés pour la sélection d’un système de fabrication

Critère Considérée	Maximiser ou Minimiser	Unité de Mesure	Limites Inférieures et Supérieures
Débit (D)	Max	Unités de produits	(1500, 3000)
Flexibilité de Capacité (FC)	Max	La capacité de production moyenne par type de produit	(100,300)
Flexibilité de la Gamme de Fabrication (FGF)	Max	Le nombre d’opérations moyen par groupe de machines	(1, 5)
Travail En Cours (TEC)	Min	Unités de produits	(100, 400)
Amélioration des Temps d’Arrêt (TA)	Max	Pourcentage d’amélioration par rapport à l’existant	(1, 15)

4.4. La méthodologie proposée

4.4.1 Utilisation de la régression linéaire floue

La sélection des systèmes flexibles de fabrication a été déjà traitée par nombreux techniques scientifiques. Nous proposons un nouveau modèle de sélection basé sur la régression linéaire floue. La justification de l'utilisation de la méthode de régression linéaire floue, est faite de façon convaincante en examinant les critères de sélection. Comme nous l'avons signalé précédemment, les problèmes de sélection des nouvelles technologies obligent les décideurs de prendre en compte un nombre important de critères qualitatives ou quantitatives. De plus, certaines critères qui semblent d'être mesurées par des méthodes expérimentales, sont recueillies de manière subjective. C'est ainsi qu'il est souvent impossible de posséder des données chiffrées exactes, les valeurs obtenues sont approximatives. Si on compte aussi des critères linguistiques, l'utilisation de la théorie des sous ensembles flous devient indispensable. Même si la régression statistique peut être utilisable avec un ensemble de données suffisamment large; la restriction du nombre des alternatives, qui est souvent le cas, constitue une base conforme à la régression floue.

Le débit qui montre concrètement le succès du système, constitue la variable dépendante de l'analyse de régression, et les autres critères définies précédemment - flexibilité de capacité, flexibilité de la gamme de fabrication, travail en cours, amélioration des temps d'arrêt- qui influencent la performance du système, constituent les variables indépendantes. Ainsi, une équation de régression qui relie le débit aux variables indépendantes est établie comme suivante:

$$y_i^* = a_0 + a_1 x_{i1} + a_2 x_{i2} + a_3 x_{i3} + a_4 x_{i4} \quad (4.1)$$

Le débit est donné sous forme d'un nombre flou triangulaire symétrique représenté par (y_i, e_i) où y_i est la moyenne et e_i représente la dispersion de la valeur du débit i . Les variables indépendantes sont supposées non-floues dans le premier cas étudié. Pour le deuxième cas, la situation est généralisée aux variables dépendantes et indépendantes

floues et donc les variables indépendantes sont aussi considérées sous forme des nombres flous triangulaires symétriques représentés par (\bar{x}_{ij}, d_{ij}) où \bar{x}_{ij} est la moyenne et d_{ij} représente la dispersion. Les variables indépendantes dont la valeur d_{ij} est nulle sont des nombres non-flous.

Les coefficients du premier cas de l'équation de régression sont calculés grâce au modèle de la régression linéaire floue de Tanaka. Un modèle plus complexe, ce de Sakawa et Yano [33], est utilisé pour le deuxième cas. Après avoir trouvé les coefficients des variables indépendantes, l'équation de régression permet de calculer les valeurs estimées y_i^* de débit. Les coefficients étant flous, les valeurs estimées deviennent floues aussi.

4.4.2. Sélection de SFF par la programmation booléenne floue

L'équation de régression trouvée permet de calculer les valeurs de débit estimées pour les systèmes flexibles de fabrication candidats en utilisant les valeurs observées des variables indépendantes. Il est naturel qu'une différence existe entre les valeurs de débit estimées et celles d'observées. Cet écart est représenté par e_i pour le i -ième SFF.

$$e_i = y_i - y_i^* \quad (4.2)$$

Les valeurs strictement positives de e_i ($e_i > 0$), impliquent que le débit estimé est inférieur au débit observé ($y_i > y_i^*$). Le i -ième SFF est plus performant que l'on prévoit et la grandeur de e_i positive indique son échelle de performance. Plus la valeur de $|e_i|$ devient grande, plus le SFF est un SFF extrême (ou "outlier") et par conséquent, le SFF est plus ou moins performant dépendant dans quelle partie de l'hyperplan de régression il se trouve [102]. Il ne faut pas oublier que notre but est de maximiser le débit des systèmes, et c'est ainsi que le SFF possédant la valeur de e_i maximale se classe au premier rang.

Comme les valeurs estimées et observées du débit sont floues, la différence entre ces deux valeurs c'est à dire e_i est obtenue sous forme d'un nombre flou. Etant donné que la différence de deux nombres flous symétriques est un nombre symétrique, e_i est un nombre flou symétrique. La différence de deux nombres flous est calculé suivant la formule (2.6).

Un modèle de programmation booléenne floue est établi pour la sélection de SFF. Comme le modèle nécessite les valeurs de classement de débit des alternatives, les méthodes de classement des nombres flous sont étudiées en premier.

4.4.2.1. Méthodes de classement des sous-ensembles flous

Un grand nombre de méthodes de classement des sous-ensembles flous sont proposées dans la littérature. Bien que leur utilisation soit assez répandue, la plupart de ces méthodes sont critiquées vivement à cause des résultats qui restent opposés à l'intuition. Le problème de classement des nombres flous est considéré très important par le fait que ces nombres flous qu'on doit classer sont en général obtenus étant des valeurs reliées à un ensemble d'alternatives dans des problèmes de décision. Par conséquent la comparaison des deux sous-ensembles est en effet la comparaison de deux alternatives dans un processus de décision [109].

Le problème traité ici est de classer n sous-ensembles flous de l'intervalle d'unité I . La restriction du domaine de définition de \mathfrak{R} à I ne cause aucune perte de généralité. La critique essentielle adressée aux méthodes de classement est qu'elles tendent à défuzzifier un classement flou pour obtenir un classement non-flou [109].

Les méthodes de classement des sous-ensembles flous peuvent être sousgroupées en diverses catégories selon l'approche utilisée. Yager [110], Chang [111] et Adamo [112] utilisent une méthode simple qui consiste à définir une fonction de classement qui envoie chaque sous-ensemble flou à une ligne réelle où une ordre naturel existe. Les autres auteurs comme Baas et Kwakernaak [113], Baldwin et Guild [114] ont pour but d'obtenir un ensemble flou d'alternatives optimales où est le degré auquel l'alternative

i peut être considéré comme optimale. Jain [115, 116] utilise l'opérateur flou de maximisation. Dubois et Prade [117] proposent un ensemble de 4 indices qui sont complètement capables de décrire la location relative de 2 nombres flous.

La plupart des premières méthodes qui ont été apparues, étaient destinées à classer les nombres flous positifs. Le classement des nombres flous négatifs qu'on rencontre pas souvent est traité dans peu d'articles. L'un des travaux importants fait sur ce sujet est l'article de Yao et Wu [118]. Ils proposent une méthode de classement des nombres flous basée sur le principe de décomposition et la distance signée. Cette méthode permet de classer les nombres flous négatifs et positifs. Chu et Tsao [119] proposent une méthode récente permettant de classer aussi les nombres flous négatifs et positifs avec l'aire entre le centre de gravité et le point d'origine. Comme ils testent l'utilisation de leur méthode par des exemples précis, cette méthode est choisie pour la première étape de notre modèle de sélection de SFF par la programmation booléenne floue.

4.4.2.1.1. Méthode de Chu et Tsao

Chu et Tsao [119] proposent une méthode nouvelle qui calcule l'aire entre le point où se trouve le centre de gravité et le point d'origine d'un nombre flou pour classer un ensemble de nombres flous entres eux. Le point du centre de gravité (\bar{x}, \bar{y}) possède ses coordonnées sur l'axe des abscisses et sur l'axe des ordonnées et a pour définition:

$$\bar{x}(\tilde{A}) = \frac{\int_a^b (x f_A^L) dx + \int_b^c x dx + \int_c^d (x f_A^R) dx}{\int_a^b (f_A^L) dx + \int_b^c dx + \int_c^d (f_A^R) dx} \quad (4.3)$$

$$\bar{y}(\tilde{A}) = \frac{\int_0^{\omega} (y g_A^L) dy + \int_0^{\omega} (y g_A^R) dy}{\int_0^{\omega} (g_A^L) dy + \int_0^{\omega} (g_A^R) dy} \quad (4.4)$$

où les fonctions d'appartenance gauche (left: L) et droite (right: R) d'un nombre flou $\tilde{A} = (a, b, c)$ sont représentées par (4.5) et (4.6).

$$f^R(x) = \frac{c-x}{c-b} \quad (4.5)$$

$$f^L(x) = \frac{x-a}{b-a} \quad (4.6)$$

La fonction $g(x)$ est la fonction inverse de $f(x)$. L'aire entre le point de centre de gravité (\bar{x}, \bar{y}) et le point d'origine $(0,0)$ d'un nombre flou \tilde{A} est égal à $S(\tilde{A}) = \bar{x} \cdot \bar{y}$. Le nombre flou possédant la plus grande valeur de S est considérée comme le plus grand des autres [119].

4.4.2.2. Modèle de programmation booléenne floue

Le problème de sélection de SFF a pour but de maximiser le débit des systèmes, c'est à dire nous essayons de trouver le SFF qui est le plus performant par rapport à l'attendu, ainsi le modèle maximise les valeurs de différence calculées, e_i . En réalisant cette objective, les sociétés doivent prendre en compte les différentes contraintes qu'elles sont d'obliger de respecter.

Les valeurs de différences calculées à partir des droites de régression seront sous forme des nombres flous comme on l'avait déjà précisé. Un problème qui sélectionne le SFF optimum en maximisant les valeurs de différence sous les contraintes déterminées peut être formulé comme un programme linéaire booléen flou ci-dessous:

$$\text{Max } z = \sum_{i \in M} \tilde{d}_i x_i \quad (4.7)$$

Sous les contraintes

$$\sum_{i \in M} \alpha_{ij} x_i \leq b_t, \quad t = 1, 2, \dots, T$$

$$x_i \in \{0, 1\}, \quad i \in M$$

Le problème est un où d_i représente la valeur de e_i calculée pour chaque SFF qui est un nombre flou, α_{ij} sont les coefficients des contraintes signalées par la société et b_i sont les valeurs limites des contraintes.

L'article de Herrera et Verdegay [120] explique en détail la résolution de ce type de problème considéré. Ils proposent de mettre à la place des coefficients flous se trouvant dans la fonction objectif, des valeurs non-floues obtenues pendant la phase de classement des nombres flous. C'est à dire les valeurs obtenues à l'aide de la méthode de Chu et Tsao [119] seront les coefficients de chaque SFF pour le calcul de la valeur de la fonction objectif. Ainsi le modèle sera transformé en un modèle non-flou assez facile à résoudre.

4.4.3. Sélection de SFF par l'optimisation de l'intervalle de la fonction objectif

Ishibuchi et Tanaka [121] proposent une autre méthode pour la résolution des problèmes flous avec contraintes. Comme, on l'observe dans le cas étudié les coefficients flous de la fonction objectif sont généralement des nombres flous triangulaires (a, b, c) avec b la valeur moyenne, qui peuvent être exprimés sous forme d'intervalles $[a, c]$; a est la borne inférieure (L:left), et c est la borne supérieure (R:right). Ainsi en utilisant les relations d'ordre sur les nombres flous et une fonction d'objectif sous forme d'intervalle, ils établissent un modèle de programmation multiobjectif. Pour résoudre ce problème et pour obtenir les solutions Pareto optimales une fonction objectif pondérée est construite:

$$\text{Max } wz_L(x) + (1 - w)z_C(x) \quad (4.8)$$

où

$$z_L(x) = \sum_{j \in N} a_j x_j \quad (4.9)$$

$$z_C(x) = \sum_{j \in N} b_j x_j \quad (4.10)$$

Après avoir ajouté les contraintes qui paraissent nécessaires, les solutions Pareto optimales sont obtenues en augmentant graduellement les valeurs de w entre 0 et 1.

4.5. Application de la méthodologie

Un ensemble de données de 10 alternatives de systèmes flexibles de fabrication, présenté dans le tableau 4.2 est formé aléatoirement suivant les critères choisies et en respectant les limites inférieures et supérieures prédéterminées.

Tableau 4.2: L'ensemble de données utilisé pour la régression floue (cas1)

SFF	Débit		FC	FGF	TEC	TA
	y_i	e_i				
1	1562	52	120	2	370	3.2
2	2680	144	249	5	167	10.9
3	2457	103	252	3	188	8.1
4	1765	27	139	1	349	2.4
5	2963	81	288	5	132	10.8
6	2108	162	186	2	265	6.2
7	1946	203	201	3	362	4.1
8	2221	79	173	3	319	7.4
9	2841	154	277	5	124	12.5
10	2199	216	162	4	212	4.6

Les variables indépendantes sont supposées non-floues dans le premier cas étudié. Les coefficients de l'équation de régression reliant le débit à les autres critères définies précédemment -flexibilité de capacité, flexibilité de la gamme de fabrication, travail en cours, amélioration des temps d'arrêt- sont calculés grâce au modèle de la régression linéaire floue de Tanaka (2.21). Ainsi l'équation obtenue est représentée par (4.11).

$$\begin{aligned} \text{Débit} = & 1583.100 + 2.297 * \text{FC} + 88.913 \text{ FGF} + (-1.220, 0.671) * \text{TEC} \\ & + (31.675, 16.472) * \text{TA} \end{aligned} \quad (4.11)$$

Le modèle écrit en code GAMS est donné dans l'appendice E. Les coefficients de régression sont données sous forme (α_j, c_j) où α_j est la valeur moyenne et c_j représente la valeur de dispersion des nombres flous. Quand c_j est nulle le nombre flou devient à un nombre non-flou dit "classique".

Pour le deuxième cas les variables indépendantes sont aussi considérées sous forme des nombres flous triangulaires symétriques représentés par (\bar{x}_{ij}, d_{ij}) où \bar{x}_{ij} est la moyenne et d_{ij} représente la dispersion. Les variables indépendantes dont la valeur d_{ij} est nulle sont des nombres non-flous. Les coefficients de régression linéaire floue qui sont calculés à l'aide du modèle de Sakawa et Yano (2.24) se forment l'équation (4.12).

Tableau 4.3: L'ensemble de données utilisé pour la régression floue (cas2)

SFF	Débit		FC	FGF	TEC	TA	
	y_i	e_i				\bar{x}_{ij}	d_{ij}
1	1562	52	120	2	370	3.2	0.6
2	2680	144	249	5	167	10.9	2.0
3	2457	103	252	3	188	8.1	1.1
4	1765	27	139	1	349	2.4	0.3
5	2963	81	288	5	132	10.8	1.8
6	2108	162	186	2	265	6.2	0.9
7	1946	203	201	3	362	4.1	0.7
8	2221	79	173	3	319	7.4	1.2
9	2841	154	277	5	124	12.5	1.9
10	2199	216	162	4	212	4.6	0.5

$$\begin{aligned} \text{Débit} = & 1054.545 + 4.579 * \text{FC} + 44.221 \text{ FGF} + (-0.394, 0.278) * \text{TEC} \\ & + 35.453 * \text{TA} \end{aligned} \quad (4.12)$$

La suite de la méthodologie proposée étant la même pour les deux cas cités, seuls les calculs et les résultats du premier cas sont présentés dans ce travail. Les valeurs estimées de débit sont calculées en utilisant l'équation (4.11) du premier cas sont données dans le tableau 4.4.

Tableau 4.4: Les valeurs estimées du débit

SFF	Débit	
	Moyenne de y_i^*	Dispersion de y_i^*
1	1686.526	300.9804
2	2741.136	291.6018
3	2455.891	259.5712
4	1641.536	273.7118
5	2870.251	266.4696
6	2061.253	279.9414
7	1999.764	310.4372
8	2092.435	335.9418
9	2908.592	289.104
10	2197.931	218.0232

Après avoir calculé les valeurs estimées de débit, les différences $e_i = y_i - y_i^*$ entre les valeurs observées et les valeurs estimées sont déterminées.

Tableau 4.5: Les valeurs de e_i

SFF	$e_i = y_i - y_i^*$	
	Moyenne de e_i	Dispersion de e_i
1	-124.526	352.9804
2	-61.1355	435.6018
3	1.1095	362.5712
4	123.464	300.7118
5	92.749	347.4696
6	46.747	441.9414
7	-53.7635	513.4372
8	128.565	414.9418
9	-67.5915	443.104
10	1.069	434.0232

Les valeurs de classement des valeurs e_i des SFFs sont déterminées en utilisant la méthode de Chu et Tsao.

Tableau 4.6: Les valeurs obtenues par la méthode de Chu et Tsao

SFF	Chu et Tsao
1	-62.2630
2	-30.5677
3	0.5548
4	61.7320
5	46.3745
6	23.3735
7	-26.8818
8	64.2825
9	-33.7957
10	0.5345

Supposons que la société possède un territoire limité, donc une contrainte d'espace doit être inclus dans l'analyse. Le coût étant un facteur toujours important constitue la deuxième contrainte du modèle. L'espace occupé par chaque SFF en m^2 ainsi que leurs coûts en millions de dollars sont présentés dans le tableau 4.7. L'espace libre dans l'atelier est limité à 600 m^2 et le budget consacré à cet investissement par la société est de 1.7 millions de dollars.

Tableau 4.7: L'espace occupé par chaque SFF et leurs coûts

SFF	Espace occupé (m^2)	Coût (millions de \$)
1	450	1.5
2	560	1.6
3	630	1.9
4	520	1.6
5	580	1.7
6	400	1.2
7	500	1.3
8	480	1.5
9	600	1.6
10	520	1.4

Supposons que la société désire acheter un système flexible de fabrication dont la flexibilité de la gamme de fabrication soit supérieure à 3 et dont le travail en cours soit inférieur à 300.

Un problème linéaire booléen flou (4.13), qui sélectionne le SFF optimum en maximisant les valeurs de différence sous les contraintes déterminées est formulé et résolu à l'aide de logiciel LINDO.

Maximiser (4.13)

$$- 62.2630 x_1 - 305677 x_2 + 0.5547 x_3 + 61.7320 x_4 + 46.3745 x_5 + 23.3735 x_6 - 26.8818 x_7 + 64.2825 x_8 - 33.7957 x_9 + 0.5345 x_{10}$$

sous les contraintes

$$1.5 x_1 + 1.6 x_2 + 1.9 x_3 + 1.6 x_4 + 1.7 x_5 + 1.2 x_6 + 1.3 x_7 + 1.5 x_8 + 1.6 x_9 + 1.4 x_{10} \leq 1.7$$

$$450 x_1 + 560 x_2 + 630 x_3 + 520 x_4 + 580 x_5 + 400 x_6 + 500 x_7 + 480 x_8 + 600 x_9 + 520 x_{10} \leq 600$$

$$2 x_1 + 5 x_2 + 3 x_3 + x_4 + 5 x_5 + 2 x_6 + 3 x_7 + 3 x_8 + 5 x_9 + 4 x_{10} \geq 3$$

$$370 x_1 + 167 x_2 + 188 x_3 + 349 x_4 + 132 x_5 + 265 x_6 + 362 x_7 + 319 x_8 + 124 x_9 + 212 x_{10} \leq 300$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} = 1$$

$$x_1, \dots, x_{10} \in \{0, 1\}$$

Le SFF 5 est sélectionné par le modèle utilisé. La méthode de Ishibuchi et Tanaka est appliquée aussi de la même façon. Les solutions Pareto optimales obtenues en augmentant graduellement w sont les mêmes. Le SFF 5 est encore le SFF choisi par la deuxième méthode.

5. Conclusion

Ce travail comportant des applications dans des domaines différentes, est une recherche détaillée sur la régression floue. Il engendre les informations générales sur l'analyse de régression y compris une recherche de littérature, la description de la méthode de régression linéaire floue et les études sur le problème de sélection de la technologie, plus particulièrement sur le problème de sélection de robot industriel et le problème de sélection de système flexible de fabrication qui constituent les domaines favorables à l'utilisation de la régression floue.

Les résultats obtenus à partir des modèles de décision proposés sont facilement interprétables. Le modèle basé sur la régression linéaire floue permet de choisir la même alternative ou dans certains cas celle qui est mieux par rapport à ce qui est déterminé par la régression statistique quand l'ensemble de données est capable de justifier les hypothèses de la régression statistique. Comme les performances des modèles de régression linéaire floue sont différentes, il est conseillé d'analyser le problème étudié pour effectuer un bon choix entre nombreux modèles.

Pour les systèmes qui ne sont pas bien définis ou dans lesquels l'effet humain est présent, les relations entre les variables sont en général imprécises et les données obtenues sont souvent sous forme des nombres flous. Pour ces cas-ci, la régression floue apparaît comme une approche plus adaptée à la situation car l'utilisation de la régression statistique n'est pas possible et en plus la régression floue permet de décrire l'imprécision observée.

La sélection de robot industriel et la sélection de système flexible de fabrication font partie d'un grand domaine de recherche scientifique: la sélection de la technologie. Ces deux problèmes constituent des bases formidables à l'utilisation de la régression floue. Les relations entre les critères de choix ne sont pas définies clairement, les

valeurs des données peuvent être floues à cause de la difficulté de mesure. En plus, le nombre d'alternatives est souvent limité.

Notre travail montre que la régression floue à l'aide des techniques d'optimisation floue donnent des bons résultats et permet de résoudre ces problèmes qui ne peuvent pas être traités avec la régression classique. La combinaison des modèles de régression linéaire floue avec la programmation booléenne floue et la programmation multiobjectif dans le cas des fonctions objectif sous forme d'intervalles testée pour le problème de sélection de système flexible de fabrication, apparait comme une technique efficace d'après le résultat obtenu. Les techniques différentes peuvent être utilisées à côté de l'analyse de régression linéaire floue dépendant au contexte du problème.

Les modèles de décision présentés dans ce travail peuvent être améliorés en utilisant des modèles de régression différents. Ces modèles sont assez généraux et ils peuvent être appliqués à plusieurs problèmes de décision. La forme de la fonction d'appartenance des nombres flous peut être choisi différemment selon le contexte du problème étudié.

La régression floue est utilisable dans des nombreux cas comme le déploiement de la fonction qualité (DFQ) possédant des propriétés citées ci-dessus. Actuellement les modèles de régression floue évoluent rapidement et leur utilisation constitue la source de nombreux travaux.

Bibliographie

[1] Gunst, R.F., Mason, R.L., *Regression Analysis and Its Application*, Marcel Dekker, New York, (1980).

[2] Kim, K.J., Moskowitz, H., Koksalan, M., “Fuzzy Versus Statistical Linear Regression”, *European Journal of Operational Research*, 92, 417-434, (1996).

[3] Redden, D.T., Woodall, W.H., “Properties of Certain Fuzzy Linear Regression Methods”, *Fuzzy Sets and Systems* 64, 361-375, (1994).

[4] Tanaka, H., Uejima, S., Asai, K., “Linear Regression Analysis with Fuzzy Model”, *IEEE Transactions on System, Man and Cybernetics*, SMC-12 (6), (1982).

[5] Zadeh, L.A., “Fuzzy Sets” , *Information and Control*, 8, 338-353, (1965).

[6] Khodja, L., “Contribution à la Classification Floue Non Supervisée”, Thèse de Doctorat, Université de Savoie, <http://www.univ-savoie.fr/Portail/Groupes/LISTIC/Theses/thesekhodja.pdf>

[7] Nguyen, H.T., Walker, E.A., *A First Course in Fuzzy Logic*, Chapman & Hall / CRC, (2000).

[8] Zimmermann, H.-J., *Fuzzy Sets Theory and Its Applications*, Kluwer Academic Publishers, (1996).

[9] H. Tanaka and P. Guo, *Possibilistic Data Analysis for Operations Research*, Physica-Verlag (Studies in Fuzziness and Soft Computing), (1999).

[10] Wen, C.-G., Lee, C.-S., “Development of A Cost Function for Wastewater Treatment Systems with Fuzzy Regression”, *Fuzzy Sets and Systems*, 106, 143-153, (1999).

[11] Yager, R.R., “Fuzzy Prediction Based on Regression Model”, *Information Sciences*, 26, 45-63, (1982).

[12] Bandemer, H., “Evaluating Explicit Functional Relationship from Fuzzy Observations”, *Fuzzy Sets and Systems*, 16, 41-52, (1985).

[13] Heshmaty, B., Kandel, A., “Fuzzy Linear Regression and its Applications to Forecasting in Uncertain Environment”, *Fuzzy Sets and Systems*, 15, 159-191, (1985).

[14] Gharpuray, M.M., Tanaka, H., Fan, L.T., Lai, F.S., "Fuzzy Linear Regression Analysis of Cellulose Hydrolysis", *Chem. Eng. Commun.*, 41, 299-314, (1986).

[15] Jajuga, K., "Linear Fuzzy Regression", *Fuzzy Sets and Systems*, 343-353, (1986).

[16] Celmins, A., "Multidimensional Least-Squares Fitting of Fuzzy Models", *Math. Modeling*, 9, 669-690, (1987).

[17] Celmins, A., "Least Squares Model Fitting to Fuzzy Vector Data", *Fuzzy Sets and Systems*, 22, 2455-269, (1987).

[18] Tanaka, H., "Fuzzy Data Analysis By Possibilistic Linear Models", *Fuzzy Sets and Systems*, 24, 363-375, (1987).

[19] Chen, S.-Q., "Analysis for Multiple Fuzzy Regression", *Fuzzy Sets and Systems*, 25, 59-65, (1988).

[20] Diamond, P., "Fuzzy Least Squares", *Information Sciences*, 46, 141-157, (1988).

[21] Tanaka, H., Watada, J., "Possibilistic Linear Systems and Their Application to the Linear Regression Model", *Fuzzy Sets and Systems*, 27, 275-289, (1988).

[22] Sakawa, M., Yano, H., "Fuzzy Linear Regression and Its Application to The Sales Forecasting", *Internat. J. Policy Inform.*, 13, 111-125, (1989).

[23] Tanaka, H., Hayashi, I., Watada, J., "Possibilistic Linear Regression Analysis for Fuzzy Data", *European Journal of Operational Research*, 40, 389-396, (1989).

[24] Bardossy, A., "Note On Fuzzy Regression", *Fuzzy Sets and Systems*, 37, 65-75, (1990).

[25] Bardossy, A., Bogardi, I., Duckstein, L., "Fuzzy Regression in Hydrology", *Water Resources Res.*, 26, 1497-1508, (1990).

[26] Hayashi, I., Tanaka, H., "The Fuzzy GMDH Algorithm by Possibility Models and Its Application", *Fuzzy Sets and Systems*, 36, 245-258, (1990).

[27] Nather, W., Albrecht, M., "Linear Regression with Random Fuzzy Observations", *Statist.*, 21, 521-531, (1990).

[28] Oh, S.-B., Kim, W., Lee, J.K., "An Approach to Causal Modeling in Fuzzy Environment and Its Application", *Fuzzy Sets and Systems*, 35, 43-55, (1990).

[29] Savic, D.A., Pedrycz, W., "Evaluation of Fuzzy Linear Regression Models", *Fuzzy Sets and Systems*, 39, 51-63, (1991).

[30] Tanaka, H., Ishibuchi, H., "Identification of Possibilistic Linear Systems by Quadratic Membership Functions", *Fuzzy Sets and Systems*, 41, 145-160, (1991).

[31] Ishibuchi, H., Tanaka, H., "Fuzzy Regression Analysis Using Neural Networks", *Fuzzy Sets and Systems*, 50, 257-365, (1992).

[32] Jozsef, S., "On The Effect of Linear Data Transformations in Possibilistic Fuzzy Linear Regression", *Fuzzy Sets and Systems*, 45, 185-188, (1992).

[33] Sakawa, M., Yano, H., "Multiobjective Fuzzy Linear Regression Analysis for Fuzzy Input-Output Data", *Fuzzy Sets and Systems*, 47, 173-181, (1992).

[34] Wang, X., Ha, M., "Fuzzy Linear Regression Analysis", *Fuzzy Sets and Systems*, 51, 179-188, (1992).

[35] Bardossy, A., Bogardi, I., Duckstein, L., "Fuzzy Nonlinear Regression Analysis of Dose-Response Relationships", *European Journal of Operations Research*, 66, 36-51, (1993).

[36] Ishibuchi, I., Tanaka, H., Okada, H., "An Architecture of Neural Networks with Interval Weights and Its Application to Fuzzy Regression Analysis", *Fuzzy Sets and Systems*, 57, 27-39, (1993).

[37] Lu, F., Xu, X.G., "A Forecasting Model of Fuzzy Self-Regression", *Fuzzy Sets and Systems*, 89, 181-192, (1993).

[38] Moskowitz, H., Kim, K., "On Assessing the H Value in Fuzzy Linear Regression", *Fuzzy Sets and Systems* 58, 303-327, (1993).

[39] Chang, P.T., Lee, E.S., "Fuzzy Linear Regression with Spread Unrestricted In Sign", *Comput. Math. Appl.*, 28, 61-70, (1994).

[40] Chang, P.T., Lee, E.S., "Fuzzy Least Absolute Deviations Regression and The Conflicting Trends In Fuzzy Parameters", *Comput. Math. Appl.*, 28, 89-101, (1994).

[41] Peters, G., "Fuzzy Linear Regression With Fuzzy Intervals", *Fuzzy Sets and Systems*, 63, 45-55, (1994).

[42] Tanaka, H., Ishibuchi, H., Yoshikawa S., "Exponential Possibility Regression Analysis", *Fuzzy Sets and Systems*, 69, 305-318, (1995).

[43] Chang, P.T., Lee, E.S., Konz, S.A., "Applying Fuzzy Linear Regression to VDT Legibility", *Fuzzy Sets and Systems*, 80, 197-204, (1996).

- [44] Chang, P.-T., Lee, E.S., "A Generalized Fuzzy Weighted Least-Squares Regression", *Fuzzy Sets and Systems*, 82 (3), 289-298, (1996).
- [45] Redden, D.T., Woodall W.H., "Further Examination of Fuzzy Linear Regression", *Fuzzy Sets and Systems*, 79, 203-211, (1996).
- [46] Chang, P.-T., "Fuzzy Seasonality Forecasting", *Fuzzy Sets and Systems*, 90 (1), 1-10, (1997).
- [47] Kim, B., Bishu, R.R., "Evaluation of Fuzzy Linear Regression Models by Comparing Membership Functions" *Fuzzy Sets and Systems*, 100 (1-3), 343-352, (1998).
- [48] Chen, Y-S., "Fuzzy Ranking and Quadratic Fuzzy Regression", *Computers & Mathematics with Applications*, 38 (11-12), 265-279, (1999).
- [49] Cheng, C.-B., Lee, E.S., "Applying Fuzzy Adaptive Network to Fuzzy Regression Analysis", *Computers & Mathematics with Applications*, 38 (2), 123-140, (1999).
- [50] Yen, K.K., Ghoshray, S., Roig, G., "A Linear Regression Model Using Triangular Fuzzy Number Coefficients", *Fuzzy Sets and Systems*, 106 (2, 1), 167-177, (1999).
- [51] Buckley J.J., Feuring T., "Linear and Non-Linear Fuzzy Regression: Evolutionary Algorithm Solutions", *Fuzzy Sets and Systems*, 112 (3), 381-394, (2000).
- [52] Duniak, J.P., Wunsch, D., "Fuzzy Regression by Fuzzy Number Neural Networks", *Fuzzy Sets and Systems*, 112 (3), 371-380, (2000).
- [53] D'Urso, P., Gastaldi, T., "A Least-Squares Approach to Fuzzy Linear Regression Analysis", *Computational Statistics & Data Analysis*, 34 (4), 427-440, (2000).
- [54] Kim, K-J, Moskowitz, H., Dhingra, A., Evans, G., "Fuzzy Multicriteria Models for Quality Function Deployment", *European Journal of Operational Research*, 121 (3), 504-518, (2000).
- [55] Özelkan, E.C., Duckstein, L., "Multi-Objective Fuzzy Regression: A General Framework", *Computers & Operations Research*, 27, 635-652, (2000).
- [56] Wang, H-F., Tsaur, R-C., "Bicriteria Variable Selection in A Fuzzy Regression Equation", *Computers & Mathematics with Applications*, 40 (6-7), 877-883, (2000).
- [57] Wang, H-F., Tsaur, R-C., "Resolution of Fuzzy Regression Model", *European Journal of Operational Research*, 126, 637-650, (2000).

- [58] Chang, Y-H.O., "Hybrid Fuzzy Least-Squares Regression Analysis and Its Reliability Measures", *Fuzzy Sets and Systems*, 119 (2), 225-246, (2001).
- [59] Chang, Y., Ayyub, B.M., "Fuzzy Regression Methods - A Comparative Assessment", *Fuzzy Sets and Systems*, 119, 187-203, (2001).
- [60] Chen, Y-S., "Outliers Detection and Confidence Interval Modification in Fuzzy Regression", *Fuzzy Sets and Systems*, 119, 259-272, (2001).
- [61] Cheng, C-B., Lee, E.S., "Fuzzy Regression with Radial Basis Network", *Fuzzy Sets and Systems*, 119(2), 291-301, (2001).
- [62] Ishibuchi, H., Nii, M., "Fuzzy Regression Using Asymmetric Fuzzy Coefficients and Fuzzified Neural Networks", *Fuzzy Sets and Systems*, 119(2), 273-290, (2001).
- [63] Lee, H.T., Chen, S.H., "Fuzzy Regression Model with Fuzzy Input and Output Data for Manpower Forecasting", *Fuzzy Sets and Systems*, 119 (2), 205-213, (2001).
- [64] Özelkan, E.C., Duckstein, L., "Fuzzy Conceptual Rainfall-Runoff Models", *Journal of Hydrology*, 253 (1-4), 41-68, (2001).
- [65] Tseng, F-M., Tzeng, G-H., Yu, H-C., Yuan, B.J.C., "Fuzzy ARIMA Model for Forecasting The Foreign Exchange Market", *Fuzzy Sets and Systems*, 118 (1), 9-19, (2001).
- [66] Yu, J-R., Tzeng, G-H., Li, H-L., "General Fuzzy Piecewise Regression Analysis with Automatic Change-Point Detection", *Fuzzy Sets and Systems*, 119 (2), 247-257, (2001).
- [67] Chang, P-T., "Fuzzy Stage Characteristic-Preserving Product Life Cycle Modeling", *Fuzzy Sets and Systems*, 126 (1), 33-47, (2002).
- [68] Cron, G., Dubuisson, B., "Pattern Characteristics of An Evolution Between Two Classes", *Fuzzy Sets and Systems*, 126 (3), 293-310, (2002).
- [69] Kahraman, C., "An Application of Fuzzy Linear Regression to The Information Technology in Turkey", *International Journal of Technology Management*, 24, 330-339, (2002).
- [70] Kao, C., Chyu, C-L., "A Fuzzy Linear Regression Model with Better Explanatory Power", *Fuzzy Sets and Systems*, 126 (3), 401-409, (2002).
- [71] Tran, L., Duckstein, L., "Multiobjective Fuzzy Regression with Central Tendency and Possibilistic Properties", *Fuzzy Sets and Systems*, 130, 21-31, (2002).

[72] Tran, L.T., Ridgley, M.A., Duckstein, L., Sutherland, R., "Application of Fuzzy-Logic Based Modeling to Improve The Performance of The Revised Universal Soil Loss Equation", *Catena*, 47 (3), 203-226, (2002).

[73] Tseng, F-M., Tzeng, G-H., "A Fuzzy Seasonal ARIMA Model for Forecasting", *Fuzzy Sets and Systems*, 126 (3), 367-376, (2002).

[74] Wu, B., Tseng, N-F., "A New Approach to Fuzzy Regression Models with Application to Business Cycle Analysis", *Fuzzy Sets and Systems*, 130, 33-42, (2002).

[75] Wünsche, A., Nather, W., "Least-Squares Fuzzy Regression with Fuzzy Random Variables", *Fuzzy Sets and Systems*, 130, 43-50, (2002).

[76] Yang, M-S., Lin, T-S., "Fuzzy Least-Squares Linear Regression Analysis for Fuzzy Input-Output Data", *Fuzzy Sets and Systems*, 126 (3), 389-399, (2002).

[77] D'Urso, P., "Linear Regression Analysis for Fuzzy/Crisp Input and Fuzzy/Crisp Output Data", *Computational Statistics & Data Analysis*, 42 (1-2), 47-72, (2003).

[78] Hong, D.H., Hwang, C., "Support Vector Fuzzy Regression Machines", *Fuzzy Sets and Systems*, 138, 271-281, (2003).

[79] Ip, K.W., Kwong, C.K., Wong, Y.W., "Fuzzy Regression Approach to Modelling Transfer Moulding for Microchip Encapsulation", *Journal of Materials Processing Technology*, 140 (1-3), 147-151, (2003).

[80] Kao, C., Chyu, C-L., "Least-Squares Estimates in Fuzzy Regression Analysis", *European Journal of Operational Research*, 148, 426-435, (2003).

[81] Soliman, S.A., Alammari, R.A., El-Hawary, M.E., "Frequency and Harmonics Evaluation in Power Networks Using Fuzzy Regression Technique", *Electric Power Systems Research*, 66 (2), 171-177, (2003).

[82] Tsauro, R-C., "Extrapolating Internet Users in Taiwan by Risk Assessment", *Computers & Mathematics with Applications*, 46 (10-11), 1725-1734, (2003).

[83] Yang, M-S., Liu, H-H., "Fuzzy Least-Squares Algorithms for Interactive Fuzzy Linear Regression Models", *Fuzzy Sets and Systems*, 135 (2), 305-316, (2003).

[84] Al-Kandari, A.M., Soliman, S.A., El-Hawary, M.E., "Fuzzy Short-Term Electric Load Forecasting", *International Journal of Electrical Power & Energy Systems*, 26 (2), 111-122, (2004).

[85] Chaves, P., Tsukatani, T., Kojiri, T., "Operation of Storage Reservoir for Water Quality by Using Optimization and Artificial Intelligence Techniques", *Mathematics and Computers in Simulation*, 67, 419-432, (2004).

[86] Hong, D.H., Hwang, C., "Extended Fuzzy Regression Models Using Regularization Methods", *Information Sciences*, 164, 31-46, (2004).

[87] Hong, D.H., Hwang, C., Ahn, C., "Ridge Estimation for Regression Models with Crisp Inputs and Gaussian Fuzzy Output", *Fuzzy Sets and Systems*, 142, 307-319, (2004).

[88] Li, C., Liao, X., Yu, J., "Tabu Search for Fuzzy Optimization and Applications", *Informations Sciences*, 158, 3-13, (2004).

[89] Nasrabadi, M.M., Nasrabadi, E., "A Mathematical-Programming Approach to Fuzzy Linear Regression Analysis", *Applied Mathematics and Computation*, 155 (3), 873-881, (2004).

[90] Sanchez, J., Gomez, A.T., "Estimating A Fuzzy Term Structure of Interest Rates Using Fuzzy Regression Techniques", *European Journal of Operational Research*, 154 (3), 804-818, (2004).

[91] Uddameri, V., Kuchanur, M., "Fuzzy QSARs for Predicting $\log K_{oc}$ of Persistent Organic Pollutants", *Chemosphere*, 54 (6), 771-776, (2004).

[92] Chang, P-T., "Fuzzy Strategic Replacement Analysis", *European Journal of Operational Research*, 160 (2), 532-559, (2005).

[93] Hojati, M., Bector, C.R., Smimou, K., "A Simple Method for Computation of Fuzzy Linear Regression", *European Journal of Operational Research*, 166 (1), 172-184, (2005).

[94] Nasrabadi, M.M., Nasrabadi, E., Nasrabad, A.R., "Fuzzy Linear Regression Analysis: A Multi-Objective Programming Approach", *Applied Mathematics and Computation*, 163 (1), 245-251, (2005).

[95] Tseng, F-M., Lin, L., "A Quadratic Interval Logit Model for Forecasting Bankruptcy", *Omega*, 33, 85-91, (2005).

[96] Carroll Janer, L.F., "Vers La Maîtrise du Développement d'Un Contrôleur Temps Réel Sûr de Fonctionnement pour Les Robots Manipulateurs", Thèse de Doctorat, L'Institut National des Sciences Appliquées de Toulouse, http://beru.univ-brest.fr/~singhoff/DOC/PAPIER_A_TRIER/these-luis-carroll.pdf.

[97] Khouja, M.J., Kumar, R.L., "An Options View of Robot Performance Parameters in A Dynamic Environment", *International Journal of Production Research*, 37 (6), 1243-1257, (1999).

[98] Khouja, M., Offodile, O.F., "The Industrial Robots Selection Problem: Literature Review and Directions for Future Research", *IIE Transactions*, 26 (4), 50-61, (1994).

[99] Karsak, E.E., "A Two-Phase Robot Selection Procedure", *Production Planning & Control*, 9 (7), 675-684, (1998).

[100] Imany, M.M., Schlesinger, R.J., "Decision Models for Robot Selection: A Comparison of Ordinary Least Squares and Linear Goal Programming Methods", *Decision Sciences*, 20, 40-53, (1989).

[101] Booth, D.E., Khouja, M., Hu, M., "Robust Multivariate Statistical Procedure for Evaluation and Selection of Industrial Robots", *International Journal of Operations & Production Management*, 12 (2), 15-24, (1992).

[102] Khouja, M., Booth, D.E., "A Decision Model for The Robot Selection Problem Using Robust Regression", *Decision Sciences*, 22 (3), (1991).

[103] Sarkis, J., Talluri, S., "A Decision Model for Evaluation of Flexible Manufacturing Systems in The Presence of Both Cardinal and Ordinal Factors", *International Journal of Production Research*, 37 (13), 2927-2938, (1999).

[104] Sarkis, J., "Evaluating Flexible Manufacturing Systems Alternatives Using Data Envelopment Analysis", *The Engineering Economist*, 43 (1), 25-47, (1997).

[105] Karsak, E.E., Kuzgunkaya, O., "A Fuzzy Multiple Objective Programming Approach for The Selection of A Flexible Manufacturing System", *International Journal of Production Economics*, 79 (2), 101-111, (2002).

[106] Shang, J., Sueyoshi, T., "A Unified Framework for The Selection of A Flexible Manufacturing System", *European Journal of Operational Research*, 85, 297-315, (1995).

[107] Proctor, M.D., Canada, J.R., "Past and Present Methods of Manufacturing Investment Evaluation: A Review of The Empirical and Theoretical Literature", *The Engineering Economist*, 38 (1), 45-58, (1992).

[108] Son, Y.K., "A Comprehensive Bibliography on Justification of Advanced Manufacturing Technologies", *The Engineering Economist*, 38 (1), 59-71, (1992).

[109] Bortolan, G., Degani, R., "A Review of Some Methods for Ranking Fuzzy Subsets", *Fuzzy Sets and Systems*, 15, 1-19, (1985).

[110] Yager, R.R., "Ranking Fuzzy Subsets Over The Unit Interval", *Proc. CDC*, 1435-1437, (1978).

[111] Chang, W., "Ranking of Fuzzy Utilities with Triangular Membership Functions", *Proc. Int. Conf. On Policy Anal. And Inf. Systems*, 263-272, (1981).

[112] Adamo, J.M., "Fuzzy Decision Trees", *Fuzzy Sets and Systems*, 4, 207-219, (1980).

- [113] Baas, S.M., Kwakernaak, H., "Rating and Ranking of Multiple-Aspect Alternatives Using Fuzzy Sets", *Automatica*, 13, 47-58, (1977).
- [114] Baldwin, J.F., Guild, N.C.F., "Comparaison of Fuzzy Sets on The Same Decision Space", *Fuzzy Sets and Systems*, 2, 213-233, (1979).
- [115] Jain, R., "Decision-Making in The Presence of Fuzzy Variables", *IEEE Trans. Systems Man Cybernet.*, 6, 698-703, (1976).
- [116] Jain, R. "A Procedure FOR Multiple-Aspect Decision -Making Using Fuzzy Sets, *Internat. J. Systems Sci.*, 8, 1-7, (1977).
- [117] Dubois, D., Prade, H., "Ranking of Fuzzy Numbers in the Setting of Possibility Theory", *Information Sciences*, 30, 183-224, (1983).
- [118] Yao, J.-S., Wu, K., "Ranking Fuzzy Numbers Based on Decomposition Principle and Signed Distance", *Fuzzy Sets and Systems*, 116, 275-288, (2000).
- [119] Chu, T.-C., Tsao, C.-T., "Ranking Fuzzy Numbers with an Area Between the Centroid Point and Original Point", *Computers and Mathematics with Applications*, 43, 11-117, (2002).
- [120] Herrera, F., Verdegay, J.L., "Fuzzy Boolean Programming Problems with Fuzzy Costs", *Fuzzy Sets and Systems*, 81, 57-76, (1996).
- [121] Ishibuchi, H., Tanaka, H., "Optimization of the Interval Objective Function", *European Journal of Operational Research*, 48, 219-225, (1990).

Appendices

Appendice A: Le modèle de Tanaka écrit en code GAMS

Sets

i competitors / 1*27 /
j independent variables / 0*2 / ;

Parameters

y(i)
/
1 -1.897119985
2 -2.995732274
3 0.2390169
4 -3.688879454
5 -1.386294361
6 -2.302585093
7 -2.302585093
8 -2.302585093
9 -1.609437912
10 -2.995732274
11 -0.693147181
12 0
13 -2.995732274
14 -2.995732274
15 0
16 0
17 0.693147181
18 -1.609437912
19 -2.995732274
20 0.693147181
21 0.693147181
22 -2.995732274
23 -2.302585093
24 -2.302585093
25 -1.609437912
26 0.2390169
27 0.708035793 / ;

Table x(i,j)

	0	1	2
1	1	1.35	4.094344562
2	1	1.1	1.791759469

3	1	1.27	3.80666249
4	1	0.66	0.405465108
5	1	0.05	3.912023005
6	1	0.3	0
7	1	1	1.609437912
8	1	1	2.708050201
9	1	1.1	2.302585093
10	1	1	1.791759469
11	1	0.9	3.401197382
12	1	0.15	2.610069793
13	1	1.2	2.302585093
14	1	1.2	3.401197382
15	1	1	3.850147602
16	1	1	4.382026635
17	1	2	2.708050201
18	1	1	2.302585093
19	1	0.3	2.302585093
20	1	0.8	0.405465108
21	1	1.7	3.295836866
22	1	1	-0.105360516
23	1	0.5	0.916290732
24	1	0.5	0.916290732
25	1	1	2.302585093
26	1	1.25	4.248495242
27	1	0.75	5.323009979

Scalar h /0.5/ ;

Variables

c(j) spread of fuzzy parameter j

m(j) mean of fuzzy parameter j

z sum of the spreads

Positive Variable c ;

Equations

spread objective function

constraint1(i)

constraint2(i) ;

spread .. z =e= sum(j ,c(j)* sum (i, abs(x(i,j)))) ;

constraint1(i) .. sum(j, (m(j)*x(i,j) + h*c(j)*abs(x(i,j)))) =g= y(i) ;

constraint2(i) .. sum(j, (m(j)*x(i,j) - h*c(j)*abs(x(i,j)))) =l= y(i) ;

Model fuzzyreg /all/ ;

Solve fuzzyreg using lp minimizing z ;

Display c.l, m.l ;

Appendice B: Le modèle de Peters écrit en code GAMS**Sets**

```
i / 1*27 /
j / 0*2 /;
```

Parameter**y(i)**

/

```
1 -1.897119985
2 -2.995732274
3 0.2390169
4 -3.688879454
5 -1.386294361
6 -2.302585093
7 -2.302585093
8 -2.302585093
9 -1.609437912
10 -2.995732274
11 -0.693147181
12 0
13 -2.995732274
14 -2.995732274
15 0
16 0
17 0.693147181
18 -1.609437912
19 -2.995732274
20 0.693147181
21 0.693147181
22 -2.995732274
23 -2.302585093
24 -2.302585093
25 -1.609437912
26 0.2390169
27 0.708035793 /;
```

Parameter**p(i)**

/

```
1 1
2 1
3 1
4 1
5 1
6 1
7 1
8 1
```

9 1
 10 1
 11 1
 12 1
 13 1
 14 1
 15 1
 16 1
 17 1
 18 1
 19 1
 20 1
 21 1
 22 1
 23 1
 24 1
 25 1
 26 1
 27 1/;

Table x(i,j)

	0	1	2
1	1	1.35	4.094344562
2	1	1.1	1.791759469
3	1	1.27	3.80666249
4	1	0.66	0.405465108
5	1	0.05	3.912023005
6	1	0.3	0
7	1	1	1.609437912
8	1	1	2.708050201
9	1	1.1	2.302585093
10	1	1	1.791759469
11	1	0.9	3.401197382
12	1	0.15	2.610069793
13	1	1.2	2.302585093
14	1	1.2	3.401197382
15	1	1	3.850147602
16	1	1	4.382026635
17	1	2	2.708050201
18	1	1	2.302585093
19	1	0.3	2.302585093
20	1	0.8	0.405465108
21	1	1.7	3.295836866
22	1	1	-0.105360516
23	1	0.5	0.916290732
24	1	0.5	0.916290732
25	1	1	2.302585093


```

26  1  1.25  4.248495242
27  1  0.75  5.323009979

```

Scalar D /0.037/ ;

Scalar T /1000/ ;

Variables

```

a(i)
c(j) spread of fuzzy parameter j
m(j) mean of fuzzy parameter j
z

```

Positive Variable c ;

a.up(i) = 1 ;

a.lo(i) = 0;

Equations

spread objective function

constraint1(i)

constraint2(i)

constraint3 ;

spread .. z =e= sum(i, a(i)*D) ;

constraint1(i) .. sum(j, (m(j) + c(j))*x(i,j)) =g= (y(i) - p(i)*(1 - a(i))) ;

constraint2(i) .. sum(j, (- m(j) + c(j))*x(i,j)) =g= (- (1 - a(i))*p(i) - y(i)) ;

constraint3 .. - sum((i,j), c(j)*x(i,j)) =g= - (1 - z)*T

Model fuzzyreg /all/ ;

Solve fuzzyreg using lp maximizing z ;

Display c.l, m.l, a.l ;

Appendice C: Les valeurs limites de la distribution du coefficient de corrélation du rang de Spearman

n	α			
	0.050	0.025	0.010	0.005
5	0.900	-	-	-
6	0.829	0.886	0.943	-
7	0.714	0.786	0.893	-
8	0.643	0.738	0.833	0.881
9	0.600	0.683	0.783	0.833
10	0.564	0.648	0.745	0.794
11	0.523	0.623	0.736	0.818
12	0.497	0.591	0.703	0.780
13	0.475	0.566	0.673	0.745
14	0.457	0.545	0.646	0.716
15	0.441	0.525	0.623	0.689
16	0.425	0.507	0.601	0.666
17	0.412	0.490	0.582	0.645
18	0.399	0.476	0.564	0.625
19	0.388	0.462	0.549	0.608
20	0.377	0.450	0.534	0.591
21	0.368	0.438	0.521	0.576
22	0.359	0.428	0.508	0.562
23	0.351	0.418	0.496	0.549
24	0.343	0.409	0.485	0.537
25	0.336	0.400	0.475	0.526
26	0.329	0.329	0.465	0.515
27	0.323	0.385	0.456	0.505
28	0.317	0.377	0.448	0.496
29	0.311	0.370	0.440	0.487
30	0.305	0.364	0.432	0.478

Source: Olds, E.G., "Distribution of sums squares of rank differences for small samples", *Annals of Mathematical Statistics*, 9, (1938).

Appendice D: Nouveau modèle de Peters écrit en code GAMS

Sets

```
i / 1*27 /  
j / 0*3 /;
```

Parameter

```
y(i)  
/  
1 0.15  
2 0.05  
3 1.27  
4 0.025  
5 0.25  
6 0.1  
7 0.1  
8 0.1  
9 0.2  
10 0.05  
11 0.5  
12 1  
13 0.05  
14 0.05  
15 1  
16 1  
17 2  
18 0.2  
19 0.05  
20 2  
21 2  
22 0.05  
23 0.1  
24 0.1  
25 0.2  
26 1.27  
27 2.03/;
```

Parameter

```
p(i)  
/  
1 1  
2 1  
3 1  
4 1  
5 1  
6 1  
7 1  
8 1
```

9 1
 10 1
 11 1
 12 1
 13 1
 14 1
 15 1
 16 1
 17 1
 18 1
 19 1
 20 1
 21 1
 22 1
 23 1
 24 1
 25 1
 26 1
 27 1/;

Table x(i,j)

	0	1	2	3
1	1	1.35	60	7.2
2	1	1.1	6	4.8
3	1	1.27	45	5
4	1	0.66	1.5	7.2
5	1	0.05	50	9.6
6	1	0.3	1	1.07
7	1	1	5	1.76
8	1	1	15	3.2
9	1	1.1	10	6.72
10	1	1	6	2.4
11	1	0.9	30	2.88
12	1	0.15	13.6	6.9
13	1	1.2	10	3.2
14	1	1.2	30	4
15	1	1	47	3.68
16	1	1	80	6.88
17	1	2	15	8
18	1	1	10	6.3
19	1	0.3	10	0.94
20	1	0.8	1.5	0.16
21	1	1.7	27	2.81
22	1	1	0.9	3.8
23	1	0.5	2.5	1.25
24	1	0.5	2.5	1.37
25	1	1	10	3.63

```

26  1  1.25  70  5.3
27  1  0.75  205  4

```

Scalar D /0.037/ ;

Scalar T /1000/ ;

Variables

a(i)

c(j) spread of fuzzy parameter j

m(j) mean of fuzzy parameter j

z

Positive Variable c ;

a.up(i) = 1 ;

Equations

spread objective function

constraint1(i)

constraint2(i)

constraint3 ;

spread .. z =e= sum(i, a(i)*D) ;

constraint1(i) .. sum(j, (m(j) + c(j))*abs(x(i,j))) =g= (y(i) - p(i)*(1 - a(i))) ;

constraint2(i) .. sum(j, (- m(j) + c(j))*abs(x(i,j))) =g= (- (1 - a(i))*p(i) - y(i)) ;

constraint3 .. - sum((i,j), c(j)*abs(x(i,j))) =g= - (1 - z)*T

Model fuzzyreg /all/ ;

Solve fuzzyreg using lp maximizing z ;

Display c.l, m.l, a.l ;

Appendice E: Le modèle de Tanaka pour la sélection de SFF écrit en code GAMS**Sets**

```
i / 1*10 /
j / 0, VF, RF, WIP, DT / ;
```

Parameter

```
y(i)/
1 1562
2 2680
3 2457
4 1765
5 2963
6 2108
7 1946
8 2221
9 2841
10 2199 /;
```

Parameter

```
e(i)/
1 52
2 144
3 103
4 27
5 81
6 162
7 203
8 79
9 154
10 216 /;
```

Table x(i,j)

	0	VF	RF	WIP	DT
1	1	120	2	370	3.2
2	1	249	5	167	10.9
3	1	252	3	188	8.1
4	1	139	1	349	2.4
5	1	288	5	132	10.8
6	1	186	2	265	6.2
7	1	201	3	362	4.1
8	1	173	3	319	7.4
9	1	277	5	124	12.5
10	1	162	4	212	4.6

```
Scalar h /0.5/ ;
```

Positive Variable c ;

Variables

c(j) spread of fuzzy parameter j

m(j) mean of fuzzy parameter j

z sum of the spreads

Equations

spread objective function

constraint1(i)

constraint2(i) ;

spread .. $z = e = \sum((i,j), c(j)*x(i,j)) ;$

constraint1(i) .. $\sum(j, (m(j) + h*c(j))*x(i,j)) = g = y(i) + h*e(i) ;$

constraint2(i) .. $\sum(j, (m(j) - h*c(j))*x(i,j)) = l = y(i) - h*e(i) ;$

Model fuzzyreg /all/ ;

Solve fuzzyreg using lp minimizing z ;

Display c.l, m.l

Biographie

Zeynep YILMAZ est née le 4 septembre 1980 à İstanbul. Elle a fait ses études de collège et de lycée au Lycée de Galatasaray d'où elle est diplômée en 1999, en étant le troisième de sa promotion d'après les moyennes scolaires.

Lors de la dernière année de ses études de génie industriel à la Faculté d'Ingénierie et de Technologie de l'Université Galatasaray, elle a passé un semestre en France à l'ENSGI (Ecole Nationale Supérieure de Génie Industriel) de l'INPG (Institut National Polytechnique de Grenoble) comme étudiante en échange dans le cadre du programme Erasmus.

En 2003, elle a obtenu son grade de licence de Génie Industriel avec le succès d'être la première parmi les élèves de son département. Depuis décembre 2004, elle est assistante au sein de la Faculté d'Ingénierie et de Technologie de l'Université Galatasaray, au département de Génie Industriel. Actuellement elle fait ses études de maîtrise dans le domaine de Génie Industriel, à l'Institut de Science de l'Université Galatasaray.

Zeynep YILMAZ continue ses travaux avec Prof. Dr. E. Ertuğrul KARSAK. Elle est co-auteur de l'article intitulé "Bulanık Doğrusal Regresyon Tekniğinin Robot Seçimine Uygulanması" qu'elle l'a aussi présenté au 4^{ème} Congrès de Statistique en Mai 2005. Elle est co-auteur d'un deuxième article nommé "Fuzzy Linear Regression-Based Decision Model for Robot Selection" qui va être présenté au 35^{ème} Congrès International d'Informatique et de Génie Industriel (International Conference on Computers and Industrial Engineering) en Juin 2005.