Modeling, Control and Optimization Of Cascade Hydroelectric-Irrigation Plants : Operation and Planning

Bassam Bou-Fakhreddine

Université Libanaise

le cnam

École doctorale Informatique, Télécommunications et Électronique (Paris)

Centre d'Études et de Recherche en Informatique et Communications
et
École Doctorale des Sciences et Technologie (Beirut)

THÈSE DE DOCTORAT

présentée par : **Bassam BOU-FAKHREDDINE**

soutenue le : **15 Mai 2018**

pour obtenir le grade de : **Docteur du Conservatoire National des Arts et Métiers et de Université Libanaise**

Discipline : **Génie informatique, automatique et traitement du signal**
Spécialité : **Informatique**

Modeling, Control and Optimization of Cascade Hydroelectric-Irrigation Plants

Operation and Planning

THÈSE dirigée par

M. FAYE Alain	*Maître de Conférences - HDR, Ecole Nationale Superieure d'Informatique pour l'Industrie et l'Entreprise*
M. MOUGHARBEL Imad	*Professeur des Universités, Université Libanaise*
M. POLLET Yann	*Professeur des Universités, Conservatoire National des Arts et Métiers*
Mme. BU SHAKRA Sara	*Maître de Conférences, Université Libanaise*

RAPPORTEURS

M. MAUTOR Thierry	*Professeur des Universités, Université de Versailles*
M. GHAZIRY Hassan	*Chargé de Recherche, Beirut Research and Innovation Center*

PRÉSIDENT

M. HAMAM Yskandar	*Professeur des Universités, Tshwane University of Technology*

EXAMINATEURS

M. HAMAM Yskandar	*Professeur des Universités, Tshwane University of Technology*
M. DOSSANTOS Pierre	*Chargé de Recherche, Centre Bruyères-Le-Chatel*

I would like to dedicate this thesis to my loving parents ...

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains around 56,500 words including appendices, bibliography, footnotes, tables and equations and has fewer than 60 figures.

Bassam BOU-FAKHREDDINE

Acknowledgements

Completion of this PhD. dissertation was possible with the support of several people. I would like to express my sincere gratitude to all of them. First of all, I am extremely grateful to my Research Director, Dr. Imad MOUGHARBEL, Professor, Faculty of Engineering - Lebanese University, for his valuable guidance, scholarly inputs and consistent encouragement I received throughout the research work. Dr. Imad is a person with a pleasant and positive disposition. He has always made himself available to clarify my doubts despite his busy schedules and I consider it as a great opportunity to do my doctoral program under his guidance. I would like also to thank my second Research Director Dr. Alain FAYE, HDR, Ecole Nationale Superieure d'Informatique pour l'Industrie et l'Entreprise, for his academic support and I was very much privileged to learn from his research expertise. I owe a lot to him for this achievement.

Furthermore, the dissertation would not have come to a successful completion without the help of my Advisor Dr. Sara Abou Chakra, Associated Professor, IUT - Lebanese University. I express my gratitude to her for being very encouraging and for the moral support of the cotutelle convention between the Lebanese University and CNAM of Paris . I am also very grateful to my second Advisor Dr. Yann POLLET, Professor, Conservatoire National des Arts et Métiers, for helping me to carry out the research work at the CNAM and for his support at various phases of the doctoral program.

Some faculty members at CNAM have been very kind enough to extend their help at various stages of this research, whenever I needed them, and I do hereby acknowledge all of them. Special thanks also go to Dr. Marie-Christine Costa for organizing a meeting to discuss my research findings and for her suggestions about a future work.

The doctoral research was financially supported by several institutions, including CNRS the National Council for Scientific Research - Lebanon, Lebanese University and CNAM de Paris.

I acknowledge Dr. Roula Bachour, for providing me with some important information regarding irrigation requirements. Thanks for her kind concern and good wishes.

I would like to acknowledge the people in charge of Water Resources Department - Litani River Authority (LRA) and Agricultural Research Institute of Lebanon (LARI) for providing us with the available hydrological and meteorological data.

Finally, I owe a lot to my parents, who encouraged and supported me at every stage of my personal and academic life, and longed to see this achievement come true.

Above all, I owe it all to Almighty God for granting me the wisdom, health and strength to undertake this research work and enabling me to reach its final phase.

Abstract

The aim of this research consists in an operational optimization of cascade hydropower plants. This general goal relies on the optimization of resources management in order to reach the best practices in hydropower and irrigation, taking into account relevant environmental constraints. The main challenge is here to find the most realistic model based on the characteristics of water resource, energy demand and irrigation profile. To meet this objective, methods integrating data exploration and mathematical programming will be the foundations for a new Decision Support Tool. The research deals with three interconnected topics which are the following ones: hydrological modeling, irrigation models and cascade hydropower operation planning. In order to test the proposed methodology, each model has been applied in the Litani project - Lebanon. The aim of hydrological modeling is to predict stream flows and lake surface evaporation. However, to establish these models, several methods standing at the intersection of machine learning, statistics, and database management systems are used. The proposed forecasting approaches are set to deal with newly addressed situation where hydro-meteorological data suffer from scarcity, non-homogeneity and asymmetry (unbalanced). In fact, the key outputs in this research are data-driven models based on some of the most wide-spread statistical data-learning techniques (Auto-regressive methods, Fuzzy inference, Least Square Method and hybrid modeling). Despite the poor quality and quantity of data retrieved from Litani river and Qararoun dam, the research managed to process in an efficient manner the available data and managed to exhibit good forecasting results. The second topic of interest in the cascade hydropower-irrigation systems relies on agriculture and irrigation models. During the literature review, all studies concerned with irrigation revenues try to optimize profit on global level by suitable water allocation and crop pattern. However, this research went further by distributing profit at the farmland level and among active farmers. In fact, two complementary models are suggested. The first model tries to reach the optimal cropping policy in order to maximize profits by taking into account water availability that is tightly linked to hydropower releases, crop production limits and cropping area constraints. The second model leverages the obtained profit and distributes it among farmers at the parcel level based on a defined cooperative policy. Implementation of both models in Bekaa Valley led to results proving their utility for agricultural decision makers. The final stage of the research is concerned with cascade hydropower operation planning. Here the research seeks an optimal operation on the medium and short term run. However, to achieve an accurate optimal solution, it is important to take into account hydrological and irrigation factors. To bridge this gap, the results obtained, with the aid of the proposed hydrological

and irrigation models, are used during the hydropower operation modeling process. Afterward, the suggested hydropower scheduling models are validated and evaluated in the Litani Hydropower Project. Based on the obtained results, the medium-term hydropower-irrigation plan succeeded to provide boundary conditions, with certain reservoir storage flexibility, to the short-term hydropower planning. On the other hand, the suggested novel short-term model objective is to reduce revenue losses resulted from power shortage, spillage and maintenance tasks. Most of the studies focus on generating power with no production bounds and no maintenance plans. Furthermore, in some studies spillage occurs with high flexibility while in others the contrary. Nevertheless, in order to have a safe and reliable operation, the introduced model takes into account safety measures related to power production and units' maintenance. In addition, spillage is considered with medium flexibility. Indeed, the proposed model managed to enhance the operability and economic benefits of the cascade hydropower system through a comprehensive series of improvements, starting with suitable turbine releases, efficient load distribution, spillages reduction and the best maintenance timing.

Finally, the outcome of this work consists in a multi-functional tool dedicated to operation planning. Its goal is to improve the performance of a multi-purpose reservoir system on several levels: appropriate water discharges, power generation, scheduled irrigation water allocation and cropping pattern. Based on the features of the decision tool, a distributed control structure is proposed. Excel is employed to develop a Human Machine Interface (HMI) while MATLAB and LINGO are used for analysis, optimization and simulation.

Keywords Cloud

Résumé

L'objectif de cette thèse est l'optimisation de la conduite de centrales hydroélectriques en cascade. Ceci passe par l'optimisation de la gestion des ressources opérationnelles et par de bonnes pratiques en termes d'irrigation tout en intégrant les charges environnementales.

L'enjeu consiste à trouver le modèle le plus réaliste basé sur les caractéristiques stochastiques des ressources en eau, de la demande en énergie et du profil d'irrigation. Pour atteindre l'objectif, les méthodes impliquant l'exploration de données et la programmation mathématique serviront de base à un outil d'aide à la décision.

La recherche porte sur trois sujets entrelacés : la modélisation hydrologique, les modèles d'irrigation et la planification des opérations hydroélectriques en cascade. Pour tester la méthodologie utilisée, chaque modèle a été appliqué au projet Litani – Liban.

L'objectif de la modélisation hydrologique est de prédire les débits et l'évaporation de la surface des plans d'eau. Pour établir ces modèles, plusieurs méthodes issues de l'apprentissage automatique, des statistiques et des systèmes de bases de données sont utilisées. Malgré la mauvaise qualité et la faible quantité de données extraites du barrage Litani et du barrage de Qaraoun, les méthodes élaborées ont réussi à traiter efficacement les données disponibles et à présenter de bons résultats de prévision.

Le deuxième sujet d'intérêt dans les systèmes d'irrigation hydroélectrique en cascade est l'agriculture et les modèles d'irrigation. En fait, deux modèles complémentaires sont suggérés. Le premier modèle cherche la politique de culture optimale afin de maximiser les profits en tenant compte de la disponibilité de l'eau qui est étroitement liée aux rejets hydroélectriques, aux limites de production des cultures et aux contraintes des zones de culture. Le second modèle utilise le bénéfice obtenu et le répartit entre les agriculteurs au niveau des terres agricoles sur la base d'une politique coopérative. Lors de la mise en œuvre des deux modèles dans la vallée de la Bekaa, les résultats ont prouvé leur utilité pour les décideurs agricoles.

La dernière partie de la recherche concerne la planification des opérations hydroélectriques en cascade. Ici, un fonctionnement optimal à moyen et court terme est recherché. Cependant, pour obtenir une solution optimale précise, il est important de prendre en compte les facteurs hydrologiques et d'irrigation. Pour cela, les résultats obtenus, à l'aide des modèles hydrologiques et d'irrigation proposés, sont utilisés lors du processus de modélisation des opérations hydroélectriques. Les modèles d'ordonnancement hydroélectrique suggérés sont ensuite validés et évalués dans le projet hydroélectrique de Litani. Sur la base des résultats obtenus, le plan d'irrigation hydroélectrique à moyen terme a réussi à fournir des conditions aux limites, avec une certaine flexibilité de stockage des réservoirs, à la planification hydroélectrique à court terme. D'autre part, le modèle à court terme a réussi à améliorer l'opérabilité et les gains du système hydroélectrique en cascade grâce à une série complète de décisions : choix de turbines appropriées, répartition efficace des charges, réduction des déversements et élaboration du meilleur calendrier de maintenance.

Au final, le résultat est un outil multifonctionnel pour la planification des opérations capable d'améliorer la performance d'un système de réservoirs polyvalent et ceci sur divers points : rejets d'eau appropriés, production d'électricité, allocation d'eau d'irrigation programmée et modèle de culture. Sur la base des caractéristiques de l'outil de décision, une structure de contrôle distribuée est proposée. Excel est employé pour développer une interface homme-machine (IHM) tandis que MATLAB et LINGO sont utilisés pour l'analyse, l'optimisation et la simulation.

Mots-clés : Hydroélectricité, Irrigation, Optimisation, Modélisation, Opération, Planification, Data Mining.

I. Introduction

La portée du problème de production hydroélectrique et d'irrigation (HIP) ne se limite pas à la production d'électricité et la répartition d'eau au sol. Il traite de tous les éléments du problème qui s'étendent du bassin en amont passant par les générateurs de puissance jusqu'aux zones agricoles. En fait, HIP est une série de composants interconnectés qui couvrent, l'infrastructure hydraulique, les relations hydrologiques, les éléments de production hydroélectrique et les activités agricoles. Ces composants sont résumés ci-dessous.

1. **Infrastructure hydraulique** : Les sujets hydrauliques couvrent des concepts tels que l'infrastructure pour l'écoulement, la conception des barrages / réservoirs et le contrôle de l'eau. Cependant, dans un HIP, différents types d'infrastructures hydrauliques produisent des effets de diverses magnitudes. Par conséquent, pour atteindre le meilleur plan de gestion de l'eau, il faut avoir une connaissance approfondie du volume de stockage du réservoir (mort, actif et maximum), des débits maximums des pipelines, de l'efficacité des turbines, du profil du réseau de distribution et du système de déversement. Au cours du processus de modélisation et de validation, le fait de négliger l'un de ces termes produit des résultats irréalistes. Par conséquent, le plan opérationnel hydroélectrique obtenu sera invalide et peu fiable.

2. **Relations hydrologiques** : L'écoulement fluvial et l'évaporation de la surface du réservoir sont des éléments clés dans l'équation du bilan hydrique du réservoir (Pulido-Calvo et Portela, 2007 ; Valiantzas, 2006). L'équation du bilan hydrique est utilisée comme base pour évaluer la disponibilité du stockage de l'eau. Ainsi, intégrer des estimations hydrologiques précises dans les modèles d'irrigation hydroélectrique proposés permet aux décideurs (DM) de comprendre comment le volume d'eau change dans le temps et l'espace. Par conséquent, des décisions éclairées peuvent être prises concernant les rejets d'eau pour l'hydroélectricité, l'irrigation et l'usage municipal. En outre, des événements non productifs peuvent être également réduits ou évités tels que les déversement inutiles et inondations.

3. **Éléments de production hydroélectrique** : Les centrales hydroélectriques utilisent l'énergie potentielle de l'eau stockée pour faire fonctionner des turbines hydrauliques. L'eau circulant amène les arbres de turbine en rotation et l'électricité est produite. De plus, il est connu que, les systèmes hydroélectriques ont une dispatchabilité élevée. En règle générale, ils ont la possibilité de démarrer une rampe arrêtée en quelques minutes, et dans certains cas en quelques secondes (FCH, 2017). Ainsi, tandis que les sources d'énergie moins facilement contrôlables (telles que nucléaire, thermique, etc…) sont utilisés comme centrales électriques de base, les HIP peuvent fonctionner en tant que centrales de suivi de charge et de pointe (Donev, 2017). Elles visent à absorber les fluctuations de demande d'électricité. En fait, cette question est importante car la surintensité cause l'échauffement des lignes affectant l'ensemble

du système et conduisant à des pannes d'électricité. D'autre part, la pénurie d'électricité génère moins de revenus financiers. Dans une telle situation, il est donc fortement recommandé d'adapter l'électricité produite à la demande du marché.

4. **Activités agricoles :** Dans le but d'obtenir des pratiques agricoles productives, il est nécessaire d'introduire de nouvelles techniques de gestion. Ces approches couvrent la bonne sélection de la culture, de la zone de culture, la gestion de la rotation des cultures, l'irrigation, la planification et la répartition des bénéfices entre les agriculteurs et les parties prenantes.

Les quatre points précédents représentent un cadre pour une usine hydroélectrique et d'irrigation typique. Néanmoins, dans les usines d'hydroélectricité et d'irrigation (CHIP), le cas est plus compliqué. Un système en cascade est formé de multi-réservoirs (alimenté par des flux multiples) reliant diverses centrales hydroélectriques à travers un réseau de distribution d'eau complexe. Malgré la complexité du système, le partage des ressources en eau, des infrastructures hydrauliques et réservoirs, reçoivent souvent une attention considérable. De tels projets peuvent générer des avantages polyvalents et sont généralement conçus afin de maximiser des objectifs spécifiques tels que les revenus de la génération d'énergie et / ou de pratiques agricoles.

II. Objectifs de la recherche et portée

De nombreux pays en développement souffrent de pénurie d'énergie et de la dégradation des zones agricoles. Cela est dû à plusieurs raisons, principalement : une mauvaise gestion des ressources et la croissance de la demande en eau et en énergie. Cependant, la fourniture de services durables à ces pays joue un rôle important pour la promotion des améliorations dans les secteurs de l'énergie et de l'eau. Une partie considérable de la littérature scientifique a été consacrée à ce sujet. Apparemment, les centrales hydroélectriques et d'irrigation sont parmi les systèmes les plus efficaces pour le développement et la gestion des ressources en eau. Elles ont la capacité de fournir un certain nombre d'avantages liés à l'eau. Ces avantages peuvent inclure l'approvisionnement en eau pour l'agriculture, l'usage domestique, l'industrie, la génération hydroélectrique et les loisirs. Cependant, pour maximiser les avantages mentionnés, une opération intelligente est obligatoire. En fait, dans ce travail, le problème de la planification opérationnelle comprend trois modules majeurs : prédictions hydrologiques, planification agricole et modélisation de l'hydroélectricité et de l'irrigation.

Prédictions hydrologiques

La prédiction précise du débit de la rivière joue un rôle essentiel dans l'exploitation du réservoir. Néanmoins, la prévision de l'écoulement fluvial reste l'une des questions les plus difficiles quand le système cible est caractérisé par un processus dynamique non-linéaire sous des conditions stochastiques et chaotiques (Huamani et al., 2011). En outre, dans les pays en développement, ce problème est exacerbé par le fait que la collecte et le partage de données hydrométéorologiques n'est pas simple. Les sources de données, dans ces régions, sont généralement limitées et peu (ou

pas) de données locales existent. Pour surmonter ces obstacles et en se basant sur la revue bibliographique, l'inférence floue (fuzzy) semble une solution intéressante. Ainsi, tout au long de la partie hydrologie, la méthode de modélisation des systèmes flous (FSM) est adoptée pour prévoir le débit de la rivière.
Les procédures de modélisation ont considéré deux situations réelles :

1. En cas de rareté des données hydrométéorologiques, le modèle Fuzzy est traité pour exploiter les données disponibles en utilisant des méthodes de prétraitement appropriées. Cependant,
2. Dans le cas d'asymétrie des mesures hydrométéorologiques (les mesures climatiques existent pour une période plus courte en comparaison avec les mesures hydrologiques), les prévisions de débit fluvial utilisent un modèle hybride basé sur la modélisation bidirectionnelle par système flou artificiel.

Un autre sujet d'importance dans la planification hydroélectrique est l'évaporation des lacs. En fait, l'évaporation de l'eau est influencée par plusieurs paramètres météorologiques tels que : l'irradiance, la température du sol, l'humidité relative, la pression atmosphérique et la vitesse du vent (Penman, 1948). Malheureusement, à ce jour, des approximations d'évaporation fiables sont extrêmement difficiles à obtenir. Ce problème est intensifié dans le cas de lacs mal surveillés. Ainsi, il y a besoin de trouver un modèle d'évaporation efficace qui soit moins exigeant sur la quantité des données d'entrée. Ici, la technique suggérée implique seulement trois grandes variables météorologiques : température, humidité relative et point de rosée.

Planification agricole

Dans cette partie, deux modèles mathématiques sont introduits. Un modèle de planification de polyculture (MCP) vise à optimiser les revenus agricoles par une bonne sélection des cultures, une planification précise de l'irrigation en tenant compte de la disponibilité de l'eau à chaque étape. Cependant, l'autre modèle est la distribution des bénéfices (PD). Il est chargé de répartir les gains entre les agriculteurs sur la base d'une politique coopérative prenant en considération la nature physique des terres agricoles et les pratiques du management. Pour les deux modèles, une recherche est menée afin de sélectionner la méthode la plus appropriée pour résoudre les problèmes d'optimisation rencontrés.

Modélisation de l'hydroélectricité

L'hydroélectricité présente un lien important entre l'énergie et l'eau. Bien que la technologie soit plus ancienne par rapport à d'autres ressources d'énergie renouvelables liées à l'eau, l'hydroélectricité a toujours un potentiel significatif principalement dans les systèmes en cascade. En effet, pour bénéficier d'une ressource en eau existante, dans la mesure du possible, il est courant de construire plusieurs centrales hydroélectriques sur la même rivière et de partager un lac d'eau commun (Yildiran et al., 2015). Par conséquent, sur la même rivière, de nouvelles centrales peuvent être ajoutées tandis que les anciennes peuvent être améliorées afin d'augmenter la production électrique et agricole. Ceci nécessite une coordination intelligente des opérations. Sinon, des déversements inutiles peuvent se produire et l'eau peut être gaspillée sans être utilisée ni pour la production d'électricité ni pour l'irrigation. Cette problématique suggère l'utilisation

d’un outil d’aide à la décision. Celui-ci est basé sur deux programmes mathématiques résolus en séquence pour un fonctionnement optimal d'un système hydroélectrique en cascade. À la première étape, le modèle de planification de l'irrigation hydroélectrique à moyen terme (MTHGIS) est résolu. Le but est de minimiser une fonction multi-objectif incluant la puissance, l'irrigation et les déficits hydriques municipaux. Sa résolution fournit une vue d'ensemble de l'opération en cascade et des conditions aux limites pour le second problème. En fait, ce dernier problème concerne les horaires de production d'électricité à court terme (STHGS). L'objectif est d'optimiser les revenus tout en respectant les contraintes liées à l'hydraulique, à l'électricité, à l'irrigation et à la maintenance. Dans l'ensemble, un calendrier judicieux d'engagement des centrales ainsi qu’une bonne répartition de l’énergie à produire peut réduire les déficits de puissance et d'irrigation de manière significative et simultanément augmenter la sécurité et la fiabilité du système en cascade.
Dans cette recherche, chacun des modules proposés est traité séparément. L'avantage de cette approche modulaire est la capacité à entrer dans plus de détails dans chaque sous-domaine et à faciliter la maintenance et la mise-à-jour. Cependant, les différents modules sont liés par l'échange, en amont et en aval, des données de sortie.

III. Méthodologie

L'objectif est de mettre en place une méthode adaptée pour la planification conjointe de l’hydroélectricité et de l’irrigation pour le moyen et court terme. En fait, pour atteindre ce but, des méthodes impliquant l’analyse de données et la programmation mathématique serviront de base à un outil d'aide à la décision (DST). Cet outil multifonctionnel, dont l’objectif est la planification des opérations, sera en mesure d’améliorer la performance d’un système d’irrigation hydroélectrique sur plusieurs points en commençant par des rejets d'eau appropriés, une irrigation adaptée et le choix de cultures adéquates. La liste suivante présente les principaux éléments de la méthodologie :

Modélisation hydrologique

1. Collecte de données
    a. Données météorologiques (précipitations, humidité, température, point de rosée)
    b. Données hydrologiques (débit fluvial, taux d'évaporation)
2. Analyse de données et traitement de données
    a. Transformation de données (Log-Transformation, Standardisation, Normalisation)
    b. Traitement du bruit à l'aide du filtre Moyenne Mobile (MA)
    c. Analyse de corrélation (Autocorrélation, Autocorrélation partielle, fonctions de corrélation croisée)
3. Modèles de prévision des débits de rivière
    a. Méthode auto-régressive (AR)
    b. Méthode de modélisation constructive des systèmes flous (C-FSM)
    c. Approche hybride basée sur la modélisation des systèmes flous constructifs à deux phases (TPC-FSM)
    d. Évaluation de la performance des modèles proposés
4. Estimations de l'évaporation du lac

a. Modèle de régression non linéaire basé sur la formule de Magnus
b. Méthode des moindres carrés non linéaire (NLS)
c. Algorithme de Levenberg -Marquardt
d. Validation et test du modèle proposé

Modélisation agricole

1. Système de culture multiple mis en œuvre dans le cadre d'une irrigation complète et déficitaire
   a. Modélisation mathématique (programmation linéaire et non linéaire)
   b. Algorithmes méta-heuristiques (recuit simulé, optimisation par essaim de particules)
   c. Test des modèles en utilisant des données expérimentales et réelles
2. Répartition des bénéfices (PD) dans les terres agricoles
   a. Présentation du problème en tant que modèle de programmation mixte non linéaire
   b. Linéarisation basée sur des méthodes algébriques
   c. Simulations et analyse des résultats

Modélisation de l'hydroélectricité

1. Distribution efficace de la charge entre les unités de travail de la centrale hydroélectrique
2. Planification de l'hydroélectricité en cascade à moyen terme
   a. Modèle d'optimisation multi-objectif non linéaire minimisant les déficits d'énergie et d'irrigation
   b. Résolution du modèle : approche par somme pondérée
   c. Simulation et analyse des résultats
3. Planification à court terme de l'hydroélectricité en cascade
   a. Présentation d'un modèle de programmation mixte non linéaire minimisant les pertes de revenus de l'opération hydroélectrique
   b. Linéarisation basée sur des méthodes algébriques
   c. Validation et évaluation du modèle

Cette recherche a porté sur une approche fondée sur l'intelligence artificielle pour le contrôle des centrales hydroélectriques et d'irrigation en cascade. En fait, l'apprentissage automatique (Machine Learning) qui met en oeuvre une analyse de corrélation et les techniques d'inférence floue, a pour but de capturer les connaissances cachées dans les données de débit de la rivière. L'objectif est d'extraire les tendances et les comportements qui permettront d'améliorer le fonctionnement de l'hydroélectricité. D'autre part, le développement récent d'algorithmes d'optimisation méta-heuristiques tels que les Algorithmes Évolutionnaires (EA) (Xue- Zhen et al., 2010) et le Recuit Simulé (SA) (Georgiou et al., 2006) offre une alternative pour surmonter certaines limites des techniques classiques comme la programmation linéaire (LP), la programmation dynamique (DP), la programmation non linéaire (PNL) pour la résolution des problèmes concernant la planification polyculture (MCP). Dans cette recherche, l'approche pour résoudre le problème se déroule comme suit. D'abord, des résultats mathématiques préliminaires sont présentés (existence de solution, modèles linéaires de référence et une

relaxation), ensuite deux algorithmes méta-heuristiques, le recuit simulé (SA) et l'optimisation par essaim de particules (PSO), sont mis en œuvre pour résoudre le problème MCP. La particularité de cette approche consiste à utiliser la solution du problème linéaire de référence comme une solution initiale pour le recuit simulé SA, tandis que pour PSO, l'essaim de particules est initialisé dans le voisinage de cette solution. En ce qui concerne le problème de répartition des bénéfices, PD, la disponibilité limitée d'outils méta-heuristiques et le temps non négligeable et nécessaire pour écrire le code à partir de zéro et le tester, ont tourné notre attention vers une approche alternative. Il s'agit de profiter de la disponibilité des solveurs de Programmation Linéaire Entière Mixte (MILP) sur le marché. Un moyen efficace pour résoudre le modèle de Programmation Non-Linéaire Entière Mixte (MINLP) consiste à linéariser le problème par une technique appropriée de changement de variable pour évacuer les termes non linéaires.

À la phase finale du processus de modélisation, connaître la demande en irrigation et la demande d'énergie peut jouer un rôle clé dans l'ajustement des rejets d'eau pour éviter la production d'énergie excessive ou les rejets d'irrigation non désirés. Le plan opérationnel sera capable de maintenir une production d'énergie continue et fiable prenant en compte le besoin en eau des cultures. Dans le problème d'ordonnancement à moyen terme, l'objectif consiste à trouver une solution capable d'établir un équilibre relatif entre différents systèmes d'irrigation. La procédure de résolution pour le problème multi-objectif a utilisé la méthode connue sous le nom de somme pondérée (WS). Cependant, pour le modèle de planification à court terme, la difficulté à résoudre le problème de programmation non-linéaire mixte est contournée en reformulant le problème en un programme linéaire mixte (MILP) à l'aide de techniques simples d'algèbre manipulant les nombres entiers. Le résultat est une structure de contrôle distribuée qui a été mise en œuvre en utilisant plusieurs frameworks logiciels. Les logiciels commerciaux LINGO et MATLAB ont été utilisés pour l'optimisation mathématique tandis que Microsoft Excel a été utilisé pour développer l'interface homme-machine (HMI) pour le contrôle séquentiel.

IV. Étude de cas : Projet de la rivière Litani - Liban

L'étude de cas examine les performances de l'application des modèles nouvellement développés sur la rivière Litani au sein du projet Litani. En effet, les modèles hydrologiques décrits dans la recherche sont expérimentés sur la rivière Litani, tandis que les modèles d'irrigation et d'hydroélectricité sont appliqués au projet Litani.

Litani est la plus longue rivière du Liban qui atteint une longueur de 170 km. Son bassin versant couvre une superficie de 2160 km^2 et il est alimenté par un niveau moyen de précipitations d'environ 764 méga-mètres cube par an (LRA, 2016). La rivière prend source près de l'ancienne ville de Baalbeck, dans la vallée centrale de la Bekaa, à 85 km à l'est de la capitale Beyrouth. Elle coule vers le sud sur 100 km, avant de tourner brusquement vers l'ouest, en Kasmieh juste au nord de la ville de Tyr. À la fin des années 1950, un développement majeur sur la rivière Litani, connu sous le nom Projet Litani, impliquait la construction d'un barrage artificiel Qaraoun (QD) et de ses structures associées : réservoirs secondaires (Anan et Joun), centrales hydroélectriques et systèmes d'irrigation.

Le lac Qaraoun, d'une capacité de stockage de 220 méga-mètres cube, est situé au milieu de la rivière Litani. Il détourne le flux de la rivière à travers un système de tunnels reliant les trois centrales hydroélectriques : Markaba (34 MW), Awali (108 MW) et Charles Helou (48 MW).

V. Résultats de la thèse

La thèse est composée de cinq chapitres. La répartition des chapitres commence par la modélisation hydrologique pour prédire le débit futur de la rivière et s'étend pour estimer la surface du lac, l'évaporation des lacs mal surveillés. Par la suite, des modèles d'irrigation et agricoles sont introduits et couvrent l'approche de planification multi-cultures mise en œuvre dans le cadre de l'irrigation et du profit déficitaires / non déficitaires et la distribution parmi les agriculteurs actifs. Une autre partie essentielle de la recherche est la cascade hydroélectrique, avec une prédiction de fonctionnement à court terme. En outre, dans la recherche, une approche de modélisation modulaire est adoptée dans laquelle des modules détaillés ont été développés séparément. La principale importance de l'approche modulaire est sa capacité d'explorer les détails dans chaque sous-domaine, et la capacité d'effectuer une mise à jour et une amélioration d'une façon indépendante. Le résultat final consiste à présenter un cadre pour la prise de décision dans le domaine de l'énergie tout en assurant une exploitation des données acquises-en amont et en aval.

Modélisation hydrologique

Les modèles hydrologiques sont des outils très utiles qui sont largement utilisés dans l'exploitation de l'énergie hydraulique. Cependant, la disponibilité de grandes quantités de données pour la formation et la validation du modèle pose problème. Selon Zemadim, les pays en développement sont confrontés à des problèmes techniques et financiers. Des contraintes qui restreignent à la fois la collecte de données et les efforts de partage (Zemadim et al., 2014). Les obstacles techniques sont liés à la taille du bassin versant surveillé et à la disponibilité d'équipements et de main-d'œuvre qualifiés. D'autre part, les contraintes financières sont étroitement liées non seulement aux dépenses liées à des systèmes de surveillance sophistiqués, mais aussi aux coûts permanents de maintenance du système. Tous ces facteurs constituent une préoccupation majeure pour les hydrologues en termes de qualité et de quantité des mesures récupérées. De ce fait, la fiabilité des prédictions hydrologiques est limitée dans ces pays, car les données hydrométéorologiques locales sont souvent rares et souvent de mauvaise qualité. Malgré ces limites, il y a une forte envie de développer de nouvelles approches de prévision. Ils doivent être capables d'exploiter efficacement les données disponibles et d'obtenir des résultats fiables utiles pour l'exploitation de centrales hydroélectriques. Sachant que les prévisions hydrologiques précises sont d'une importance vitale pour la gestion et le contrôle efficaces de l'eau du réservoir.

En fait, dans cette recherche, l'objectif principal est d'estimer deux quantités cruciales : le taux d'évaporation et la prévision du débit fluvial. Ce sont des problèmes considérés parmi les plus difficiles en sciences hydrologiques car ils ont un aspect dynamique, incertain et non linéaire

(Huamani et al., 2011). Ces problèmes concernent un système qui reçoit des milliers d'entrées interagissant dans un environnement complexe et bruité. Concernant les estimations d'évaporation d'eau, l'inconvénient persiste également. De nombreux chercheurs ont développé différents modèles pour estimer l'évaporation (Penman, 1948, Penman, 1963, Priestley et al., 1972, Linacre, 1977). Malheureusement, des approximations précises et à jour sont extrêmement difficiles à obtenir. Cela est dû aux interactions complexes entre de nombreux facteurs météorologiques interdépendants. L'objectif principal de cette recherche est de présenter, d'analyser et de discuter de divers modèles hydrologiques. En fait, ces modèles ont été mis en œuvre en utilisant différentes méthodes d'exploration de données, alors que plusieurs scénarios ont été testés dans la rivière Litani afin d'évaluer leur performance.

Premièrement, dans ce travail, une analyse préliminaire des données statistiques a été effectuée afin de déterminer la qualité et la quantité des données hydrométéorologiques collectées du bassin de Litani. Malheureusement, les données souffrent d'insuffisance, d'imprécision, d'hétérogénéité, d'asymétrie et sont parfois peu fiable en information fournies par les stations de jaugeage.

Cependant, l'étude bibliographique montre que l'inférence Fuzzy semble être utile pour traiter les inconvénients mentionnés (Bouchon-Meunier et al., 2008, Cheng et Li, 2012). De ce fait, une variante de l'inférence Fuzzy, connue sous le nom de Modélisation de système floue constructive (C-FSM) (Luna et al., 2007), a été utilisée pour modéliser le débit quotidien de la rivière Litani.

En outre, le filtre « Moving Average », adressé dans le modèle C-FSM_MA, a fourni un outil de soutien pour la modélisation Fuzzy. Il n'a pas seulement réduit le bruit inhérent aux données pluviométriques, mais il a aussi préservé la variabilité du débit dû à la pluie. Dans l'ensemble, l'analyse présentée a montré qu'une variante du modèle C-FSM_MA a une meilleure précision que le reste des modèles, AR et C-FSM, dans les prévisions de débit fluvial utilisant différents indicateurs statistiques.

Malgré la rareté, l'hétérogénéité et la non-normalité des données météorologiques et hydrologiques, mises à part les incertitudes héritées des activités illégales signalées le long du Litani et en raison de facteurs tels que l'urbanisation et l'industrialisation, les résultats des modèles C-FSM et C-FSM_MA sont très raisonnables.

Par conséquent, la mise en œuvre réussie du modèle C-FSM était un motif pour aller de l'avant pour faire face au problème de l'asymétrie. En conséquence, un modèle hybride est présenté sur la base du modèle Fuzzy Constructive Two-Phase Modeling System (TPC-FSM).

L'approche TPC-FSM proposée rejoint la modélisation hybride avec la méthode C-FSM afin de gérer l'asymétrie des données hydrométéorologiques. En fait, la série temporelle d'écoulement de la rivière est décomposée en une composante linéaire et non linéaire. Les deux composantes sont équipées de l'approche C-FSM. Tout d'abord, la composante linéaire est estimée, puis le second terme non linéaire est obtenu sur le résiduel. L'objectif de l'approche TPC-FSM est d'exploiter au mieux les données disponibles pour surmonter les inconvénients du C-FSM conventionnel dans des conditions d'asymétrie de données. L'objectif de TPC-FSM est d'obtenir

une plus grande précision de prévision par rapport au C-FSM pour le débit journalier de la rivière en séries chronologiques.

Le modèle TPC-FSM suggéré a montré une meilleure performance par rapport au modèle C-FSM. Les résultats étaient prometteurs avec une bonne concordance globale (93%) entre les valeurs prédites et observées pour le jour le plus long.

Concernant l'estimation de l'évaporation de surface, le problème de la rareté des données persiste. Cependant, la recherche a réussi à fournir un modèle, basé sur la méthode non linéaire Least Square (NLS) et sur la formule « Simplified Penman », pour estimer l'évaporation dans les lacs faiblement surveillés.

En fait, dans une première étape, utilisant l'algorithme de Levenberg- Marquardt (LM), un modèle de régression non linéaire multi- variée (MNR) $D = f(T, RH)$ (f est une fonction basée sur la formule de Magnus) est formé et testé pour l'estimation du point de rosée. Par conséquent, le rayonnement solaire (R_s) peut être approximé en utilisant une fonction spéciale qui lie R_s à D. A l'étape deux, en considérant les trois entrées T, RH et R_s, la version simplifiée de la formule de Penman fournit une estimation du taux d'évaporation du lac (Valiantzas, 2006). Par conséquent, la méthode proposée fournit une solution applicable dans les lacs mal surveillés. En fait, les prévisions habituelles de la température et de l'humidité seront suffisantes pour produire une prédiction d'évaporation plausible.

La méthodologie présentée afin d'estimer les pertes d'eau dues à l'évaporation, n'utilise que des données climatiques communément mesurées (température et humidité relative). Pour illustrer l'efficacité et les capacités de l'approche suggérée, le lac Qaraoun a été choisi comme étude de cas. Dans une première étape, un modèle de régression non linéaire multi- variée (MNR) a été formé en utilisant la méthode des moindres carrés. On a pu prédire avec précision le point de rosée avec $R^2 = 0,99$. Par la suite, la sortie du modèle est utilisée comme entrée pour la version simplifiée de l'équation de Penman. Le résultat est une estimation pour l'évaporation de surface. En effet, les valeurs obtenues ont été comparées à l'évaporation mensuelle moyenne récupérée d'une région proche du lac connu sous le nom de Tal Amara. Considérant le montant limité de données, les taux estimés sont fiables avec un coefficient de corrélation égal à 0,8.

Les résultats obtenus ont été suffisamment précis pour permettre, plus tard dans la thèse, une évaluation plus poussée des pertes par évaporation du réservoir de Qaraoun sur la production hydroélectrique et sur l'irrigation. À ce niveau, les pertes financières associées motiveront les ingénieurs à soumettre des solutions pour réduire le volume d'évaporation afin d'augmenter les revenus.

Bien que les données disponibles souffrent de différents types d'inconvénients, les modèles pilotés par les données et basés sur la modélisation du système flou constructif à deux phases et la méthode des moindres carrés non linéaires ont été appliqués avec succès. TPC-FSM et NLS visent à établir respectivement le débit et l'évaporation du réservoir de Qaraoun QD. Leurs résultats prometteurs ont été une motivation pour adopter le modèle TPC-FSM pour générer de futurs débits et le modèle NLS pour estimer les pertes d'eau d'évaporation dans le bilan hydraulique afin d'améliorer la programmation de l'opération hydroélectrique. Les prévisions

aideront à trouver une politique opérationnelle optimale précise, à différentes étapes, à travers les bonnes décisions de décharge.

Modèles d'irrigation et d'agriculture

Traditionnellement, les modèles agricoles visaient principalement à maximiser la production et le rendement économique par unité de surface en allouant de l'eau à différentes cultures en fonction de leurs besoins en eau (Afshar et Mariño, 1989, Onta et al., 1995, Garg et Ali, 1998). Avec le temps, les études abordaient l'irrigation déficitaire et son impact sur la production des cultures. L'objectif était de réglementer l'irrigation déficitaire de manière à économiser l'eau en soumettant les cultures à des périodes de pénurie hydrique avec un effet minimum sur les rendements. Dans le cadre de cette approche et sur la base des rapports de la FAO, la réduction du rendement peut être minimisée, par rapport aux avantages obtenus, en détournant l'eau économisée pour couvrir de plus larges zones cultivées (FAO - Rapport sur l'eau, 2002). Garga affirme qu'une irrigation optimale est utile pour augmenter la production, la superficie irriguée et les bénéfices économiques nets (Garga et Dadhich, 2014).

Dans cette recherche, deux modèles mathématiques complémentaires ont été développés pour la planification agricole. L'objectif est de fournir des réponses aux questions suivantes:

1. Quelle est la répartition optimale de l'eau, les ressources en terres et la configuration des cultures en tenant compte de la disponibilité de l'eau, de la production végétale et des contraintes des zones de culture ?
2. Quel est le profit maximal réalisable ?
3. Comment le bénéfice obtenu peut-il être réparti entre les agriculteurs actifs ?
4. Comment optimiser la distribution des cultures sur les terres agricoles en tenant compte de la rotation des cultures ?

En fait, les modèles fournis ont été formulés en tenant compte des questions abordées. Le premier modèle est la planification multi-cultures (MCP) mise en œuvre sous irrigation déficitaire. L'objectif est divisé en deux sous-objectifs : 1- maximiser le profit du système en utilisant les ressources en eau disponibles pour le modèle multi-cultures proposé au cours de la période de planification ; 2- Générer un profil d'irrigation optimal qui peut être utilisé lors de la planification hydroélectrique à moyen terme. Le deuxième modèle est relatif au paradigme de la distribution des profits (PD). Il vise à répartir le bénéfice obtenu par le modèle MCP parmi les agriculteurs actifs.

Au cours de la dernière décennie, des études approfondies ont été menées sur les algorithmes évolutifs (EA) pour résoudre les problèmes de programmation non-linéaire concernant la planification optimale des cultures et de l'irrigation. L'Algorithme génétique (GA) a été utilisé pour résoudre le problème de l'irrigation (Alvarez et al. 2004, K. Srinivasa Raju et Ashok, 2006), alors que dans (Georgiou et al., 2006) on recherche l'exploitation du réservoir d'irrigation

optimale à l'aide Recuit simulé (SA). D'un autre côté, l'algorithme d'optimisation par essaim de particules (PSO) a été appliqué pour trouver l'opération optimale du réservoir pour l'irrigation de plusieurs cultures (Kumar et Reddy, 2007, Noory et al., 2012). Cependant, l'approche problème-solution présentée dans l'étude de MCP est donnée dans le paragraphe suivant.

L ' approche numérique pour résoudre le problème de la planification multi - cultures (MCP) est la suivante. Tout d'abord, un modèle non-linéaire PNL est établi pour décrire le problème MCP. Par la suite, des résultats mathématiques préliminaires sont établis qui impliquent l'existence de solutions, l'extraction de modèles linéaires et la formulation d'une relaxation de la PNL. Ensuite, deux algorithmes méta-heuristiques SA et PSO sont implémentés en tant que technique numérique pour résoudre le problème MCP. L'approche utilise la solution du problème linéaire comme solution initiale pour le SA, tandis que pour PSO, l'essaim de particules est initié au voisinage de cette solution, plutôt que de le générer aléatoirement (Georgiou et al., 2006 ; Kumar et Reddy, 2007). L'efficacité des modèles proposés est testée et évaluée à l'aide de données expérimentales. Basé sur l'exemple numérique fourni, les résultats obtenus par les algorithmes Recuit Simulé et essaim de particules initialisés à l'aide de la solution d'un problème linéaire spécifique, ont révélé une diminution significative du temps d'exécution de l'algorithme, et une augmentation de la superficie cultivée et du résultat financier total relatif à l'agriculture sous irrigation déficitaire.

Les résultats de calcul nous amènent à considérer une situation de cas réel pour étudier les capacités réelles de l'approche suggérée. Ceci a été illustré à travers une mise en œuvre utilisant des données réelles obtenues dans la région de Bekaa- Valley près du réservoir Qaraoun alors que la disponibilité de l'eau à chaque étape est étroitement liée à la production hydroélectrique. Toutefois, le calendrier d'irrigation obtenu n'a pas utilisé tout le montant disponible de l'eau. Pour surmonter ce problème, une version ajustée (AMCP) du modèle MCP a été introduite. Le modèle AMCP a réussi à augmenter les revenus en réorganisant le profil d'irrigation. En réalité, le modèle AMCP peut jouer un rôle important en fournissant au modèle à moyen terme les demandes d'irrigation chaque mois. De cette façon, le risque que les rejets d'eau pour l'irrigation ne soient pas pleinement utilisés est réduit.

En ce qui concerne le modèle PD, l'intérêt principal est dans la méthode de distribution du profit financier, obtenue par le modèle MCP, parmi les agriculteurs actifs dans cette vallée au niveau local. Selon Staatz, les coopératives agricoles s'efforcent de maximiser les bénéfices qu'elles génèrent pour leurs membres (ce qui implique généralement une opération sans but lucratif) et le bénéfice est réparti entre les membres agriculteurs sous la forme d'un dividende proportionnel à la part de l'agriculteur (Staatz, 1987). Par conséquent, le modèle de distribution des bénéfices (PD) adopte la mesure coopérative pour identifier le modèle de culture au niveau de la ferme ou de la parcelle.

L'idée de base de ce problème de PD est une ré-initialisation du modèle de programmation linéaire mixte entier présenté dans le travail de Pap (Pap, 2008) en le mettant dans un meilleur contexte, en gardant à l'esprit deux choses principales : 1- une répartition efficace des cultures sur les terres agricoles ; 2- la rotation des cultures. Ici, on peut prétendre que le mode de rotation de culture peut être établi dans le problème MCP. En fait, cette procédure augmentera la

complexité du problème et le rendra plus difficile à résoudre. Par conséquent, il était plus pratique d'introduire la contrainte de rotation des cultures dans le problème PD moins complexe.

La méthode proposée est basée sur une politique coopérative définie pour répartir le bénéfice global réalisé par le modèle MCP parmi les agriculteurs. Les capacités du modèle PD suggéré ont été illustrées par sa mise en œuvre dans la vallée de la Bekaa au cours de deux saisons consécutives où les cultures considérées sont la pomme de terre, le maïs, l'ensilage et la tomate. La formulation MILP du problème PD a réussi à maintenir le même schéma de distribution des bénéfices dans la politique de rotation des cultures présence / absence. Cette technique est efficace pour faire face à la complexité de la gestion du profit entre plusieurs propriétaires fonciers impliqués dans le même projet agricole.

L'un des inconvénients potentiels se manifeste dans les pays à économie non-libre alors que cette approche accorde le profit le plus élevé à l'agriculteur ayant la plus grande zone d'activité. Cependant, dans le cadre des travaux futurs, la résolution de ce problème pourrait être réalisée en remplaçant la zone d'activité par le rapport de superficie totale avec un facteur de pondération. Ce facteur de pondération serait proportionnel non seulement à la zone d'activité, mais également à la situation financière de l'agriculteur. Dans ce cas, les agriculteurs débutants pourraient améliorer leurs exploitations et soutenir leurs activités agricoles.

Exploitation des centrales hydroélectriques d'irrigation en cascade

Bien que l'hydroélectricité soit une technologie mature, elle a encore un potentiel important surtout dans les systèmes en cascade. Il est courant de voir plusieurs centrales hydroélectriques construites sur la même rivière et partageant un lac d'eau commun (Yildiran et al., 2015). Sur une même rivière, de nouvelles centrales peuvent être jointes et d'anciennes peuvent être améliorées afin d'augmenter la production d'électricité et les productions agricoles. Ces améliorations ne sont réalisées que grâce à des opérations et une coordination intelligente. Sinon, des déversements peuvent se produire et l'eau peut être gaspillée sans n'être utilisée ni pour la production d'électricité ni pour l'irrigation.

Cette recherche suggère des modèles détaillés afin d'améliorer l'opération d'irrigation hydroélectrique à moyen et court terme. Pour atteindre l'objectif souhaité, il est important de prendre en compte les éléments clés suivants :

1. Entrées : demandes de charge ; estimations hydrologiques ; exigences d'irrigation ; stockage initial du réservoir ; les prix de l'électricité ; coûts de maintenance.
2. Contraintes : électrique (capacité de puissance des centrales, transition de puissance, etc…) ; hydraulique (pipelines, système de déversement, stockage, etc.) sécurité (entretien, effet de coup de bélier, stockage, etc.) ; accords d'irrigation ; environnemental.

3. Objectifs : équilibrage de la puissance, satisfaction des demandes d'irrigation, minimisation des pertes de revenus de l'exploitation hydroélectrique, répartition efficace de la charge, maintien de l'efficacité élevée des turbines.

Le modèle d'ordonnancement à moyen terme est généralement utilisé pour déterminer la stratégie optimale pour l'opération hydro-irrigation sur une année (avec un incrément de temps d'un mois). L'objectif est de gérer efficacement les ressources en eau pendant la production hydroélectrique et l'irrigation. De plus, le modèle à moyen terme, dans le cas d'une topologie similaire au modèle à court terme, joue un rôle important dans la fourniture de conditions aux limites appropriées. En fait, après résolution, le rôle de l'opérateur hydroélectrique est de transformer les résultats du processus d'ordonnancement à moyen terme en une forme appropriée dans le processus d'ordonnancement à court terme.

Le modèle à court terme est mis en œuvre dans un horizon temporel d'un jour à une semaine et utilise des incréments d'une heure. Il supporte un modèle de système plus détaillé que dans le modèle à moyen terme. Sa tâche principale est de suggérer le calendrier de production journalière pour les centrales hydroélectriques d'irrigation. Ceci implique différents sous-problèmes : engagement de l'unité, répartition économique des charges, courbe d'efficacité, tâches de maintenance, etc.

L'objectif global est de développer un ensemble de nouveaux modules pour aider les exploitants à prendre les meilleures décisions possibles en matière de planification et d'exploitation à moyen et court terme. Ici, quatre sujets sont étudiés:

1. Évaluation financière de l'évaporation de l'hydroélectricité et de l'irrigation dans le projet Litani

Cette partie consiste à évaluer l'impact financier annuel de l'évaporation sur le projet Litani. Tout d'abord, sur la base de la méthodologie présentée dans la section modélisation hydrologique, le volume d'évaporation est estimé au cours de l'année 2013. Le résultat obtenu a montré que la perte d'eau estimée est d'environ 17,88 méga-mètres cube. Ensuite, l'effet économique de l'eau perdue due à l'évaporation est évalué. Malheureusement, le volume évaporé a entraîné une perte financière tangible sur les secteurs de l'hydroélectricité et de l'irrigation. On estime que la perte d'énergie hydraulique est d'environ 0,76 million d'euros, alors que la perte d'irrigation est énorme atteignant une valeur de 39,22 millions d'euros. Sur la base de ces résultats, le terme d'évaporation ne peut pas être simplement négligé. Ainsi, dans cette recherche, il est important d'inclure l'évaporation dans les planifications à moyen et court terme.

2. Distribution de charge électrique pour un rendement optimal

Cette partie présente une méthode pour améliorer la solution au problème de l'engagement des unités de production électrique et au problème de répartition des charges économiques. En fait, en utilisant le schéma de la courbe d'efficacité, deux améliorations sont apportées aux problèmes

décrits : 1- maintenir l'efficacité de la turbine presque optimale à travers les bonnes limites de libération de l'eau. De cette façon, la solution du problème d'engagement d'unité est améliorée en maintenant l'unité en fonctionnement (en ligne) à haute efficacité, sinon, elle est mise hors ligne ; 2- répartir efficacement la charge entre les groupes d'unités. Ici, il est mathématiquement prouvé que la répartition des charges économiques peut être mieux réalisée par une répartition égale de la charge entre les unités de travail. Par conséquent, en considérant ces mesures, toute l'unité travaillera à son rendement le plus élevé pendant son planning prévu. En conséquence, en incorporant ces deux améliorations dans les modèles de planification, moins d'eau est libérée et plus d'énergie est produite.

3. Ordonnancement à moyen terme de CHIP

De nombreuses études traitaient de la planification à moyen terme d'un système de réservoir à cascade à usages multiples (Reddy et Kumar, 2006, Bai et al., 2015). Apparemment, la planification opérationnelle implique des interactions et des compromis entre divers objectifs qui peuvent parfois être complémentaires mais qui sont souvent concurrentiels et conflictuels. Par exemple, il existe essentiellement des conflits d'intérêts entre : 1- la part de l'eau municipale en amont et en aval ; 2- l'approvisionnement en eau dans différents secteurs (hydroélectricité, irrigation, loisirs) ; 3- la sécurité structurale des différents réservoirs ; 4- les problèmes environnementaux. Par une mauvaise coordination entre les réservoirs en cascade, l'eau peut être gaspillée sans être utilisée efficacement.

Ici, l'objectif du modèle d'ordonnancement de la production hydro-électrique à moyen terme (MTHGIS) est de minimiser les déficits d'électricité, d'irrigation et d'eau à usage municipal. Le problème est résolu en utilisant la méthode de la somme pondérée. Pour tester la plausibilité du modèle, celui-ci a été validé en utilisant l'étude de cas réel : Litani Project. En effet, une comparaison est effectuée entre l'opération réelle mise en place en 2011 et la simulation. Les résultats de la simulation se sont rapprochés des données de l'opération réelle. En conséquence, le modèle a réussi le test de fiabilité. Ensuite, plusieurs scénarios de différents profils de charge sont simulés. Dans le cas libanais, l'analyse effectuée reposait sur deux conditions : 1) les exigences en matière d'énergie et d'irrigation doivent être satisfaites ; 2- le réservoir de stockage ne doit pas être à proximité du niveau de stockage mort.

Les résultats indiquent que la production d'électricité ne doit pas dépasser 0,15 écart-type au-dessus de la moyenne au cours d'une année normale, sinon les deux conditions sont violées. Dans l'ensemble, le modèle MTHGIS introduit a permis, dans une certaine mesure, de donner un aperçu général de l'opération d'hydro-irrigation à moyen terme. Il a également initialisé les conditions aux limites avec une certaine flexibilité de stockage du réservoir pour le modèle à court terme plus détaillé.

4. Planification à court terme de CHIP

Dans la section précédente, le modèle MTHGIS pour les systèmes en cascade a fourni avec succès des limites de stockage de réservoir et un aperçu général de l'opération d'hydro-irrigation. Ici, dans cette partie, l'intérêt principal réside dans un modèle plus détaillé concernant la planification de la production hydroélectrique à court terme (STHGS) des centrales

hydroélectriques en cascade. Cependant, pour construire un modèle réaliste détaillé, il est essentiel de considérer tous les éléments liés à la disponibilité de l'eau, à la production d'électricité et à l'hydraulique du système.

L'hydroélectricité peut fonctionner en s'adaptant aux pics de charge. Dans une telle situation, il est recommandé d'adapter l'alimentation électrique à la demande pour la production d'hydroélectricité. En fait, la minimisation du déficit énergétique a été considérée dans le travail de Zhang, l'objectif étant de trouver les rejets d'eau horaires optimaux (R. Zhang et al., 2013). Cependant, l'approche suggérée peut favoriser un déversement inutile. Ce problème est résolu en évitant le déversement (hors turbine) tout en veillant à ce que l'écoulement dans la turbine soit en dessous de son maximum (Zhang et al., 2016) ou en introduisant le déversement (hors turbine) comme terme de pénalité quadratique à la fonction objective (Yu et al., 2015). Dans notre travail, ce terme de déversement est linéaire pour plus de flexibilité opérationnelle. En outre, on peut prétendre que le déversement dans les réservoirs supérieurs peut aider à améliorer l'efficacité de la production d'énergie dans les réservoirs inférieurs. Ce problème est résolu par : 1- le maintien du rendement élevé de la turbine à travers les bonnes limites de rejet d'eau. De cette façon, la solution améliorée du problème d'engagement d'unité se manifeste en maintenant l'unité en fonctionnement à haute efficacité en ligne, sinon, l'unité est mise hors service. En fait, une approche similaire est présentée dans le travail de Lu (Lu et al., 2015) ; 2- la distribution efficace de la charge. Ici, il a été prouvé mathématiquement que la répartition de la charge économique peut être mieux réalisée par une répartition égale de la charge entre les unités de travail.

De plus, pour une production d'énergie stable, une gestion efficace de la maintenance est requise. Elle peut jouer un rôle majeur dans le maintien de la fiabilité de l'usine. Selon la bibliographie, la plupart des études portant sur l'optimisation en cascade ne tiennent pas compte de la maintenance préventive (Guedes et al., 2015). Leur travail modélise l'ordonnancement de la maintenance en tant que variable continue en utilisant une fonction auxiliaire non linéaire. Le volume du réservoir et le calendrier d'entretien ont été optimisés, simultanément, pour minimiser le complément de production thermique lorsque l'hydroélectricité est utilisée comme charge de base (Guedes et al., 2015). Dans notre approche, la non-linéarité est considérée dans la procédure de maintenance en l'introduisant comme un sous- problème binaire avec un ensemble de contraintes linéaires. De plus, les coûts de maintenance sont incorporés dans la fonction l'objective. On sait que les réservoirs hydroélectriques polyvalents sont conçus pour fournir des services au-delà de la production d'énergie, tels que l'approvisionnement en eau pour l'irrigation (Branche, 2017). Villavicencio a présenté un modèle de programmation mixte linéaire et en nombres entiers pour la planification de l'exploitation hydroélectrique à court terme et prend en compte les besoins d'irrigation (Villavicencio et al., 2015). Par conséquent, dans une usine d'hydroélectricité-irrigation, il est important de considérer des accords d'irrigation afin d'avoir un modèle réaliste. L'objectif de la recherche est d'atteindre le revenu optimal. En fait, un objectif similaire est présenté dans le travail de (Pérez- Díaz et al., 2010). Leur modèle de programmation non linéaire vise à maximiser les revenus générés par la vente de l'énergie produite dans un marché de l'électricité déréglementé sans aucune contrainte sur la production d'électricité. Néanmoins, si le système hydroélectrique en cascade fonctionne comme un suivi de charge, il devrait suivre un certain profil de charge. Tout déséquilibre entre l'offre et la demande rend le système électrique instable, ce qui peut entraîner des effets économiques et

techniques graves (Moseley et Garche, 2014). Par conséquent, pour optimiser le revenu, il est recommandé de minimiser les pertes de revenus dues à la pénurie d'électricité.

En fait, la recherche suggère un modèle détaillé d'ordonnancement à court terme des centrales hydroélectriques (STHGS). Les innovations significatives du modèle proposé comprennent principalement trois points: 1- la fonction objective représente les pertes de revenus dues au déficit de puissance, au déversement et aux tâches de maintenance; 2- en fonction des recommandations des points 1 et 2 précédents , les solutions des sous-problèmes d'engagement de l' unité et de répartition de la charge sont améliorées dans le modèle STHGS; 3- les procédures de maintenance sont introduites dans le modèle comme un ensemble de contraintes linéaires et les coûts de réparation sont intégrés dans la fonction objective. Au cours du processus de modélisation, le modèle STHGS obtenu a trouvé une forme plutôt compliquée avec de nombreux termes non linéaires. Cependant, pour réduire sa complexité, le problème STHGS a été formulé comme un modèle MILP en utilisant des techniques numériques et algébriques de linéarisation. Par la suite, la performance du MILP suggéré est étudiée dans le cadre du projet Litani lors d'événements spéciaux (comme le débit de pointe de la rivière, l'inondation du réservoir principal pendant la saison d'irrigation) et dans les cas extrêmes (comme le débit de pointe de toutes les ressources en eau et les centrales hydroélectriques fonctionnant presque à pleine capacité). Les résultats obtenus sont évalués en comparant des résultats de simulation entre les méthodes suggérées et les méthodes existantes. Par exemple, dans l'opération actuelle, une quantité importante de déversement s'est produite dans le système à différentes dates. Cependant, avec l'utilisation d'un terme de déversement dans la fonction objective, le déversement est évité sauf dans des situations extrêmes. En même temps, l'intégration intelligente de la maintenance dans le modèle garantit la rentabilité économique de la centrale hydroélectrique. Ceci se manifeste à travers la détermination du meilleur moment pour l'inspection de la machine avec des pertes économiques minimales. Dans l'ensemble, la mise en place du système hydroélectrique Intelligent Control-Maintenance-Management a amélioré le fonctionnement à court terme grâce à une série de décisions optimales commençant par les décharges d'eau, la séquence de mise en ligne ou non des unités, la répartition des charges et les meilleures dates de maintenance.

En général, pour un fonctionnement optimal des générateurs en cascade, une interface homme-machine (IHM), couplée à un noyau de décision basé sur le paradigme MTHGIS et STHGS, sera reliée à un système SCADA pour la surveillance et le contrôle à distance. Il permettra de réduire efficacement les fautes commises par les opérateurs humains en centralisant l'ensemble du processus avec un système entièrement automatisé.

VI. Conclusion

La planification des systèmes hydro-électriques d'irrigation reste un domaine de recherche actif et de nombreuses études sont publiées chaque année. Cependant, l'aspect intéressant de ce travail est qu'il réussit à diviser le problème principal d'hydroélectricité-irrigation en plusieurs sous-problèmes importants. La solution pratique de chaque sous-problème agit comme un outil de soutien dans l'amélioration de l'opération hydro-irrigation. On sait qu'au Liban et dans de nombreux pays en développement, l'opération hydro-irriguée est déterminée en utilisant une

approche par essais et erreurs. Les principales raisons sont : 1- qualité et quantité de données médiocres ; 2- absence de systèmes d'acquisition de données fiables ; 3- manque d'implication des experts ; 4- absence de plans d'ordonnancement efficaces ; 5- obstacles financiers au développement de l'hydroélectricité. En fait, cette recherche a réussi à fournir des approches peu coûteuses pour surmonter tous les obstacles mentionnés et elles sont résumées ci-dessous : 1 - Différents modèles hydrologiques sont suggérés pour résoudre les problèmes de données tels que rareté, incohérence, hétérogénéité et asymétrie. 2- Un modèle mathématique est présenté capable de fournir des profils d'irrigation fiables ; 3.- Un outil complet d'aide à la décision (DST) peu coûteux, est proposé pour aider les opérateurs de systèmes d'irrigation hydroélectrique à établir des plans d'ordonnancement d'experts à court et à moyen terme.

En conclusion, l’outil DST exposé représente une solution attrayante pour les pays en développement dans le secteur de l'hydroélectricité-irrigation. Malgré les défis existants, le DST a réussi à générer des plans efficaces et utiles pour la gestion des ressources en eau à objectifs multiples.

Publications and Activities

A. Refereed international journal articles

1. Bou-Fakhreddine B., Mougharbel I., Faye A., Abou-Chakra S. and Pollet Y· (2018). Daily River Flow Prediction Based on Two Phase Constructive Fuzzy Systems Modeling: A Case of Hydrological - Meteorological Measurements Asymmetry. Submitted to Journal of Hydrology (Accepted on 15 January 2018).

2. Bou-Fakhreddine B., Abou-Chakra S., Mougharbel I., Faye A. and Pollet Y· (2018). Daily River Flow Prediction Coupled with Data Processing Techniques: A Comparative Study between Constructive Fuzzy Systems and Autoregressive Models. Submitted to Hydrology Research on 09 December 2017 (under review).

3. Bou-Fakhreddine B., Mougharbel I., Faye A. and Pollet Y· (2018). Optimal Short-term Cascade Hydropower Scheduling and Maintenance Planning. Submitted to Journal of Hydro-Environment Research on 16 December 2017 (under review).

B. Papers in refereed conference

1. Bou-Fakhreddine B., Abou-Chakra S., Mougharbel I., Faye A. and Pollet Y· An Optimal Multi-Crop Planning Approach Implemented Under Deficit Irrigation. MELECON 2016 April 18 - 20, 2016, Limassol, Cyprus. IEEE conference.

2. Bou-Fakhreddine B., Abou-Chakra S., Mougharbel I., Faye A. and Pollet Y· Short-Term Hydro Generation Scheduling Of Cascade Plants Operating On Litani River Project Lebanon. REDEC 2016 July 13-14, 2016, Notre-Dame University, Zouk Mosbeh, Lebanon. IEEE conference.

3. Bou-Fakhreddine B., Abou-Chakra S., Mougharbel I., Faye A. and Pollet Y· Estimating Daily Evaporation from Poorly – Monitored Lakes using limited Meteorological Data. SDEWES 2017 October 4-8, 2017, Dubrovnik, Croatia.

C. Activities and Presentations

1. Bou-Fakhreddine B. Daily River Flow Prediction Coupled with Data Processing Techniques: A Comparative Study between Constructive Fuzzy Systems and Autoregressive Models. Poster Presentation. JFL3 Troisièmes Journées Franco-Libanaises, 2015.

2. Bou-Fakhreddine B. An Optimal Multi-Crop Planning Approach Implemented Under Deficit Irrigation. Poster Presentation. MELECON 2016 April 18 - 20, 2016, Limassol, Cyprus.

3. Bou-Fakhreddine B. An Optimal Multi-Crop Planning Approach Implemented Under Deficit Irrigation. Doctoral day, EDST, 2016.

4. Bou-Fakhreddine B., Mougharbel I. LabVIEW System Applied to Distributed Hydropower Stations. Oral Presentation, NIDays Conference, 2016.

5. Bou-Fakhreddine B. Short-Term Hydro Generation Scheduling Of Cascade Plants Operating On Litani River Project Lebanon, Oral Presentation, REDEC 2016 July 13-14, 2016, Notre-Dame University, Zouk Mosbeh, Lebanon.

6. Bou-Fakhreddine B. Modeling, Control and Optimization Of Cascade Hydro-Electric-Irrigation. Oral Presentation, CEDRIC Lab, CNAM de Paris, 2016.

7. Bou-Fakhreddine B. Estimating Daily Evaporation from Poorly – Monitored Lakes using limited Meteorological Data. SDEWES 2017 October 4-8, 2017, Dubrovnik, Croatia.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms / Abbreviations

AI	Artificial Intelligent
AMCP	Adjusted version of Multi-Cropping Planning problem
ANN	Artificial Neural Network
AR	Auto-Regressive
ARIMA	Auto-Regressive Integrated Moving Average
ARMA	Auto-Regressive Moving Average
ARMAX	Auto-Regressive Moving Average with Exogenous inputse
AWS	Adaptive Weighted Sum
C-FSM_MA	Constructive Fuzzy System Modeling coupled with a Moving Average filtering
C-FSM	Constructive Fuzzy System Modeling
CCF	Cross Correlation Function
CDR	Council for Development and Reconstruction
CHIP	Cascade Hydropower-Irrigation Plants
CV	Coefficient of Variation
DM	Decision Makers
DP	Dynamic Programming
DST	Decision Support Tool
EA	Evolutionary Algorithms

EDL	Electricité du Liban
EM	Expectation Maximization
FIS	Fuzzy Inference Systems
GA	Genetic Algorithm
HIP	Hydropower-Irrigation Plants
LM	Levenberg-Marquardt
LP	Linear Programming
LRA	Litani River Authority
MA	Moving Average
MCM	Millions of Cubic Meters
MCP	Multi-Cropping Planning
MILP	Mixed Integer Linear Programming
MINLP	Mixed Integer NonLinear Programming
MISO	Multiple Input Single Output
ML	Machine Learning
MNR	Multi-variate Nonlinear Regression
MTHGIS	Medium Term Hydro Generation-Irrigation Scheduling
NLP	NonLinear Programming
ACF	Autocorrelation Function
PACF	Partial Autocorrelation Function
PD	Profit Distribution
PSO	Particle Swarm Optimization
QD	Qaraoun Dam
RNLP	Relaxed Non-Linear Programming

SA	Simulated Annealing
SC	Subtractive Clustering
STHGS	Short Term Hydro Generation Scheduling
TPC-FSM	Two Phase Constructive Fuzzy System Modeling
WNN	Wavelet Neural Network
WS	Weighted Sum

Chapter 1

1. Introduction

> Electric power is everywhere present in unlimited quantities and can drive the world's machinery without the need of coal, oil, gas or any other of the common fuels
>
> *Nikola Tesla*

1.1 Overview

Cascade Hydropower-Irrigation Plants (CHIP) generate numerous inter-regional and national impacts at different levels: socioeconomic, health, institutional, environmental, ecological and cultural. The direct benefits provided by large CHIP are the provision of irrigation water, electricity generation, municipal, industrial water supply and flood control.

Nowadays, hydopower is the main contributor in the world's renewable energy production sector with a share of around 72.8% (REN21., 2015). In fact, hydroelectricity's low cost, near-zero emissions, and the ability to be dispatched rapidly to meet highly variable electricity demand have made it one of the most valuable renewable energy sources worldwide. Hydropower plants provide, in addition to clean electricity, benefits in the agriculture sector and a variety of recreational opportunities. For several years, the awareness of all these factors has created the necessity to present effective water management plans.

An efficient management approach of a hydropower-irrigation system must be capable to increase the productivity of water by continuously adjusting allocation decisions according to energy demands, irrigation requirements and hydrological status of the system. Several studies were dedicated to hydrological modeling (Porporato and Ridolf, 2001; Dibike and Solomatine, 2001; Pulido-Calvo and Portela, 2007; Huamani et al., 2011). Other studies were dedicated to irrigation modeling (Wardlaw and Barnes, 1999; Georgiou et al., 2006) and to hydropower operation modeling (Reddy and Kumar, 2006;

R. Zhang and et al., 2013; Pérez-Díaz et al., 2010). In this work, a modular approach is adopted in which each model is first treated independently. Afterward, these modules are linked together in order to exchange the output data. In fact, the first module consists in modeling hydrological regimes under conditions related to data scarcity, non-homogeneity and asymmetry. The second module provides optimal cropping patterns and irrigation profiles to Decision Makers (DM). In addition, it provides a technique to distribute profit at farmland level based on a defined cooperative policy. The third module comprises a novel cascade hydropower scheduling which was constructed using details from different studies. The outcome of the research is an optimal scheduling scheme for making the cascade reservoir system more reliable through a robust management of water resources at the watershed scale. Furthermore, the research exhibits computer models applied in a real case study in order to evaluate the performance of the suggested models under various conditions. Subsequently, it is a matter of designing a coordinated mechanism to drive forward any un-coordinated structure and inefficient operation towards an ideal centralized solution performance.

Normally, the problem of hydropower engineering is how to come out with a plan that, to a high degree of precision, represents the best solution for the hydropower-irrigation system. The optimal policy is highly connected to economic issues such as optimizing financial-revenues, satisfying energy-irrigation demands and reducing water losses. In fact, optimization of water allocation from a reservoir requires a comprehensive and deep understanding of hydropower system elements, irrigation demands, crop patterns, flood control, environmental flows, electricity generation and conflicting objectives of stakeholders.

This chapter gives a brief description covering components of a hydropower plant, research objectives, adopted methodologies and the case under study.

1.2 Components of a Hydropower-Irrigation Plant

The scope of Hydropower-Irrigation Plant (HIP) is not restricted to electricity generation and the application of water to the soil. It deals with all elements of the problem extending from the watershed passing to power generators till the agricultural zones. In fact, the HIP is a series of interconnected components that cover, Hydraulic Infrastructure, Hydrological Relations, Hydro Generation Elements and Agricultural Activities. These components, depicted in Figure 1.1, are summarized in the following:

1. **Hydraulic Infrastructure:** Hydraulic topics cover concepts such as pipe flow, Dam/Reservoir design and fluid control. However, in a HIP, different types of hydraulic infrastructure produce different magnitudes of effects. Therefore, to achieve the best water management plan, it requires having a comprehensive knowledge of the reservoir storage volume (dead, active and maximum), pipelines maximum flow rates, turbine efficiency, distribution network profile and the spillage system. During modeling and validation process, neglecting one of these terms will produce

unrealistic outcomes. Consequently, the obtained hydropower operational plan will be invalid and unreliable.

2. **Hydrological Relations:** River flow and reservoir's surface evaporation are key components in the reservoir's water balance equation (Pulido-Calvo and Portela, 2007; Valiantzas, 2006). The water balance equation is used as a basis for assessing the availability of water storage. Thus, integrating accurate hydrological estimations within the suggested hydropower-irrigation models enable DM to understand how water volume changes across time and space. As a result, informed decisions can be made regarding water releases for hydropower, irrigation and municipal. Besides that, non-productive events can be also reduced or avoided such as unnecessary spillage and floods.

3. **Hydro Generation Elements:** Hydropower plants use the potential energy of the stored water to operate hydraulic turbines. The flowing water causes the turbine shafts enclosed within the generating unit to rotate and electricity is produced. Moreover, it is known that, hydroelectric systems have high dispatchability. Typically, they have the ability to start-ramp up-shutdown within minutes, and in some cases seconds (FCH, 2017). Thereby, while less controllable energy sources (such as nuclear, thermal,...etc.) are used as baseload power plants, HIP can operate as load following and peaking power plants (Donev, 2017). It aims to soak up fluctuation in electricity demand. In fact, this issue is important since power over-generation causes electrical lines to heat up and sag affecting the whole system and leading to blackouts. On the other hand, power shortage generates less financial revenues. In such situation, it is highly recommended to fit the generated electricity with market demand.

4. **Agricultural Activities:** Farmlands serve as a platform for agricultural activities. However, for productive agricultural practices, there is a great urge to introduce new management technologies. These approaches cover right crop selection, cropping area, crop rotation, precise irrigation scheduling and profit distribution among farmers and stakeholders.

The addressed components represent a framework for a typical hydropower-irrigation plant. Nevertheless, in CHIP, the case is rather more complicated. A cascade system is formed of multi-reservoirs (fed by multiple stream inflows) interconnecting various hydropower plants through a complex water distribution network. Despite system complexity, sharing water resources, water infrastructures and reservoirs, often receive considerable attention. Such projects can generate multipurpose benefits and they are typically designed in order to maximize specific objectives such as revenues from energy generation or/and from agricultural practices.

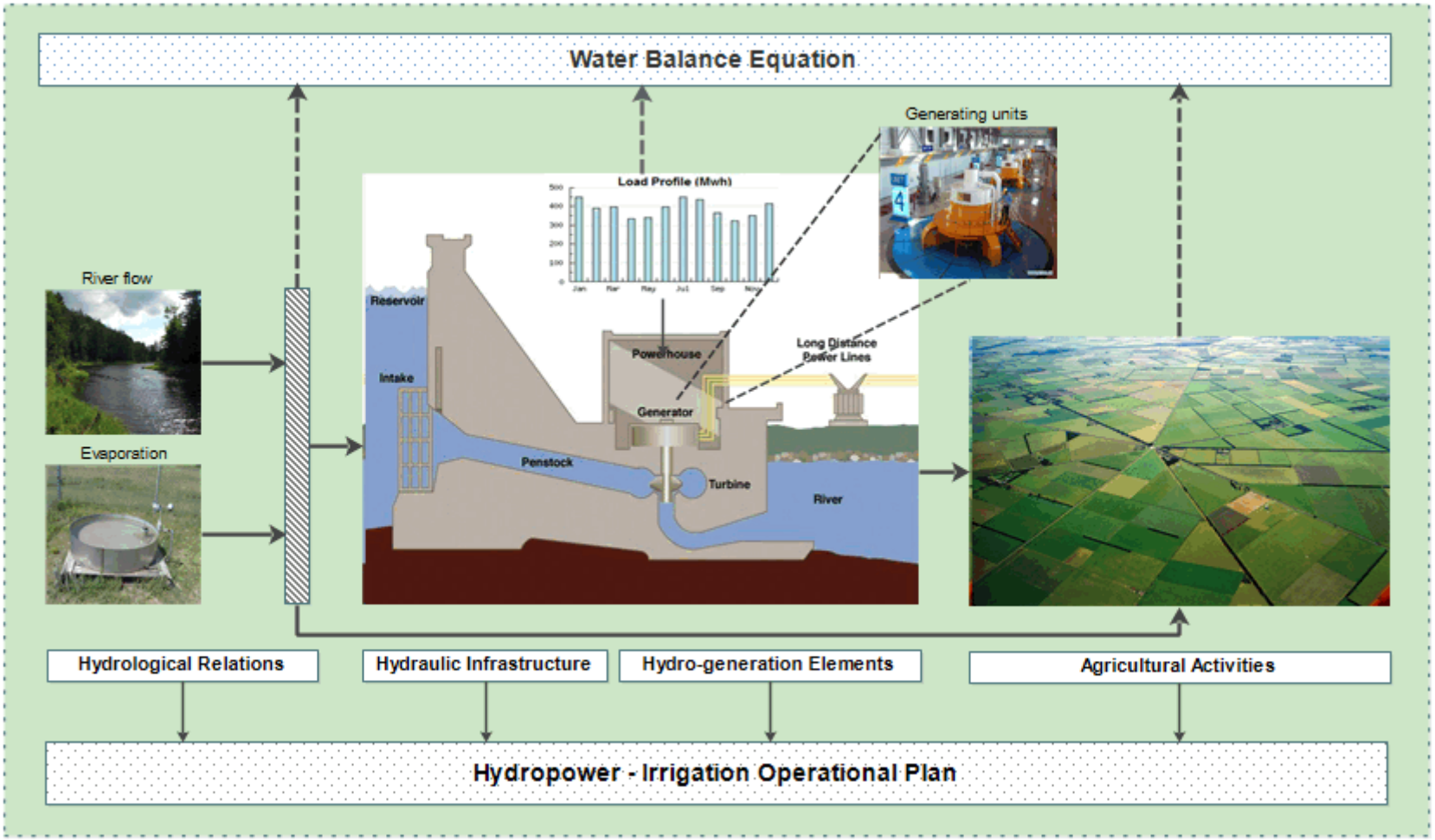


Figure 1.1 Different components of a hydropower-irrigation plant.

1.3 Research Objectives and Scope

Many developing countries suffer from energy shortage and from the degradation of agricultural areas. This is due to several reasons, mainly: poor resource management and the growing demand for water and energy. However, providing sustainable services to these countries have a significant role in promoting improvements in both energy and water sectors. A considerable portion of the scientific literature was devoted to this topic with a large number of papers. Apparently, hydropower-irrigation plants are among the most effective systems for integrated water resources development and management. They have the capability of providing a number of water related advantages. These advantages may include water supply for agriculture, domestic, industrial, hydroelectric power generation and recreation. However, to maximize the mentioned benefits, an intelligent operation is required. In fact, in this work, the operational planning master problem involves three major modules: Hydrological Predictions, Agricultural Planning and Hydropower-Irrigation Modeling.

1.3.1 Hydrological Predictions

Accurate prediction of river flow plays a vital role in reservoir operation. Nevertheless, forecasting river flow remains one of the very difficult matters in the field of hydrological sciences, especially when the target system is characterized by a nonlinear dynamic process under stochastic and chaotic conditions (Huamani et al., 2011). Besides that, in developing countries, this problem is exacerbated by the fact that hydro-meteorological data collection and sharing are not straightforward. Data sources, in that region, are typically limited and less (or no) local data often exist or are publicly available. To overcome these obstacles and based on the bibliographic review, Fuzzy inference seemed an interesting solution. Thereby, throughout the hydrology part, Fuzzy System Modeling (FSM) method is adopted to forecast river flow. The modeling procedures considered two real situations:

1. During Fuzzy modeling the literature (Huamani et al., 2011; Mehmmet, 2009; Galavi and Shui, 2012) did not consider different data drawbacks related to scarcity and non-homogeneity. In this work, using suitable data pre-processing methods, the Fuzzy model is treated in order to exploit the limited and heterogeneous hydro-meteorological data. However,

2. in a case of hydro-meteorological measurements asymmetry (the climatic measurements exist for a shorter period in comparison with the hydrological ones), the river flow forecast uses a hybrid model based on Two-Phase Constructive Fuzzy System Modeling.

Another topic of vital importance in hydropower planning is lake evaporation. In fact, open water evaporation is influenced by several meteorological parameters such as: irradiance, soil temperature, relative humidity, atmospheric pressure, and wind speed (Penman, 1948). Unfortunately, up to date, reliable evaporation approximations are extremely difficult to obtain. This problem is intensified in poorly monitored lakes. Thereby, there is an urge to find an effective evaporation model that is less

demanding in input requirements. Here, the suggested technique involves only major three weather variables: Temperature, Relative Humidity and Dew point.

1.3.2 Agricultural Planning

In this part, two mathematical models are introduced. A Multi-Cropping Planning (MCP) model that aims to optimize agricultural revenue by right crop selection, precise irrigation scheduling taking into consideration water availability at each stage. However, the other one is the Profit Distribution (PD) model. It is responsible for distributing economical benefits among farmers based on a defined cooperative policy taking into consideration the physical nature of the farmland and the carried management practices. For both models, an investigation is carried in order to select the most suitable method for solving the encountered optimization problems.

1.3.3 Hydropower Modeling

Hydropower exhibits an important linkage between energy and water. Although being a mature technology when compared to other water related renewable energy resources, hydropower has still a significant potential mostly in cascade systems. In fact, to benefit from an existing water resource as far as possible, it is common to notice several hydropower plants constructed on the same river and sharing a common water lake (Yildiran et al., 2015). Therefore, on the same river, new plants can be sub-joined and old ones can be upgraded in order to increase electricity and agricultural productions. It can only be achieved through intelligent operations and coordination. Otherwise, unnecessary spillages may occur and water may be wasted without being utilized neither for electricity generation nor for irrigation. This research suggests a decision-support tool. It is based on two stage mathematical programming models for optimal operation of a cascade hydropower-irrigation system. At stage one, a model for Medium Term Hydro Generation-Irrigation Scheduling (MTHGIS) is derived. The aim is minimizing a multi-objective functional concerning power, irrigation and municipal water deficits. Its resolution provides a general overview of the cascade operation and boundary conditions for stage two problem. In fact, this problem deals with Short Term Hydro Generation Scheduling (STHGS). Its goal is to optimize revenues while prioritizing critical hydraulic, electrical, irrigation and maintenance constraints.

Overall, a wise commitment schedule and proper power distribution among hydropower plants can decrease power and irrigation deficits significantly and simultaneously increasing the safety and reliability of the cascade system.

In this research, each of the suggested modules is treated separately. The advantage of the modular approach is the ability to go into more details in each sub-field and the ability to be independently

enhanced and updated. However, during the research progress, different modules are linked through the backward and forward exchange of output data.

1.4 Methodology

In this section, the objective is to implement a suitable method for medium and short term optimal hydropower-irrigation planning. In fact, to achieve this aim, methods involving and joining data mining and mathematical programming will serve as a base for a Decision Support Tool (DST). This multi-functional tool consists in operation planning will be capable of enhancing the performance of a hydropower-irrigation system on several levels starting with appropriate water releases, scheduled irrigation and cropping pattern. The following list represents the methodology main elements:

Hydrological Modeling

1. Data Collection
   (a) Meteorological data (rainfall, humidity, temperature, dew point)
   (b) Hydrological data (river flow, evaporation rates)
2. Data Analysis and Data Processing
   (a) Data transformation (Log-Transformation, Standardization, Normalization)
   (b) Noise treatment using Moving Average (MA) filter
   (c) Correlation analysis (Autocorrelation, Partial Autocorrelation, Cross-Correlation Function)
3. River Flow Forecasting Models
   (a) Auto-Regressive (AR) method
   (b) Constructive Fuzzy Systems Modeling (C-FSM) method
   (c) Hybrid approach based on Two-Phase Constructive Fuzzy Systems Modeling (TPC-FSM)
   (d) Performance evaluation of the suggested models
4. Lake Evaporation Estimates
   (a) Nonlinear Regression model based on Magnus formula
   (b) Multi-variate Nonlinear Least Square (NLS) method
   (c) Levenberg-Marquardt algorithm

(d) Validation and testing of the proposed model

Agricultural Modeling

1. Multi-cropping system implemented under full and deficit irrigation
   (a) Mathematical modeling (Linear and Nonlinear Programming)
   (b) Meta-heuristic algorithms (Simulated Annealing, Particle Swarm Optimization)
   (c) Testing models using experimental and real data
2. Profit Distribution (PD) within farmlands
   (a) Introducing the problem as a nonlinear mixed integer programming model
   (b) Linearization is based on algebraic methods
   (c) Simulations and results analysis

Hydropower Modeling

1. Efficient load distribution among working units within the hydropower plant
2. Medium term cascade hydropower scheduling
   (a) Nonlinear Multi-Objective Optimization model that involves minimizing power and irrigation deficits
   (b) Model's resolution: Weighted Sum approach
   (c) Simulation and results analysis
3. Short term cascade hydropower scheduling
   (a) Presenting a Nonlinear Mixed Integer Programming model that involves minimizing the revenue losses of the hydropower operation
   (b) Linearization is based on algebraic and numerical methods
   (c) Model's validation and evaluation

This research addressed an artificial intelligence based approach for the control of cascade hydropower-irrigation plants. In fact, Machine Learning (ML) that involves correlation analysis and Fuzzy inference techniques are used to capture the hidden knowledge in the river flow data. The goal is to extract trends and behavior patterns that would allow improvements to hydropower operation.

On the other hand, the recent development of meta-heuristic optimization algorithms such as Evolutionary Algorithms (EA) (Xue-Zhen et al., 2010) and Simulated Annealing (SA) (Georgiou et al., 2006) have emerged as an alternative approach to overcome some of the limitations of classical techniques like Linear Programming (LP), Dynamic Programming (DP), Non-Linear Programming (NLP) for solving problems concerning Multi Crop Planning (MCP). In this research, the problem-solution approach is as follows: at first preliminary mathematical tools are presented involving existence, benchmark linear models and a relaxation formulation, second two meta-heuristic algorithms Simulated Annealing (SA) and Particle Swarm Optimization (PSO) are implemented as a numerical technique for solving the MCP problem. The particularity of this approach consists in using the solution of the linear problem as an initial guess for the SA, while for PSO, the particle swarm is initiated in the neighborhood of that solution. Regarding the PD problem, the limited availability of meta-heuristic tools and the dullness of writing the code from scratch, testing and evaluation, turned our attention to an alternative approach. It is a matter of taking advantage of the availability of the Mixed Integer Linear Programming (MILP) solvers in the market. A favorable way to solve Mixed Integer Non-Linear Programming (MINLP) model is by linearizing through a suitable change of variable technique to handle nonlinear terms.

At the final phase of the modeling process, knowing irrigation and energy demand can play a key role in adjusting the water discharges to avoid over power production or unwanted irrigation releases. The operational plan will be capable of maintaining a continuous and reliable energy production taking into account the crops water demand. In the medium term scheduling problem, the attempt is to find a solution that has the ability to establish a relative balance between different power-irrigation objectives. The solution procedure for solving the multi-objective problem has considered the prior method which is known as Weighted Sum (WS). However, for the short term planning model, the difficulty of resolving the Mixed-Integer Non-Linear Programming (MINLP) problem is overcame by reformulating it into a Mixed Integer Linear Programming (MILP) problem using ordinary integer algebra and numerical techniques. The outcome is a distributed control structure that was implemented using multiple software frameworks. The commercial package LINGO and MATLAB were used for mathematical optimization while Microsoft Excel was employed to develop the Human Machine Interface (HMI) for sequential control.

1.5 Case Study: Litani River Project - Lebanon

The case study takes a closer look at how the newly developed models are performing within Litani river and Litani Project. In fact, the hydrological models described in the research are tried out in the Litani river, while irrigation and hydropower models are applied to the Litani Project. The Litani basin is selected as it represents a typical case for basins that are poorly monitored. This situation is encountered in many parts of the developing world.

Litani is the longest river in Lebanon reaches a length of 170 *km*. Its watershed covers an area of 2160 km^2 and it is fed by an average level of rainfall around 764 *MCM* per year (LRA, 2016). It rises near the ancient city of Baalbeck in the central Bekaa Valley, 85 *km* east of the Capital Beirut. It flows southward for 100 *km* or so, before bending sharply toward the west, entering the Mediterranean at Kasmieh just north of the city of Tyre. In the late 1950s, a major development on the Litani River known as Litani Project involved constructing an artificial Qaraoun Dam (QD) and its associated structures: secondary reservoirs (Anan and Joun), hydropower plants and an irrigation system. The Qaraoun lake with a storage capacity of 220 *MCM* is located in the middle reach of the Litani River. It diverts the river flow through a system of tunnels inter-connecting three hydropower plants (Figure 1.2): Markaba (34 *MW*), Awali (108 *MW*) and Charles Helou (48 *MW*).

- **Markaba:** the first power station in the network. It is located 660 *m* above sea level and 11 *km* away from the Qaraoun reservoir. It takes water in from the Qaraoun lake through Markaba tunnel cutting an underground distance 6.4 *km* along the right riverbank.

- **Awali:** it is located in the basin of Bisri River, at 228.50 *m* above sea level. Its waterfall is located at an altitude of 400 *m*. Water from the Qaraoun lake, after it discharges from the turbines of Markaba station, join the water of Ain Zarqa and other springs, through a 17 *km* tunnel crossing Jabal Niha-Jezzine all the way to its pool: Anan lake (capacity 170, 000 m^3). The basic role of Awali plant is regulating the frequency of the public transportation network related to the difference in consumption needs, within its power supply capacity.

- **Charles Helou :** it is located on the left bank of Awali river, 32 *m* above sea level. Its waterfall is located at an altitude of 194 *m*. It pulls, through a 6800 *m* tunnel, the water discharged by the turbines of Awali power plant, in addition to the water flowing from Bisri/Awali river to the balancing lake (capacity 300,000 m^3) at the foot of the Awali power plant.

In the irrigation sector, Litani Project is responsible for irrigating more than 1,400 hectares of agricultural lands in the Bekaa Valley through Canal 900 as well as 36,000 hectares of agricultural lands in the South (Lebaa and Kasmieh projects) (Source: LRA, 2016).

1.6 Dissertation Organization

This dissertation has been structured into five chapters. The path for connecting individual chapters begins from the hydrological modeling to predict future river flow and extends to estimate lake surface evaporation from poorly monitored lakes. Afterward, irrigation and agricultural models are introduced that covers multi-crop planning approach implemented under deficit/non-deficit irrigation and profit

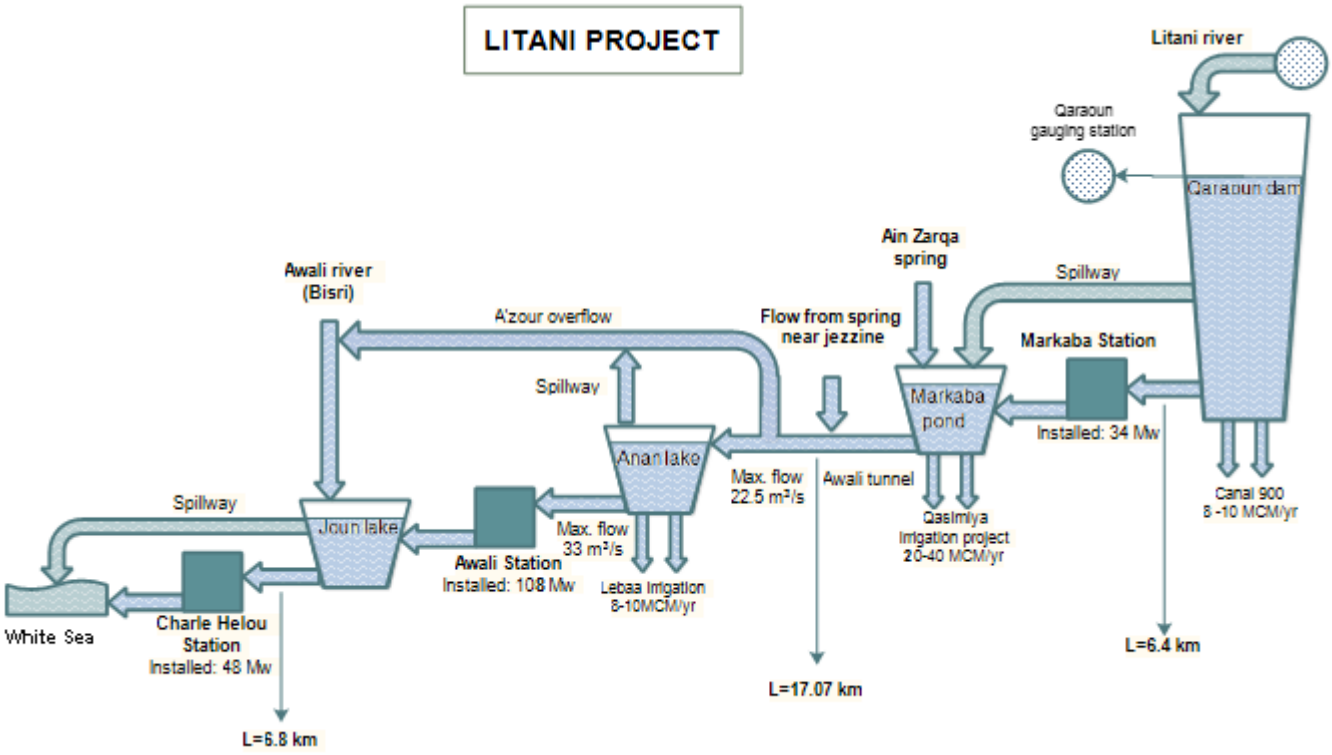


Figure 1.2 Litani Project Profile.

distribution among active farmers. Another essential part of the research is the cascade hydropower operation. Its role is to present a framework for decision making in energy sector in the medium and short run operation. In fact, all these topics are addressed in the following chapters:

Chapter 2 presents three hydrological models. The first model deals with daily river flow forecast in a region where meteorological and hydrological data are insufficient, inaccessible and sometimes unreliable. The data-driven model is based on Constructive Fuzzy Systems. The second model is based on hybrid modeling implemented using a Two-Phase Fuzzy Inference. The modeling approach for river flow forecasting is capable to deal with hydro-meteorological measurements asymmetry (climatic measurements exist for a shorter period in comparison with the hydrological ones). The third model aims to estimate daily evaporation from poorly – monitored lakes using limited meteorological data.

Chapter 3 introduces a multi-crops planning optimization models for cropping pattern and water allocation. The problem-solution approach of the suggested models is discussed extensively using experimental and real data. In the final section of Chapter 3, it proposes a profit distribution model based on a defined cooperative policy to distribute profit among active farmers.

Chapter 4 exhibits mathematical programming models for medium/short term scheduling of cascade hydropower-irrigation plant. It aims to design a coordinated mechanism to drive forward any un-coordinated structure and inefficient operation towards an optimal centralized control.

Chapter 5 summarizes the most relevant insights and conclusions of this thesis. Future research work and possible enhancements are also identified.

Chapter 2

Hydrological Models

> Water is the driving force of all nature
>
> *Leonardo da Vinci*

2.1 Introduction

Hydrological models are very useful tools that are widely used in hydropower operation planning. However, availability of large amounts of data for model's training and validation is a problem in the developed countries. According to Zemadim, developing countries face technical and financial constraints that restrain both data gathering and sharing efforts (Zemadim et al., 2014). The technical obstacles are related to the watershed size being monitored, and to the availability of both monitoring equipment and skilled labor. On the other hand, financial constraints are tightly linked not only to the expenses related to sophisticated monitoring systems, but also to the ongoing cost of supervision and maintenance of the system. All these factors present a major concern to hydrologists regarding the quality and quantity of the retrieved measurements. Thereby, the reliability of hydrological predictions is limited in these countries, because the local hydro-meteorological data are often sparse, scarce, and most of the time of poor quality. Despite these limitations, there is a great urge to develop new forecasting approaches. They must be capable of exploiting the available data efficiently and provide reliable results to be used by hydropower plant operators, farmers, and most importantly by DM. Accurate hydrological predictions are of vital importance for efficient reservoir water management and control. In fact, in this research, the main focus is to estimate two crucial quantities: river flow and evaporation rates. Forecasting river flow is one of the very difficult issues in hydrological sciences because it is characterized by a dynamic, uncertain and nonlinear problem (Huamani et al., 2011). This problem deals with a system that receives thousands of inputs interacting in a complex and noisy environment. Concerning water evaporation estimations, the drawback also persists. Many researchers

developed different models to estimate evaporation (Penman, 1948; Penman, 1963; Priestley et al., 1972; Linacre, 1977). Unfortunately, up to date, accurate approximations are extremely difficult to obtain. This is due to the complex interactions between many interrelated meteorological factors. The main scope of this research is to present, analyze and discuss various hydrological models. In fact, these models were implemented using different data mining methods, whereas several scenarios were tested in the Litani river in order to assess their performance.

List of Symbols

Notation	Description	Unit
AR model		
y_k	Output of the AR model at time k	-
e_k	White-noise disturbance at time k	-
ρ	Autocorrelation coefficient of the random process y_k	-
C-FSM model		
$\mathbf{x}^k$	Input vector of C-FSM model at instant k	-
$\hat{y}^k$	Output of the C-FSM model at instant k	-
c_i	Center of partition i	-
V_i	Covariance matrix of partition i	-
g_i^k	Membership degree	-
TPC-FSM model		
Y_k	Any given time series	-
L_k	Linear component of Y_k	-
N_k	Nonlinear component of Y_k	-
d_k	Residual at instant k	-
NLS model		
$\mathbf{x}$	Input vector of NLS model	-
y	Output of the of NLS model	-
f	Model curve function	-
Miscellaneous variables		
P	Rainfall	mm
T	Temperature	^oC
Q	River flow	m^3/sec
D	Dew point	^oC
RH	Relative humidity	$\%$
T_{max}	Maximum temperature	^oC
T'	Daily average temperature	^oC
E_0	Lake evaporation rate	mm/day
R_S	Solar radiation	$MJ/m^2/day$
E_P	Pan evaporation rate	mm/day
Miscellaneous parameters		
R_A	Extraterrestrial radiation	$MJ/m^2/day$
K_{R_S}	Empirical radiation adjustment coefficient	$^oC^{-0.5}$

In order to satisfy water requirements in different sectors such as hydropower, agriculture and recreation, a water program needs to be developed. It should be based on accurate hydrological models and forecasts. Here in this chapter and as a first step, several approaches are presented that cover Auto-Regressive (AR) methods, Constructive Fuzzy System Modeling (C-FSM), Constructive Fuzzy System Modeling coupled with Moving Average filter (C-FSM_MA), Two-Phase Constructive Fuzzy System Modeling (TPC-FSM) and Nonlinear Least Square (NLS) approach. The first four methods are used for streamflow forecasting while the last one is utilized for evaporation estimation. Afterward, different pre-processing tools are exhibited for data treatment and input selection. Nevertheless, the core of the chapter is the actual application of these mathematical models. To achieve this aim, data were gathered from the Litani river basin, then they were subjected to a preliminary analysis for better understanding of the hydrological behavior in the river catchment. Subsequently, the adopted methodologies were tested and evaluated. In fact, the successful implementation of the suggested approaches into practice had encouraged us to integrate them later into the hydropower DST. It is manifested through the simulation of water availability at QD during planning and management of water resources. The framework of Chapter Two is illustrated in Figure 2.1.

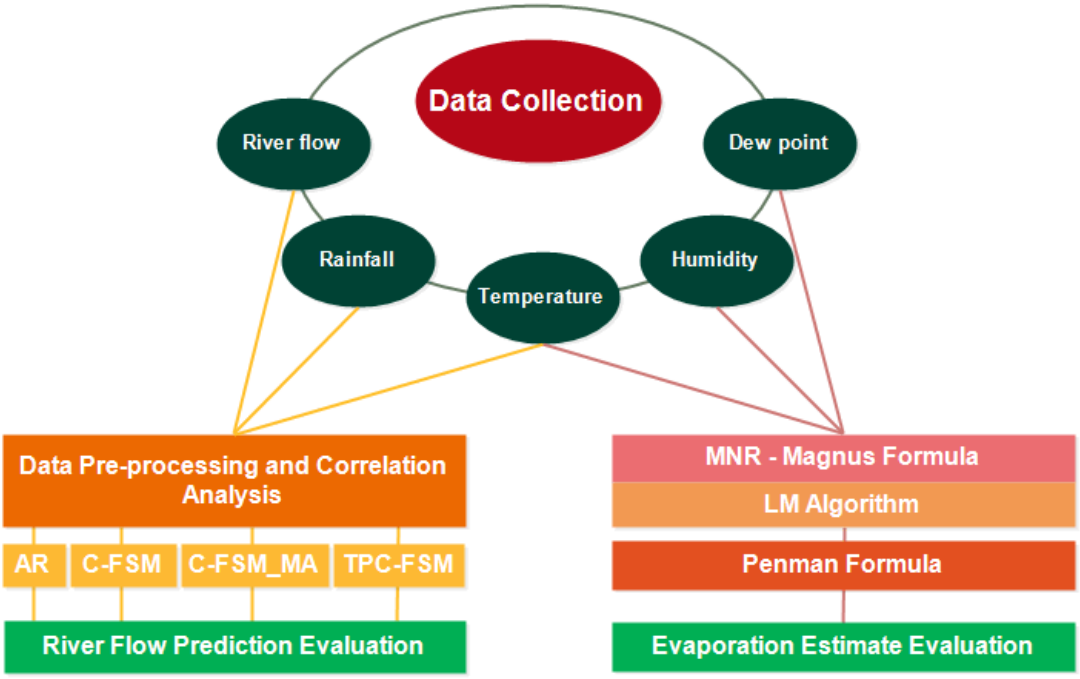


Figure 2.1 Hydrological modeling framework.

2.2 Suggested Hydrological Models

Over the past few decades, several types of stochastic models have been suggested for hydrological time series modeling such as Box and Jenkins methods for Auto-Regressive (AR), Auto-Regressive

Moving Average (ARMA), Auto-Regressive Integrated Moving Average (ARIMA), Auto-Regressive Moving Average with Exogenous inputs (ARMAX) models (Box and Jenkins, 1970). They were generally utilized in the linear sense for estimating future river flow. Later on, several studies were dedicated to the formularization and the development of nonlinear river flow models that aim to improve the quality of hydrological forecasting. In fact, Porporato dealt with local linear models with time-dependent parameters (Porporato and Ridolf, 2001), whereas Dibike and Pulido-Calvo have considered data-driven nonlinear models based on Artificial Neural Network (ANN) (Dibike and Solomatine, 2001; Pulido-Calvo and Portela, 2007) or on Wavelet Neural Network (WNN) (Cuia et al., 2015). In the work of Huamani, the followed methodology was based on Fuzzy inference systems (Huamani et al., 2011). However, others have discussed extensively Neuro-Fuzzy hybrid models that have the capability of preserving the learning abilities of ANNs and the reasoning of Fuzzy systems (Coulibaly and Baldwin, 2005; Firat, 2008; Kisi et al., 2012). Although a variety of forecasting approaches have been successfully formulated, choosing the proper model to accurately predict river flows still up to date imposes a challenge to hydrologists. However, in this research, the main focus is on Auto-Regressive methods and Constructive Fuzzy System Modeling.

2.2.1 Auto-Regressive Model

AR structure

An Auto-Regressive model (AR) defines the next random variable in a sequence as an explicit linear function of previous ones within a time frame. The structure of AR model of order p is given in Equation 2.1 (Wei, 2006):

$$y_k + a_1 y_{k-1} + ... + a_p y_{k-p} = e_k \tag{2.1}$$

where y_k is the output at time k, $a_1, ..., a_p$ are the parameters of the AR model to be estimated from the data, $y_{k-1}, ..., y_{k-p}$ are the previous outputs on which the current output depends and e_k is the white-noise disturbance.
As a matter of fact, the name “autoregressive" comes from the fact that the output y_k is regressed on the past values of itself.

Optimization Algorithm

There are many ways to estimate the coefficients of Equation 2.1, such as the ordinary least squares procedure, method of moments, Markov chain - Monte Carlo or Yule–Walker methods. However, in this work Yule–Walker equations are used to relate the Auto-Regressive model parameters to the Autocorrelation coefficient ρ of the random process y_k (Wei, 2006).

The values of $a_1, ..., a_p$ are determined by solving the matrix Equation 2.2:

$$\begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \rho_2 & \rho_1 & \cdots & \rho_{p-3} \\ \vdots & \vdots & \ddots & \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{pmatrix} = - \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_p \end{pmatrix} \tag{2.2}$$

2.2.2 Constructive Fuzzy System Modeling

C-FSM Structure

Typically, the Multiple Input Single Output (MISO) model structure based on first order Takagi-Sugeno Fuzzy system is composed of a set of M fuzzy rules. Its representative power is manifested through its capability of describing a highly complex nonlinear system using a small number of simple rules. In this research, Constructive Fuzzy System Modeling is adopted (Luna et al., 2007) and it is described in the following:
Let us denote by $\mathbf{x}^k = [x_1^k \; x_2^k \; ... \; x_p^k] \in \mathbb{R}^p$ the input vector at instant k , $k \in \mathbb{Z}^+ - \{0\}$; $\hat{y}^k$ is the output of the model, for a given $\mathbf{x}^k$. The aim is subdividing the input space into M fuzzy sub-regions and approximating the system in each subdivision by a simple linear model. Each partition is defined by its center $c_i \in \mathbb{R}^p$ and its covariance matrix $V_i \in \mathbb{R}^{p \times p}$, whereas a data point can belong to all partitions with different membership degree g_i^k that lies between 0 and 1, such that the sum of all membership values is equal to 1. Afterward, an IF-THEN rule is set to each sub-region and it is defined in the form:

R_i : **IF** $< \mathbf{x}^k$ belong to the i^{th} region with a membership degree $g_i^k >$ **THEN**

$$y_i^k = \varphi^k \times \theta_i^T \tag{2.3}$$

where $\varphi^k = [1 \; x_1^k \; x_2^k \; ... \; x_p^k] \in \mathbb{R}^{p+1}$, and $\theta_i = [\theta_{i0} \; \theta_{i1} \; ... \; \theta_{ip}] \in \mathbb{R}^{p+1}$ is the coefficients vector (parameter) for the local model (Figure 2.2). Every input pattern has a membership degree associated to each sub-region of the input space and is calculated by the formula:

$$g_i(x^k) = g_i^k = \frac{\alpha_i P\left[i \left| x^k \right.\right]}{\sum_{q=1}^{M} \alpha_q P\left[q \left| x^k \right.\right]} \tag{2.4}$$

where α_i is a positive parameter that is considered as an indirect measure of the relevance of each rule and satisfies $\sum_{i=1}^{M} \alpha_i = 1$. $P\left[i\left|x^k\right.\right]$ is the conditional probability of activating the i^{th}rule given the input vector $\mathbf{x}^k$ and is defined as:

$$P\left[i\left|x^k\right.\right] = \frac{1}{(2\pi)^{p/2}\det(V_i)^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}^k - c_i)V_i^{-1}(\mathbf{x}^k - c_i)^T\right\}$$

where $\det(.)$ is the determinant function.

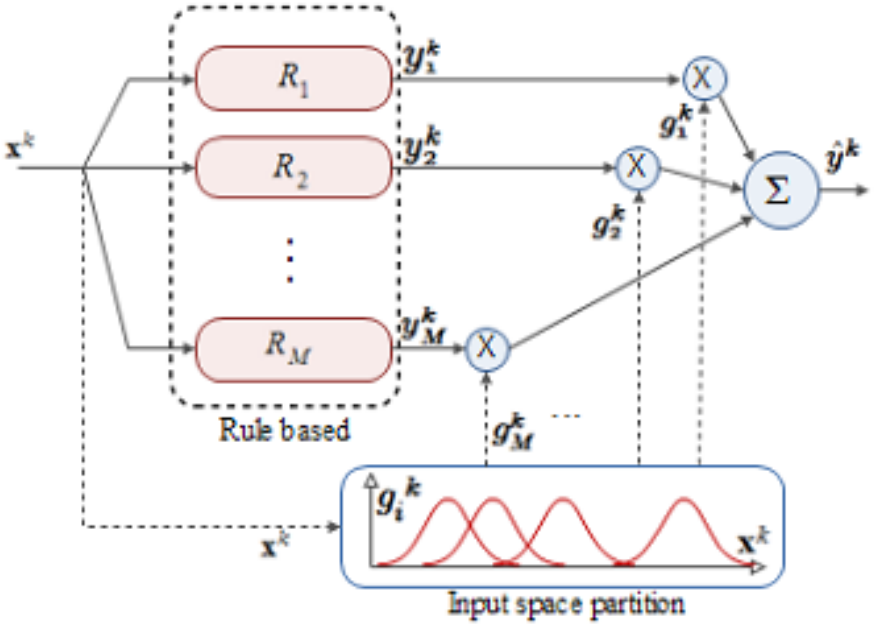


Figure 2.2 C-FSM general structure with M fuzzy rules (Luna et al., 2007).

The final model output is computed by a nonlinear weighted average of the aggregated local outputs and their respective membership degrees. Thus, the estimated output value of the global model for the future time instant k is:

$$\hat{y}^k = \sum_{i=1}^{M} g_i^k y_i^k \tag{2.5}$$

Optimization Algorithm

The constructive offline learning process for building a FIS model determines automatically the number of fuzzy rules as well as its internal parameters c_i , V_i , θ_i and α_i for $i = 1,...,M$. In fact, the procedure is carried over two stages: model initialization and structure modification stage (Luna et al., 2007).

At stage one, the model is initialized by using the well known Subtractive Clustering (SC) algorithm (Chiu, 1994). Its goal is to determine the initial structure of the fuzzy system that will serve as a starting point for the next stage. The input-output pattern constructed from the available historical data are fed into the SC routine that is available in MATLAB package.

The function returns the cluster centers in the matrix C and the vector S which contains the sigma values that specify the range of influence of a cluster center in each of the data dimensions. r_a is a real number ranging between 0 and 1 that specifies the cluster center's range of influence, assuming that data falls within a unit hypercube.
Suppose the initial number of rules M^0 is the length of the matrix C. Then, the C-FSM structure is initialized as follows:

- $c_i^0 = C_i|_{1...p}$ the first p components of the i^{th} center found by the SC algorithm.
- $V_i^0 = r_a^2 I$, the covariance matrix codifying the spread where r_a is the spread parameter that is used in the SC algorithm and I is the $p \times p$ identical matrix.
- $\theta_i^0 = \left[C_i^{p+1}\ 0...0\right]$, C_i^{p+1}: last $p+1$ component determined by the SC algorithm.
- $\sigma_i^0 = 1.0$, the initialized standard deviation for each local output y_i^k.
- $\alpha_i^0 = 1/M^0$.

Once the model initialization is completed, parameters are re-adjusted based on EM algorithm with the objective of maximizing the log-likelihood $\mathscr{L}$ (Equation 2.6) of the observed values of y^k at each step M of the learning process.

$$\mathscr{L}(D,\Omega) = \sum_{k=1}^{N} \ln\left(\sum_{i=1}^{M} g_i(x^k,\mathbf{C}) \times P(y^k \left| x^k, \theta_i\right.)\right) \tag{2.6}$$

where $D = \left\{\left(x^k, y^k\right); k = 1,...,N\right\}$ is the training set, Ω contains all the model parameters and bold $\mathbf{C}$ contains the centers and the covariance matrix parameters. However, for maximizing $\mathscr{L}$ it is necessary to estimate h_i^k: the posterior probability of x^k belong to an active region of the i^{th} local model that is computed for $i = 1,...,M$ by:

$$h_i^k = \frac{\alpha_i P\left[i \left| x^k\right.\right] P\left[y^k \left| x^k, \theta_i\right.\right]}{\sum_{q=1}^{M} \alpha_q P\left[q \left| x^k\right.\right] P\left[y^k \left| x^k, \theta_q\right.\right]} \tag{2.7}$$

The conditional probability $P\left[y^k \left| x^k, \theta_i\right.\right]$ is defined as:

$$P\left[y^k \left| x^k, \theta_i\right.\right] = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\left[y^k - y_i^k\right]^2}{2\sigma_i^2}\right) \tag{2.8}$$

The variance of the local output y_i^k can be estimated by:

$$\sigma_i^2 = \left(\sum_{k=1}^{N} h_i^k \left[y^k - y_i^k \right]^2 \right) / \sum_{k=1}^{N} h_i^k \tag{2.9}$$

The EM algorithm for finding the parameters is summarized by:

1. E step: Estimate h_i^k via Equation 2.7

2. M step: Maximize Equation 2.6 and update the model parameters:

$$\alpha_i^{new} = \frac{1}{N} \sum_{k=1}^{N} h_i^k \tag{2.10}$$

$$c_i^{new} = \left(\sum_{k=1}^{N} h_i^k \mathbf{x}^k \right) / \sum_{k=1}^{N} h_i^k \tag{2.11}$$

$$V_i^{new} = \left[\sum_{k=1}^{N} h_i^k (\mathbf{x}^k - c_i)^T (x^k - c_i) \right] / \sum_{k=1}^{N} h_i^k \tag{2.12}$$

   for $i = 1, ..., M$. An optimal solution for θ_i is obtained by solving the equation:

$$\sum_{k=1}^{N} \frac{h_i^k}{\sigma_i^2} \left(y^k - \varphi^k \times \theta_i^{new} \right) . \varphi^k = 0 \tag{2.13}$$

   After adjusting the parameters, $\mathscr{L}(D, \Omega)$ is re-calculated and saved as $\mathscr{L}_{new}(D, \Omega)$.

3. Convergence: Stop the process if:

$$\mathscr{L}_{new}(D, \Omega) - \mathscr{L}_{old}(D, \Omega) < \varepsilon$$

   else return to step 1.

2.2.3 Constructive Fuzzy System Modeling coupled with Moving Average

To enhance performance of the forecasting model, the C-FSM_MA adopts the C-FSM structure, whereas its inputs are fed by treated data obtained through the use of the Moving Average (MA) filter presented in Subsection 2.3.2.

2.2.4 Two-Phase Constructive Fuzzy System Modeling

Relationships in real-world are usually involved with a linear and a nonlinear component. Thus, it is more practical for a time series to encompass linear and nonlinear models. Such modeling procedure gives the ability to capture different aspects of the underlying patterns (Zhang, 2003).

In fact, Zhang considered for a given time series Y_k, a composition of linear and nonlinear patterns as shown in Equation 2.14 (Zhang, 2003), where L_k is the linear and N_k is the nonlinear component:

$$Y_k = L_k + N_k \tag{2.14}$$

In his approach, an Auto-Regressive Integrated Moving Average (ARIMA) model was employed to map linear patterns whereas, ANN is employed to perform nonlinear mappings in the residual data. However, in the Two-Phase Constructive Fuzzy System Modeling (TPC-FSM) case, the procedure is as follows: at phase one, C-FSM infers the linear component L_k. Then, the residuals from the linear model will contain only the nonlinear relationship. Denote by d_k the residual at time k from the linear model, then:

$$d_k = Y_k - \widehat{L}_k \tag{2.15}$$

where $\widehat{L}_k$ is the forecast value for time k.
At phase two, the residuals are also modeled using C-FSM, such that the nonlinear relationship can be discovered. The C-FSM model for the residuals will be of the form:

$$d_k = f(d_{k-1}, d_{k-2}, ..., \mathbf{x}_k, \mathbf{x}_{k-1}, ...) + e_k \tag{2.16}$$

where f is a nonlinear function determined by the Fuzzy inference, $\mathbf{x}_k$ is the input vector and e_k is the random error. Designating by $\widehat{N}_k$ the forecast from Equation 2.16, the combined forecast will be:

$$\widehat{Y}_k = \widehat{L}_k + \widehat{N}_k \tag{2.17}$$

2.2.5 Multi-variate Nonlinear Least Square Method

NLS Structure

Nonlinear Least Squares (NLS) is the form of least squares analysis used to fit a set of m observations $(\mathbf{x}_1, y_1)$, $(\mathbf{x}_2, y_2)$, ..., $(\mathbf{x}_m, y_m)$ with a Multi-variate Nonlinear Regression (MNR) model $y = f(\mathbf{x}, \theta)$, where $\theta = (\theta_1, \theta_2, ..., \theta_n)$ is an unknown parameter. The aim is to find the vector θ that minimize the sum of squares:

$$\min_{\theta} \; S(\theta) = \sum_{i=1}^{n} [y_i - f(\mathbf{x}_i, \theta)]^2 \tag{2.18}$$

where the residuals (errors) e_i given by:

$$e_i = y_i - f(\mathbf{x}_i, \theta) \tag{2.19}$$

should follow a Normal distribution with zero mean.

Optimization Algorithm

The method used to estimate the model parameters is the Levenberg-Marquardt Algorithm (LM). LM algorithm is an iterative procedure that solves nonlinear least square problems by combining the Steepest Descent and Gauss-Newton methods. To initiate the minimization process, the user has to provide an initial guess for the parameter vector and in each iteration step, the parameter vector θ is replaced by a new estimate $\theta + \delta$. To determine the value of δ, the functions $f(\mathbf{x}_i, \theta + \delta)$ are approximated by their linearizations:

$$f(\mathbf{x}_i, \theta + \delta) \approx f(\mathbf{x}_i, \theta) + J_i \delta \tag{2.20}$$

J is the Jacobian matrix whose i^{th}row is $J_i = \frac{\partial f(\mathbf{x}_i, \theta)}{\partial \theta}$. Thus, the convergence is achieved at each step by calculating δ that minimizes:

$$S(\theta + \delta) = \|Y - f(\theta) - J\delta\|^2 \tag{2.21}$$

Taking the derivative of $S(\theta + \delta)$ with respect to δ and setting the result to zero gives:

$$(J^T J)\delta = J^T (Y - f(\theta)) \tag{2.22}$$

where the terms f and Y are vectors with i^{th} component $f(\mathbf{x}_i, \theta)$ and y_i respectively. In fact, it is a set of linear equations which can be solved for δ. Levenberg's contribution is to replace this equation by a "damped version":

$$(J^T J + \lambda I)\delta = J^T (Y - f(\theta)) \tag{2.23}$$

I is the identity matrix.

The (non-negative) damping factor λ is adjusted at every iteration according to the following: If the reduction of S is rapid, the algorithm sets $\lambda = \lambda/10$; otherwise, the algorithm sets $\lambda = \lambda \cdot 10$ (Mathworks, 2013).

The LM algorithm iterative process is terminated when $\left\|J^T (Y - f(\theta))\right\|_\infty$ drops below a threshold ε or when the maximum number of iterations $k_{\max}$ is completed.

2.3 Data Processing – Correlation Analysis

The available weather and streamflow measurements corresponding to the continuous period were split into two subsets:

1. A training data set is used to select inputs and to estimate model's parameters.
2. A testing set formed of the remaining data utilized to test the performance of the suggested models.

2.3.1 Standardization/Normalization

Standardization is crucial in the improvement of both Fuzzy and Auto-Regressive models (Firat, 2008; Luna et al., 2007; Wu et al., 2012). In fact, all series presented here: rainfall (P), temperature (T) and river flow (Q) have a periodic and seasonal components. They were removed by standardizing the original data through the following transformation:

$$z_m^k = \frac{y_m^k - \mu(m)}{\sigma(m)} \tag{2.24}$$

where z_m^k is the stationary version of the time series y^k at instant k, $\mu(m)$ is the monthly average value, $\sigma(m)$ is the monthly standard deviations and m is the month number. Moreover, in the course of Fuzzy System Modeling, the model is initialized using Subtractive Clustering (SC) with spread radius $r_a \in [0,1]$, Thus, it is necessary to re-scale or normalize the trained data set within a unit hypercube, using the formula:

$$Z_{norm}^k = \frac{z^k - z_{\min}}{z_{\max} - z_{\min}} \tag{2.25}$$

where Z_{norm}^k is the normalized data at time k, z^k is the observed value, $z_{\min}$ and $z_{\max}$ are the minimum and maximum in the data set.

2.3.2 Data Filtering via Moving Average

MA filters data by replacing each data point with the average of the neighboring k data points, where k is the size of the memory window. The method is based on the idea that any large irregular component at any point in time will exert a smaller effect if we average the point with its immediate neighbors (Newbold et al., 2003). The equally weighted MA is the most commonly used method, where each value of the data carries the same weight in the data filtering process.
The k-term unweighted moving average y_t^* can be calculated by:

$$y_t^* = \frac{1}{k}\sum_{i=0}^{k-1} y_{t-i} \tag{2.26}$$

where $t = k; \ldots, N$.

2.3.3 Data Transformation

According to Aqil, networks trained on transformed data attain better performance. Then a "'log"' transformation has been considered to bring the observed data as possible to resemble a Normal distribution (Aqil et al., 2007). The "'log"' transformation is performed on each input and output variable independently, using the following equation:

$$Y = a\log_{10}(X+b) \tag{2.27}$$

The forecasted results are then back-transformed using the inverse transformation:

$$X = 10^{Y/a} - b \tag{2.28}$$

where a and b are arbitrary constants.

2.3.4 Correlation Analysis

One of the most important steps in the forecasting model development process is the determination of significant input variables. The employed statistical approach in this study was suggested by Sudheer to identify the appropriate input vector. The method is based on the heuristic that the possible influencing variables, related to different time lags, can be identified through correlation analysis (Sudheer et al., 2002). Basically, Cross Correlation Function (CCF) and Partial Autocorrelation Function (PACF) between the variables are utilized.

2.4 Performance Metrics

In order to study the hydrological models performance, three statistical indicators are considered: the Root Mean Square Error (*RMSE*), Mean Absolute Error (*MAE*) and the Mass Curve Coefficient (E) (Dawson et al., 2007). Their corresponding formulas are given below:

$$RMSE = \sqrt{\frac{\sum_{k=1}^{n}(\hat{y}^k - y^k)^2}{n}} \tag{2.29}$$

$$MAE = \frac{1}{n}\sum_{k=1}^{n}\left|\hat{y}^k - y^k\right| \tag{2.30}$$

$$E = \frac{\sum_{k=1}^{n}(y^k-\bar{y})^2 - \sum_{k=1}^{n}(y^k-\hat{y}^k)^2}{\sum_{k=1}^{n}(y^k-\bar{y})^2} \tag{2.31}$$

2.5 Data Collection and Preliminary Data Analysis

Long-term hydrological and meteorological data are essential for investigating any hydrological regime. However, in developing countries like the case of Lebanon, such data are either scarce, incomplete, missing or not available due to civil wars and other logistic constraints. Unfortunately, even the available data was not accessible via internet portals. Thereby, most of it was gathered personally, through a dull administrative process, from different sites. The modeling approaches presented in this chapter were applied on data collected from existing climate-streamflow stations in Litani river basin. The obtained data, described in Table 2.1, covers rainfall, temperature, dew point, pan evaporation and Litani river flow.

Table 2.1 Basic information on existing stations.

Station Name	Location			Climate		Data type	Data availability
	Latitude	Longitude	Altitude	Summer	Winter		
Joub Jannine	33.38	35.46	871	hot, dry	cold	River flow	2009-2013 (daily)
Qaraoun	33.55	35.69	800	hot, dry	cold	River flow	1990-2013 (daily)
Machghara	33.52	35.64	940	hot, dry	cold	Temperature, Humidity, Dew point, Rainfall	2009-2013 (daily)
Chtoura	33.81	35.88	910	hot, dry	cold	Rainfall	2009-2013 (daily)
Tal Amara	33.86	35.98	922	hot, dry	cold	Pan evaporation	monthly average

Due to urbanization and industrialization, the Litani river basin is today experiencing increasing water demands, groundwater over-exploitation, and extensive pollution. A walk along the riverside shows extensive garbage dumping, direct release of urban sewage water, agricultural run-off, uncontrolled industrial discharges, lack of riverbed maintenance, infringements and prohibited diversions (International Resources Group (IRG), 2012). All these activities are often illegitimate but there are rarely available possibilities for water users to behave differently.

During the on site visit to QD, an interesting piece of information was revealed: the director in charge claims that the Joub Jannine streamgage station is not fully automated which may result in frequent gaps and data inconsistency. However, to fill the gaps, the operators of Joub Jannine station (upstream) acquire river flow data from the Qaraoun reservoir (downstream) 5-6 km away or vice versa. This matter introduces non-homogeneity into data series that was confirmed using Pettitt and Von Neumann homogeneity tests. Thus, besides the uncertainties associated with extreme events (meteorological, hydrological and illegal activities), numerous data limitations affect the accuracy of the results, including insufficient data and inconsistency due to the fact that some measurements were taken from different sources. All these factors suggest a river system with high variability.

Litani river catchment (Figure 2.3) receives annually 500-600 mm of rainfall (Verner et al., 2013). The peak of rainy season is between December and April where 75% of the rain occurs (Figure 2.4). Average temperatures range between 9 ^oC in the winter to 27 ^oC in the summer. Figure 2.5 shows the mean monthly river flow values (Joub Jannine station) with the day-to-day variability at every month. It can be noticed that, high streamflow with high variability occurs in the wet season for the period

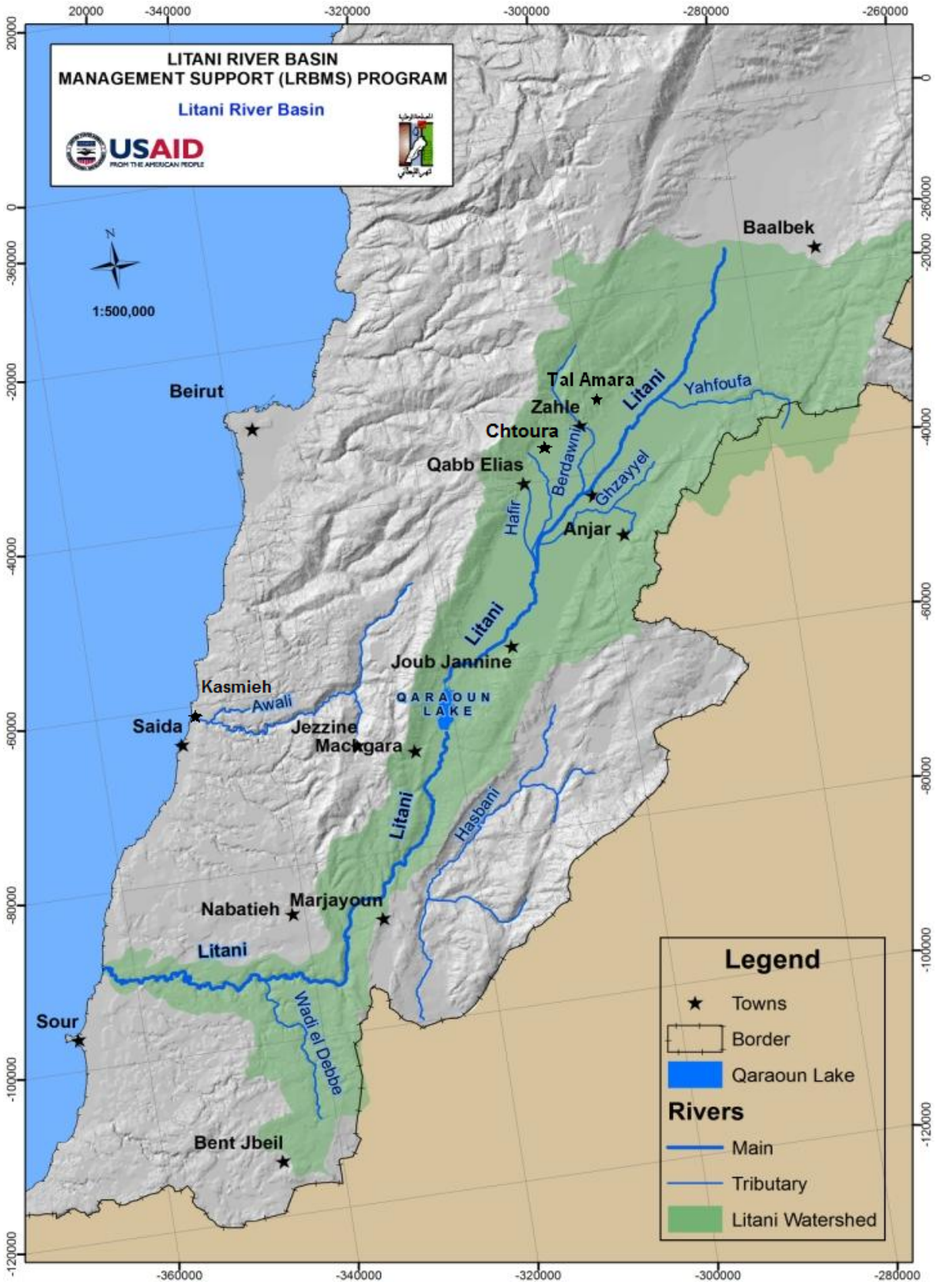


Figure 2.3 Litani Basin - Lebanon.

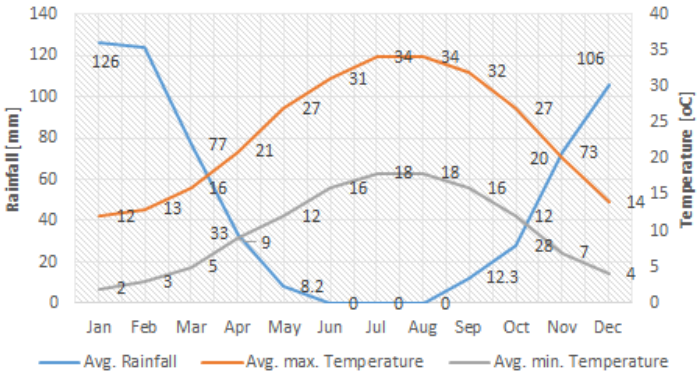


Figure 2.4 Average rainfall and average max./min. temperature (Source: Chtoura weather station).

starting January to April and during December. Furthermore, peak flow occurs during February while the river is almost dry from July till October.

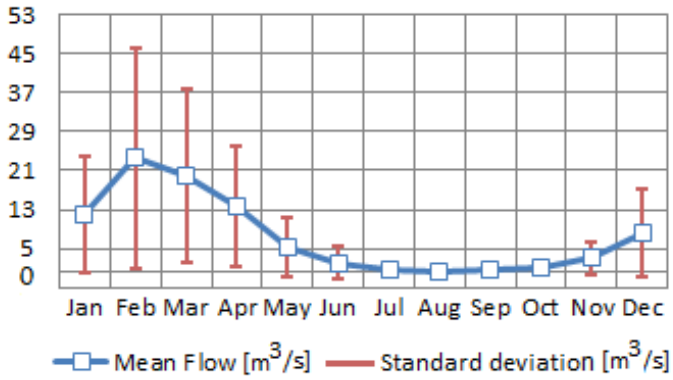


Figure 2.5 Average and standard deviations of the monthly flow - Litani river (Joub Jannine station).

Based on available data, Litani river is characterized by a strong seasonal pattern: high water flow in winter and spring while a low discharge in summer. Furthermore, it possesses great inter-annual variability (Coefficient of Variation (Everitt, 1998) : $CV = 1.336 > 1$ at Joub Jannine and $CV = 1.53 > 1$ at Qaraoun) with a rather weak flow. According to Leopold, this is mainly due to the fact that the river follows a pluvial regime (Leopold et al., 1995).

It was found by Rushworth that under different climate conditions, the influence of precipitation on flow variability arises due to several reasons (Rushworth et al., 2013): 1- antecedent ground wetness, 2- time-delay in rainfall caused by spatial separation, 3- snow accumulation and melt. Therefore, the rainfall is not the only term that induces variation in the streamflow. In fact, by calculating the coefficient of determination R^2 (Stat Trek, 2017) of the available data collected from Machghara and Joub Jannine, only 4.2% of the variation in streamflow is explained by the variation of rainfall.

Figure 2.6 depicts streamflow, rainfall, and temperature of the entire dataset (Machghara and Joub Jannine). The following can be noticed: at the middle of the wet season (around January), the fast responding "runoff" causes a more instantaneous response of streamflow to rainfall. In fact, surface

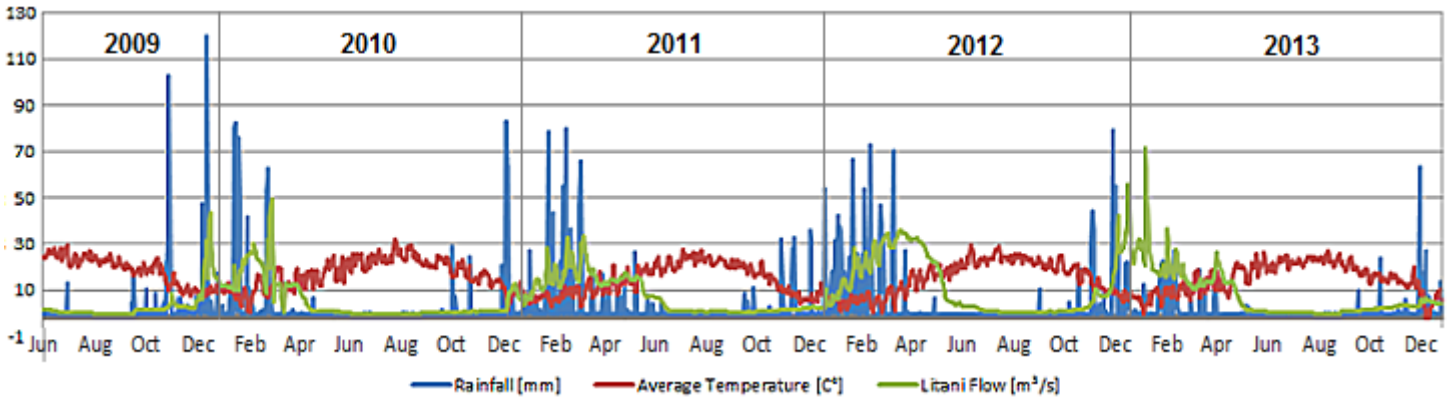


Figure 2.6 Daily river flow (m^3/s), temperature (^0C) and rainfall (mm) starting 01 Jun 2009 till 31 Dec 2013 (data from Joub Jannine - Machghara stations).

runoff accounts for much of the flow during prolonged rainy periods, which wasn't the case at early days of the season. Fast runoff arises when antecedent soil moisture increases to a level where rainfall can move faster near the soil surface without being absorbed. It can result in a rapid increase in flow over a short time period.

During the rainy season, runoff in Litani river catchments is one of the most important drivers of variation in flow levels (Rushworth et al., 2013). It is affected by physical factors including soil and subsurface composition, surrounding land usage, evaporation, and transpiration. However, at the early spring months, as the weather gets warmer and the rainfall starts to tapper off, the snowmelt becomes the main driver of the Litani river. Starting April, the river begins to exhibit a decrease in flow and it continues in this manner until it is almost dry around June month when all the accumulated snow at the mountains tops melts off.

2.6 River Flow Modeling based on Auto-Regressive Method and Constructive Fuzzy Inference

In this section, the study is carried on upstream Litani river (Figure 2.7) using data retrieved from Joub Jannine-Machghara stations. The main concern was dealing with meteorological and hydrological data suffering from insufficiency and also from certain inaccuracies and sometimes unreliability in the information provided by the gauging stations.

Besides that, in the past decades as it was mentioned in Section 2.5, Litani river experienced many major outlaw actions. Thus, the river flow prediction model features a highly dynamic and nonlinear structure. In addition, it accompanies forecasting errors related to noisiness and non-homogeneity of data. However, during the literature review, Fuzzy theory appears to be quite effective for handling these aspects, especially when the inherent physical relationships are not fully understood (Nayak and

Figure 2.7 Litani river flow (Source: wearelebanon.org).

Sudheer, 2008). In addition, according to Cheng, Fuzzy Time Series (FTS) has attracted more interest due to its capabilities of dealing with the uncertainty and the vagueness that are often inherent in real-world data resulting from imprecision in measurements, imperfect sets of observations, or difficulties in acquiring measurements under uncertain circumstances (Cheng and Li, 2012). Bouchon-Meunier also claim that fuzzy logic provides an interesting tool in the field of data mining, mainly because of its ability to represent flaw information, which is crucial when databases are complex, large, imprecise and contain heterogeneous data (Bouchon-Meunier et al., 2008). Therefore, in this study, the proposed Fuzzy inference approach suggested in the literature review (Luna et al., 2007) is adopted for daily river flow time series modeling.

Indeed, the presented method is based on a Constructive Fuzzy System Modeling (C-FSM) and it is formed of two steps: First, the model is initialized by applying the Subtractive Clustering (SC) algorithm on the available historical data to determine the initial structure of the system (Huamani et al., 2011). In fact, this procedure had provided an extra tool to divide the heterogeneous data into more homogeneous sub-populations which in turn improves the forecasting accuracy (Asadia et al., 2013). Second, the initial structure is modified and refined based on constructive offline learning where a classical Expectation Maximization (EM) algorithm is used for adjusting the parameters of the model.

Up-to-date, the major concern in Fuzzy modeling, is the identification of the suitable input vector. Traditionally, the family of Auto-Regressive models has been widely used for modeling water resources time-series. The order of these models is typically estimated by examining the plots of

the Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF) and Cross Correlation Function (CCF). Based on the literature review, the statistical parameters ACF, PACF and CCF could be also utilized in Fuzzy modeling (Sudheer et al., 2002; Galavi and Shui, 2012). Concerning Litani river, the determination of the number of antecedent rainfall, temperature and river flow values involves the computation of time lags that have a significant influence on the forecasting process.

Once relevant inputs are selected, three models are considered: an Auto-Regressive model (AR), Constructive Fuzzy System Modeling (C-FSM) and a Constructive Fuzzy System coupled with a Moving Average filtering method (C-FSM_MA). The MA aims to reduce rainfall fluctuation and filter out noise. The filtered rainfall data are then fed into the C-FSM forecasting model. As a matter of fact, this technique has been used extensively in work of Vos and Wu for predicting runoff and precipitation respectively via ANN modeling (Vos and Rientjes, 2005; Wu et al., 2012). The time series model of AR type is developed in this study as a benchmark model since the correlation analysis used to determine its structure was also adopted for the Fuzzy modeling.

The main scope of this section is not solely comparative. It aims to analyze and discuss stochastic modeling of river flow time series using FIS coupled with traditional correlation analysis in case of data scarcity and heterogeneity. Huamani have presumed the Normality of the gathered data (Huamani et al., 2011), and in the course of the C-FSM training process they didn't state a clear approach for choosing the appropriate cluster radius . In this work, and to boost the performance of the C-FSM, the collected data had been brought to near a Normal distribution using a suitable transformation. In addition, a calibration phase is introduced before validation to select the suitable cluster radius. Moreover, several scenarios were tested by simulating the streamflow associated with different data processing techniques in order to assess their performance.

2.6.1 Data Processing and Input Selection

In the following, we proceed by utilizing the proposed statistical methods with the aim of constructing a framework that allows us to approximate the flow generating processes with an attempt to identify rainfall-flow, temperature-flow and flow-flow (present-previous flow) relationships.

The available hydro-meteorological measurements (Joub Jannine - Machghara stations), corresponding to the continuous period starting from June 2009 to December 2013, were split into two subsets: a training data set composed of all data preceding January 2013 and a testing set formed of the remaining data.

Correlation Analysis

Using the training data, the PACF suggests a significant correlation at 95% confidence level up to 6 days of river flow lag (Figure 2.8). One may notice that lag 6 shows better significance than lags 4 and 5. This anomaly is closely related to the limited data since lag 6 had dropped below threshold once the PACF is carried on the whole dataset (training and testing).

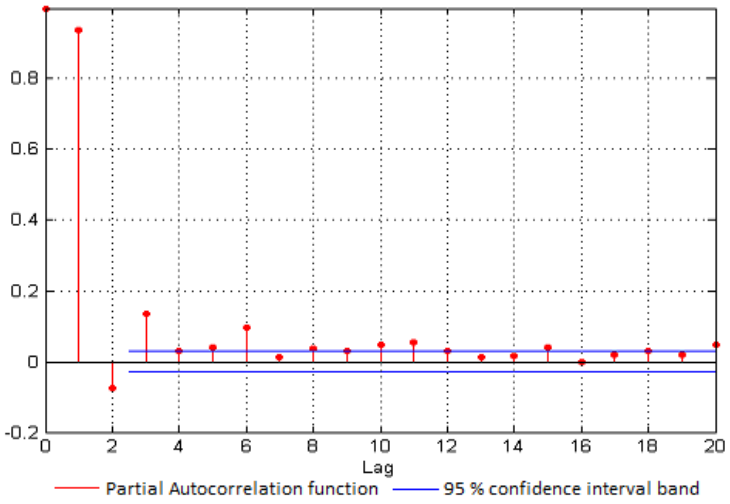


Figure 2.8 Partial Autocorrelation Function of Litani flow.

During the transformation of rainfall into streamflow, the rainfall input to the system goes through two operators: 1- "translation" in time; 2- "attenuation" due to the storage characteristics of the watershed (Chow et al., 1988). The sophistication and complexity of these two operations may explain the weak cross correlation between rainfall and streamflow (Figure 2.9). In order to improve the

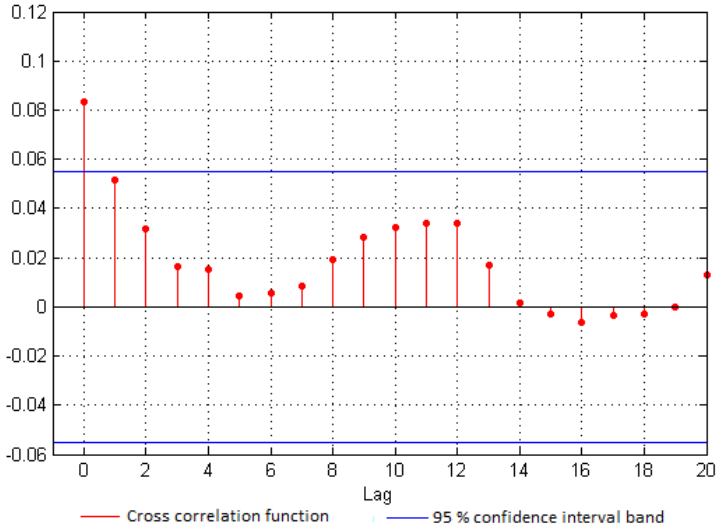


Figure 2.9 Cross Correlation Function between unfiltered rainfall and Litani flow.

similarity between rainfall and streamflow, Wu explored the efficiency of various data pre-processing methods in improving the input-output mapping of the ANN model by filtering raw data (Wu et al., 2012; Wu et al., 2009). One of the used techniques is the Moving Average (MA). The MA operation entails the window size k in Equation 2.26 to filter the raw rainfall data. A suitable k was found by a systematic increase of k from 1 to 12, where at every step, the filtered data are cross-correlated with the river flow data. The targeted value of k corresponds to the optimal zero-lag CCF.

Physically, it is known that, cross correlation measures the similarity between two signals. Thus, k was chosen in a way that reveals the best similarity between rainfall and streamflow.

The plot in Figure 2.10 shows that the best zero lag correlation occurs at a window size 12. Thus, MA(12) is adopted for the filtering process. It is clear that the filtered rainfall data exhibits better

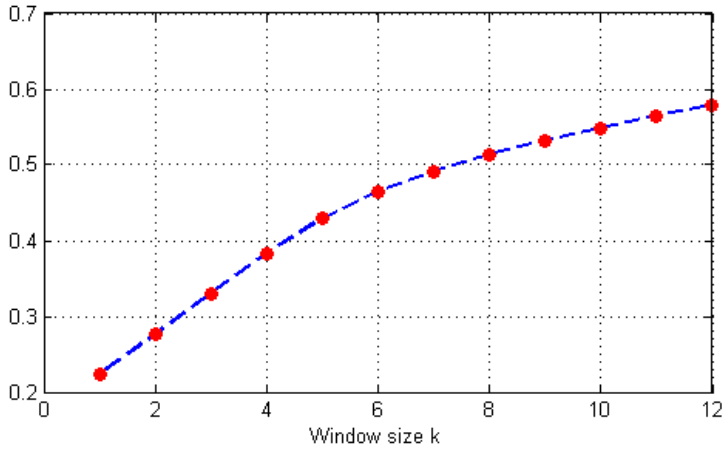


Figure 2.10 Moving Average window size k versus zero lag Cross Correlation.

correlation than the unfiltered one when cross correlated with the river-flow (Figures 2.9, 2.11). We

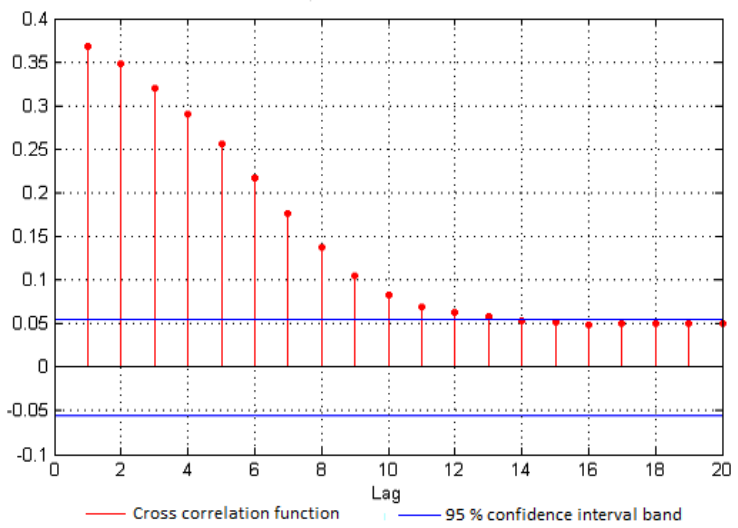


Figure 2.11 Cross Correlation Function between filtered rainfall and Litani flow.

note here that wider window size was not considered since the improvement in cross correlation was negligible. Figure 2.12 exhibits the enhanced similarity between the river flow and the filtered rainfall data.

Furthermore, the temperature (T) and the river-flow (Q) were also cross-correlated and the result showed a negative correlation up to lag 20 which can be interpreted as: T varies in opposite sense with Q. Therefore, (T) is also considered as an input in the suggested models.

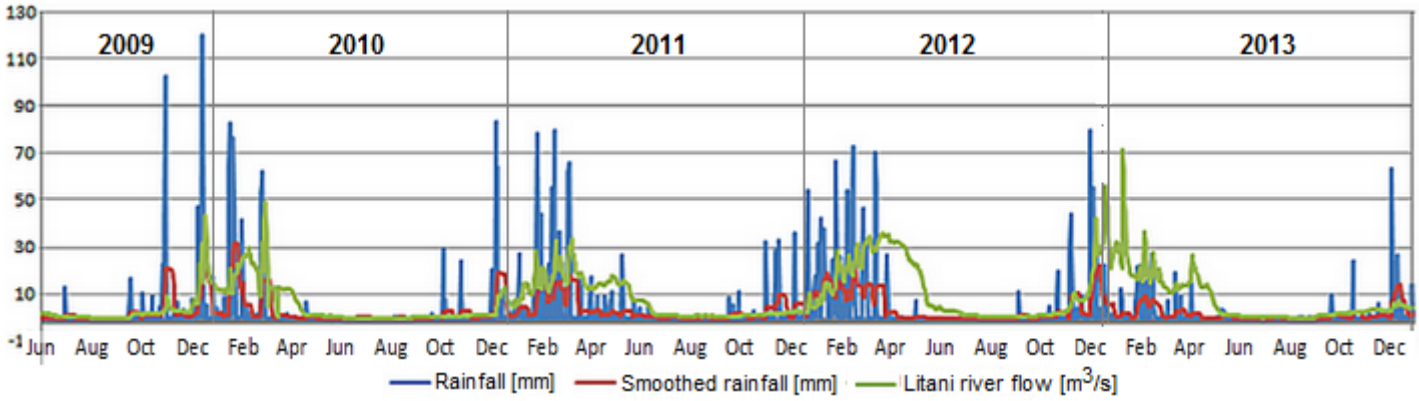


Figure 2.12 Daily river flow (m^3/s), rainfall (mm) and filtered rainfall (mm) starting 01 Jun. 2009 till 31 Dec. 2013.

Data Transformation

The coefficients a and b of Equation 2.27 are obtained on trial-and error basis, until the data follow a Normal distribution. For $a = 0.5$ and $b = 1$, the descriptive statistics of the entire data are shown in Table 2.2.

It can be noticed from Table 2.2 that the standard deviation, skewness and kurtosis show high values in observed data, then found to be reduced significantly after logarithmic transformation. However, regarding the temperature, the skewness and kurtosis were relatively small, thus in this study there is no need to consider data transformation for the temperature.

Input Selection

This part aims at modeling river flow process by AR and FSM models by using recorded rainfall, temperature and streamflow data. Based on the graphical interpretation of PACF and CCF, several input combinations of river flow, rainfall and temperature were examined in the modeling process. The input pattern considers both past and present precipitations $(...,P_{t-2},P_{t-1},P_t)$ and temperatures $(...,T_{t-2},T_{t-1},T_t)$ but only past stream data $(...Q_{t-3},Q_{t-2},Q_{t-1})$ for the river flow. The output corresponds to the present river flow (Q_t), where the subscript t represents the time step. As a consequence, different input combinations of Q, P and T were constructed and listed in Table 2.3.

Another major input that needs to be identified in the Fuzzy modeling is the cluster radius. It is important to recall that the radius specifies the range of influence of the cluster center on each input-output point. Knowing that the cluster radius falls within the unit hypercube, a smaller cluster radius yields an increase in clusters and thus a greater number of rules which will increase the model's complexity. However, Velasquez suggests that the best value for a given radius is usually between 0.2 and 0.5, so the clustering radius is identified through a trial and error procedure by varying the cluster radius from 0.2 to 0.5 with an increment of 0.01 to get the best performance during the calibration phase (Velasquez and Palade, 2013).

Table 2.2 Statistical properties of raw and logarithmic transformed daily data.

Daily Rainfall $[mm]$				
Dataset	Entire (Observed)	Entire (transformed)	Training (transformed)	Testing (transformed)
Statistics	(2009-2013)	(2009-2013)	(2009-2012)	(2013)
Mean	2.993	0.092	0.098	0.070
St. Deviation	10.797	0.213	0.223	0.170
Skewness	5.262	2.485	2.405	2.653
Kurtosis	32.578	5.171	4.681	6.143
CV	3.608	2.320	2.282	2.417
Smoothed Daily Rainfall $[mm]$				
Mean	2.985	0.175	0.189	0.125
St. Deviation	5.268	0.211	0.221	0.157
Skewness	2.253	1.016	0.909	1.222
Kurtosis	4.985	-0.288	-0.581	0.446
CV	1.765	1.203	1.170	1.263
Daily river flow $[m^3/s]$				
Mean	7.463	0.318	0.314	0.329
St. Deviation	9.973	0.253	0.255	0.244
Skewness	1.670	0.402	0.428	0.312
Kurtosis	2.926	-1.283	-1.296	-1.215
CV	1.336	0.796	0.812	0.741

Table 2.3 Models' structure: input-output configuration.

Model	Input structure	Output
AR	$Q_{t-1}, Q_{t-2}, ..., Q_{t-7}$	Q_t
C-FSM	$Q_{t-1}, Q_{t-2}, ..., Q_{t-7}, P_t, P_{t-1}, T_t, T_{t-1}$	Q_t
C-FSM_MA	$Q_{t-1}, Q_{t-2}, ..., Q_{t-7}, P_t, P_{t-1}, ..., P_{t-8}, T_t, T_{t-1}$	Q_t

2.6.2 AR and C-FSM Implementation

Some Tweaks

For the sake of being fair with all Fuzzy models, a calibration phase was considered and performed on one month data (December 2012). Since C-FSM is sensitive to the number of clusters; the best radius r_a, that corresponds to the optimal efficiency E for 1 day lead forecast, is achieved by varying it from 0.2 to 0.5 with an increment of 0.01 and limiting the number of clusters between 2-8 to avoid over fitting (Nayak and Sudheer, 2008).

Table 2.4 displays different scenarios with the utilized data processing method and the optimization algorithm. Further it presents, in case of Fuzzy modeling, the used cluster radius (obtained during calibration) and whether the model is coupled with an Adding Operator (Luna et al., 2007).

Table 2.4 AR, C-FSM and C-FSM_MA Models.

Scenario		Data Pre-processing	Optimization Algorithm			
AR(7)		Std	Yule-Walker			
			Expectation-Maximization			
Group			SC→	r_a	Cluster no	AO
1	C-FSM 1	Std/Norm		0.36	7	No
	C-FSM_MA 1	MA/ Std/ Norm		0.47	4	No
2	C-FSM 2	Std/Norm		0.36	7	Yes
	C-FSM_MA 2	MA/ Std/ Norm		0.47	4	Yes
3	C-FSM 3	Dtrans / Std/ Norm		0.37	7	No
	C-FSM_MA 3	MA/ Dtrans/ Std/ Norm		0.36	8	No
4	C-FSM 4	Dtrans / Std/ Norm		0.37	7	Yes
	C-FSM_MA 4	MA/ Dtrans/ Std/ Norm		0.36	8	Yes

Std: Standardization, Norm: Normalization, Dtrans: Data Transformation, MA: Moving Average, SC: Subtractive Clustering, AO: Adding Operator

Regarding Normality, Huamani asserted that the data have to be Normally distributed before the model coefficients can be estimated (Huamani et al., 2011), while Mehmmet claimed that the Normality assumption is not restrictive and good results can be obtained by using real world observations directly (Mehmmet, 2009). In the current application, this issue is investigated by comparing the models performance on transformed (into the Normal domain) and non-transformed data.

Skewness and kurtosis values lying between -2 and +2 are considered acceptable in order to prove Normal univariate distribution (George and Mallery, 2010). Thus, after the transformation (Table 2.2), the smoothed rainfall and streamflow data satisfy the claim of George concerning Normality (George and Mallery, 2010). On the other hand, the raw rainfall data was pushed as much as possible to resemble a Normal distribution.

2.6.3 Results and Discussion

Table 2.5 shows the performance metrics of 12 prediction horizons for the 9 scenarios carried in the real world (i.e. results were restored to the original space). After fitting the historical flow data to the benchmark AR model of order 7, it shows, for all lead days, the poorest forecasting among the other models. This is due to the fact that the AR model is unlikely able to capture any nonlinear dependency and it is fragile to data non-homogeneity (but gave plausible results). However, the performance of the Fuzzy models accompanied with different pre-processing techniques were more useful in detecting nonlinearities in the streamflow and in dealing with data non-homogeneity.

Furthermore, due to estimation errors of the previous steps that are fed into the input pattern for the next step ahead, one can notice from observing the performance indices of all scenarios a decreasing trend in the Mass Curve Coefficient (E) and an increase in the Root Mean Square Error ($RMSE$) and

Table 2.5 Performance measures of forecasting daily river flow for a horizon h varying from 1 to 12.

Scenario	Performance index	Horizon [days] 1	2	3	4	5	6	7	8	9	10	11	12
AR(7)		3.454	3.700	4.075	4.192	5.003	5.240	4.811	4.633	4.862	5.094	5.034	6.668
C-FSM 1		3.377	4.276	4.623	5.206	4.687	4.950	4.523	5.099	**4.564**	4.912	4.640	4.924
C-FSM_MA 1		3.349	3.534	3.868	3.763	4.987	4.491	4.637	4.373	**4.640**	**4.754**	4.827	5.774
C-FSM 2	*RMSE*	3.346	4.283	4.615	5.184	**4.510**	4.985	4.777	5.040	4.708	5.060	4.822	5.125
C-FSM_MA 2	$[m^3/s]$	**3.301**	3.486	3.816	3.698	4.899	4.625	4.213	**4.181**	4.644	4.913	4.948	4.960
C-FSM 3		3.324	**3.413**	3.717	3.613	4.941	4.183	4.455	4.471	4.995	5.063	5.119	4.647
C-FSM_MA 3		3.344	3.496	3.823	3.722	4.891	4.351	4.448	4.281	4.807	4.885	4.979	5.127
C-FSM 4		3.352	4.178	3.691	3.527	5.020	4.371	4.831	4.585	5.360	5.297	5.354	4.715
C-FSM_MA 4		3.348	3.433	**3.632**	**3.474**	4.885	**3.757**	**4.203**	4.203	4.807	4.798	5.054	**4.350**
AR(7)		0.951	1.203	1.458	1.524	1.824	1.957	1.939	1.961	1.916	2.090	2.046	2.695
C-FSM 1		0.924	1.260	1.438	1.743	1.759	1.786	1.743	2.160	1.873	2.048	**1.745**	**1.801**
C-FSM_MA 1		0.917	1.156	1.369	1.403	1.751	1.760	1.804	1.826	1.821	1.880	1.887	2.412
C-FSM 2	*MAE*	0.928	1.291	1.409	1.762	1.727	1.833	1.923	2.213	1.970	2.240	1.950	2.136
C-FSM_MA 2	$[m^3/s]$	0.907	1.138	1.347	1.379	1.767	1.764	**1.658**	1.722	1.804	2.027	1.954	2.227
C-FSM 3		0.842	**1.031**	**1.241**	**1.285**	**1.715**	1.540	1.700	1.715	1.818	1.965	1.887	1.943
C-FSM_MA 3		0.884	1.110	1.326	1.368	1.734	1.673	1.753	1.746	1.812	1.900	1.867	2.168
C-FSM 4		**0.841**	1.229	1.289	1.301	1.783	1.620	1.834	1.799	2.053	2.209	2.198	1.973
C-FSM_MA 4		0.898	1.122	1.264	1.311	1.788	**1.482**	1.725	**1.705**	**1.800**	**1.831**	1.922	1.870
AR(7)		0.878	0.860	0.830	0.820	0.743	0.719	0.763	0.780	0.758	0.734	0.740	0.544
C-FSM 1		0.888	0.820	0.791	0.734	0.784	0.762	0.799	0.747	**0.797**	0.765	**0.789**	0.764
C-FSM_MA 1		0.885	0.872	0.847	0.855	0.745	0.796	0.780	0.806	0.782	**0.771**	0.762	0.663
C-FSM 2		**0.890**	0.820	0.791	0.736	**0.800**	0.758	0.776	0.753	0.784	0.751	0.772	0.744
C-FSM_MA 2	E	0.888	0.876	0.852	0.860	0.754	0.783	**0.819**	**0.823**	0.781	0.755	0.750	0.751
C-FSM 3		0.887	**0.881**	0.858	0.866	0.750	0.821	0.797	0.795	0.744	0.737	0.731	0.779
C-FSM_MA 3		0.885	0.875	0.850	0.858	0.755	0.806	0.797	0.812	0.763	0.755	0.746	0.731
C-FSM 4		0.885	0.821	0.860	0.873	0.742	0.804	0.761	0.785	0.706	0.712	0.706	0.772
C-FSM_MA 4		0.885	0.879	**0.865**	**0.876**	0.755	**0.855**	**0.819**	0.819	0.763	0.764	0.738	**0.806**

the Mean Absolute Error (*MAE*). Based on the obtained results, the models efficiency (E) in explaining the hydrological process range between 54.4% and 89%.

Effect of data transformation on model performance

For 12 days lead, the AR(7) model gave a *RMSE* of 6.668 m^3/s, the C-FSM 1 model with un-transformed inputs gave *RMSE* of 4.924 m^3/s, i.e. a reduction of 26.15% versus AR. However, the C-FSM 3 model with transformed inputs reduces the *RMSE* by 30.31% with an improvement of 4.16% more than C-FSM 1 model. In a way, this result supports the claim of Mehmmet (Mehmmet, 2009) that using un-transformed data can still provide good results.

In general, results presented in Table 2.5 show that the C-FSM 3 model whose inputs are transformed are more accurate (in terms of the Mass Curve Coefficient E) than C-FSM 1, where it emerges as a better performer for most lead days.

On the other hand, C-FSM 4 did not exhibit a clear better performance than C-FSM 2, neither did C-FSM 2. Apparently, the adding operator didn't work well for both models with transformed and un-transformed inputs.

Impact of MA filter on model performance

Upon using the MA filter, a significant observation was made. The coefficient of determination R^2 of the sub-series joining the filtered rainfall and the river flow indicates a value of 0.2786. That is, 27.86% of the variability in river flow is explained by that of the filtered rainfall. Therefore, the explained variability has increased from 4.18% to 27.86% when using filtered instead of raw rainfall data. This can be interpreted by the fact that, Moving Average contains within a "memory" that has the ability to record, to a certain extent, the variation caused by snow melt and antecedent ground wetness resulted from previous precipitations. Thus, MA didn't just remove the noise but it has improved the explained variance by more than 27%.

The impact of the MA filter on the performance of the C-FSM model is described as follows: each of the three models C-FSM_MA 1, C-FSM_MA 2 and C-FSM_MA 4 exhibits a noticeable prediction efficiency in many lead days in terms of *RMSE*, *MAE* and *E* compared with C-FSM 1, C-FSM 2 and C-FSM 4. But the remarkable performance is achieved by C-FSM_MA 4 that shows almost the lowest *RMSE*, the lowest *MAE* and the highest *E*.

Figure 2.13 shows the time series plot for 1, 4, 9 and 12 days ahead forecast associated with C-FSM_MA 4 model. During the beginning and the end of the wet season (October, November, December and March, April, June), the flow variability is low and the model shows a noticeable fit with the actual flow. However, it wasn't the case during January and February months. Furthermore, the actual flow is characterized with a very sharp spike (12 Jan. 2013). This can be interpreted as an anomaly in the observation due to inaccurate measurements and can't be considered as a flood for two main reasons:

- The accumulated rainfall, a week before the spike occurrence date, was only 21 mm, and
- The average temperature during this week was below 6.5 degrees Celsius.

These two reasons are not enough to produce a sudden elevation in the river flow from 20 m^3/s to 71 m^3/s based on previous observations (Figure 2.6).

Besides inaccurate measurements, the river flow forecast is disrupted by vast sources of noise due to: illegal activities previously mentioned (International Resources Group (IRG), 2012), some meteorological conditions (wind, evaporation, irradiance,...etc.) and urbanization. All these factors distort the accuracy of the river flow model and cause a decrease in the forecasting precision. To reduce the noise effect, one can consider different noise filters. However, for the mentioned sources of distortion, quantitative data are not available. Thus, the attention is turned to the noise existing within the rainfall data and MA filter. In this case, the Noise to Signal ratio (Jayawardena and Gurung, 2000) was calculated for the two time series (12 days ahead) C-FSM 4 and the denoised one C-FSM_MA 4.

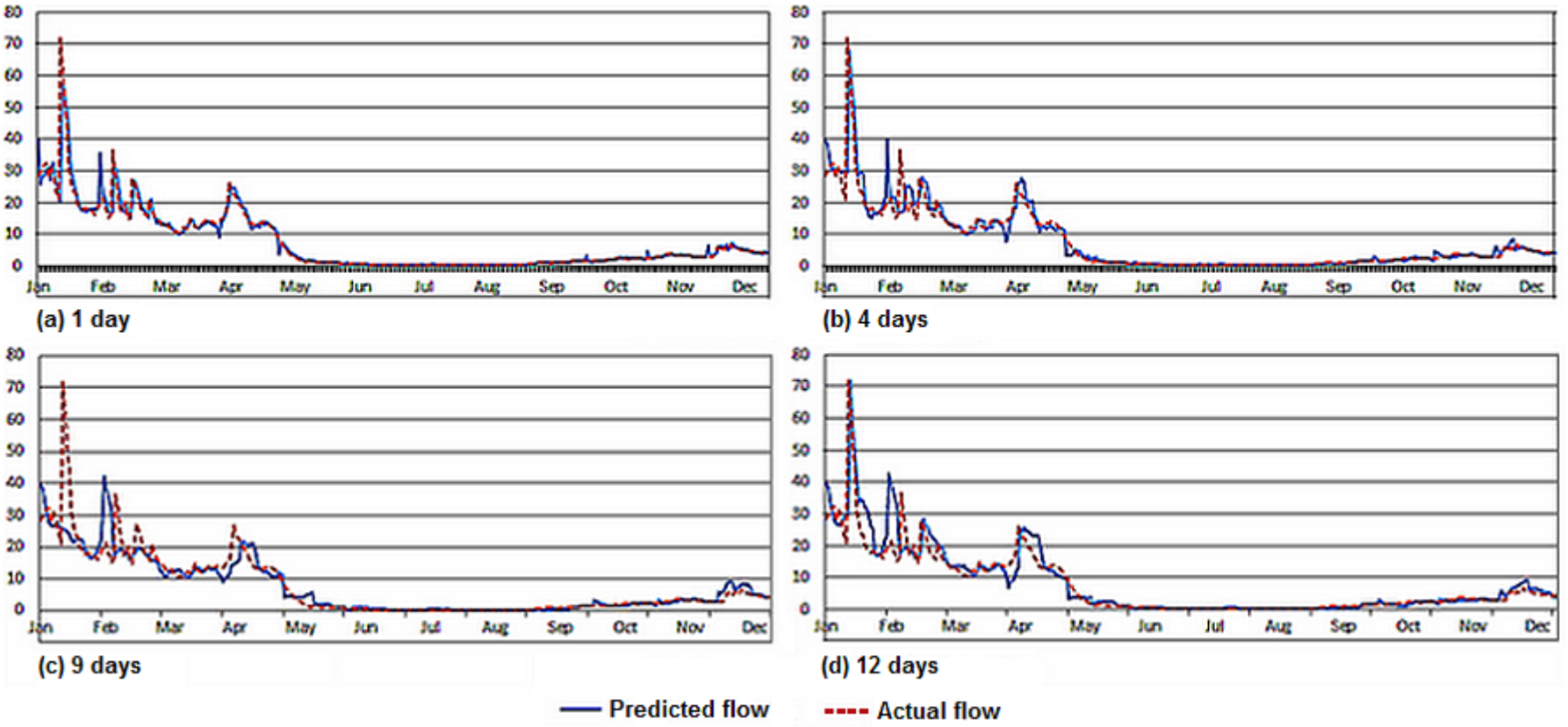


Figure 2.13 C-FSM_MA 4 river flow estimates (a) 1 day (b) 4 days (c) 9 days (d) 12 days ahead along with the observed flow for Litani river over the testing period (1 January - 31 December, 2013).

The obtained respective values were 0.48 and 0.44. Hence, the MA filter applied to rainfall has reduced the noise in the streamflow time series by 8.33% . This percentage is an acceptable value bearing in mind the scarcity of noise sources data. Furthermore, some people may claim that filtering may remove noise as well as variability. Therefore, it might not be a good choice for daily streamflow forecasts. In fact, this issue is explained and discussed in the next paragraph.

Although, the performance of C-FSM_MA 4 model was more than satisfactory with an efficiency reaching 80.6% for the 12 days ahead forecast. The model was also able to explain 84.27% of the actual river flow variability. The C-FSM_MA 4 model managed to reproduce the day-to-day variability almost within naturally occurring ranges by taking the “memory" advantage of MA filter while the C-FSM 4 model that was fed with unfiltered rainfall could capture 79.52 % only.

Furthermore, the nonlinearity of streamflow processes is also investigated. Brock introduced a test for the existence of nonlinearity in streamflow processes (Brock et al., 1996). It is found that the shorter the time-scale, the stronger the nonlinearity. All annual time series models are linear, whereas all daily streamflow processes posses strongly nonlinear characteristics (Wang, 2006). Looking backward to the linear benchmark model, the AR time series forecast was correlated with the original river flow. It revealed that, for 12 days ahead forecast, the coefficient of correlation is 0.86, which means that the nonlinear identification was difficult and the AR, as expected, manifests not much of accurate results.

However, regarding C-FSM_MA 4 model, the coefficient of correlation between the observed and the forecasted flows for 12 days lead is equal to 0.92. Thus, this model is more competent in capturing the nonlinearity in river flows at different lead days.

Figure 2.14 shows the scatter plot of both the observed and the predicted flows obtained by using the C-FSM_MA 4 model on the testing period for 12 days lead. The line $y = x$ represents the perfect fit case when the predicted and the observed river flows are equal. In fact, the reader can notice that, along the line $y = x$, a tight dispersion for the low flows and a wide one for the high flows (within the circle). Thus, based on the data distribution for high and low flows, the forecasting model showed a good prediction accuracy for the low values of the flow but it was unable to maintain the same accuracy for the high values.

In general, models performance in reproducing and inferring river flow for the testing year were more than satisfactory, given the limitations descending especially from the quality and quantity of the historical observations. If somehow, an advanced data acquisition system was installed on the river’s site, it will have the ability to obtain more accurate and reliable meteorological and hydrological measurements. Thus, by sweeping off uncertainties related to missing or inaccurate observations, the models would have delivered even better results. Furthermore, reliable and longer climate and discharge measurements would have allowed a proper training and testing of the model performance. The data scarcity did not allow to account for other sources of uncertainty, such as factors related to climate change and urbanization. However, the C-FSM models proved to be accurate enough to provide plausible results and a reasonable agreement with the observed streamflow. Thus, they were robust enough to be used in a situation where data possess a certain level of heterogeneity.

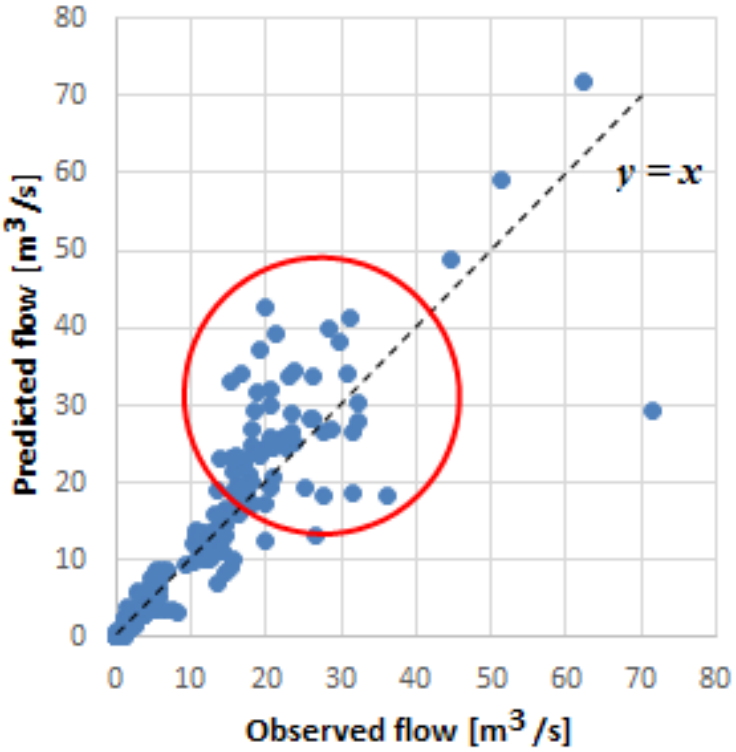


Figure 2.14 Observed versus predicted river flow for the testing period: 01 Jan. till 31 Dec. 2013.

Unfortunately, besides the data scarcity and heterogeneity, the reader may notice in Table 2.1 that the river flow measurements at Qaraoun and rainfall measurements at Chtoura are of different sizes. Consequently, to apply the C-FSM approach, the data should be re-sampled either by: 1- using only river flow data; or 2- using part of the river flow data such that both streamflow and rainfall data have matching time periods (balanced). In both situations, not the whole data are fully exploited. The main disadvantage is that we may lose potentially relevant information from the left-out data. However, the next section provides a resolution of this issue based on Two-Phase Constructive Fuzzy System Inference.

2.7 River Flow Modeling based on Two-Phase Constructive Fuzzy System Inference

In this section, a modeling approach is suggested for river flow forecasting that have ability to deal with hydro-meteorological measurements asymmetry. In fact, the term hydro-meteorological measurements asymmetry is a critical issue that occurs in poorly-gauged basins. It is interpreted in a sense that the climatic measurements exist for a shorter period in comparison with the hydrological ones. The challenge, in this work, is to enhance river flow prediction by developing a daily river flow forecasting model that can exploit the whole available data rather than using only the hydrological measurements

or the balanced hydro-meteorological data set (re-sampling data). Before going through the modeling process, managing such hydrological system in developing countries requires further knowledge of the hydrological developments and their application under data-scarce conditions.

Although a variety of forecasting approaches based on Auto-Regressive methods and Artificial Intelligence (AI) have been successfully formulated, further studies have applied hybrid systems to the problem of river flow prediction in order to generate more accurate models (Jain and Kumar, 2007; Wang et al., 2006). However, a hydrological model application under data scarcity may be sometimes unreliable due to the inability to specify the model components or to estimate efficiently the parameter values that consistently represent the dominant hydrological processes in a particular basin. Neale claims that there is a number of important research directions still being pursued (Neale et al., 2009), such as priori methods that help to infer model's parameter values either directly by observing the basin characteristics (example geology, topography, soils, land cover,...etc.), or indirectly from regionalized parameter-to-watershed characteristics obtained from "hydrologically similar" gauged basins. Other techniques are based on remote sensing technologies (satellites, radars) and coupled with systems of hydro-meteorological prediction models as in the literature review (Chang and Tsai, 2016; Gao et al., 2017). Satellites and radars offer opportunities to acquire temporal and spatial details about the water cycle components (such as rainfall, soil moisture, evaporation,...etc.) for the use in the hydrological model development and validation. These methodologies help in reducing the negative consequences of data scarcity and thereby improve water resources management.

Unfortunately, due to technical and financial barriers, developing countries suffer from limited knowledge in hydrology (Hughes et al., 2014) and from the lack of concrete geological and topographical studies. In addition, remote sensing data are often hard to obtain, mainly due to inaccessibility or high cost. Consequently and for the time being, the proposed methodologies that are based on hydrologically similar watersheds or remote sensing technologies can't be applied efficiently for hydrological modeling in these regions.

Faced with such embarrassing conditions, the current efforts were turned to a cheaper approach based on soft computing and implemented on the available data. In fact, considering the previous study that was presented in Section 2.6, Fuzzy inference appears to be effective in handling aspects such as shortage and heterogeneity (Zadeh, 1965) of the hydro-meteorological data. As well, Fuzzy inference has attracted more interest due to its capabilities of dealing with vagueness that are often inherent in real-world data resulting from imprecision in measurements, imperfect sets of observations, or difficulties in acquiring measurements under uncertain circumstances (Cheng and Li, 2012). Furthermore, from the knowledge gained from recent hybrid modeling (Han et al., 2017; Wang et al., 2017; Banihabiba et al., 2017; Barak and Sadegh, 2016) in different forecasting fields (wind speed, $PM_{2.5}$ concentration, daily inflow and energy consumption), hybrid model seems to outperform individual models being used alone. These hybrid models are namely, ARMA-NP (NP: Non-parametric), ARIMA-SVM (SVM: Support Vector Machine), ARIMA-NARX (NARX: Nonlinear Auto-Regressive model with eXogenous inputs) and ARIMA-ANFIS (ANFIS: Adaptive Neuro Fuzzy Inference System). However, during the

literature review, all hybrid models were trained over symmetrical measurement data sets while the asymmetric case was not considered. Thus, relying on the former work in this research, a Two-Phase Constructive Fuzzy System Modeling (TPC-FSM) approach is introduced to handle measurements asymmetry. It consists of training the linear and nonlinear components of the hybrid model using C-FSM.

In fact, the main contribution of this section can be summarized in the following:

1. It highlights, in poorly gauged basins, a new problematic data issue that may arise during river flow forecasting where hydrological data exist for a longer period than meteorological data (data measurements asymmetry). In fact, one may claim that data re-sampling could resolve measurements asymmetry problem, however the main drawback is that potentially useful information are lost. On the other hand, taking advantage of the whole data set can enhance the forecasting accuracy.

2. A TPC-FSM approach is proposed that couples hybrid modeling with C-FSM method in order to handle hydro-meteorological data asymmetry. In fact, the river flow time series is decomposed into a linear and nonlinear components. Both components are fitted by the C-FSM approach. First, the linear component is estimated, and then the second nonlinear term is obtained on the residuals. The aim of the TPC-FSM approach is to exploit the foremost of the available data to overcome the drawbacks of the conventional C-FSM under data-asymmetry condition. The TPC-FSM goal is to achieve a higher prediction accuracy versus the C-FSM for daily river flow time series.

3. In order to determine the effectiveness of the suggested approach, a comparative study between TPC-FSM and C-FSM is carried using real data retrieved from Litani basin.

2.7.1 Data Preprocessing and Input Selection

In this section, the available Litani flow measurements correspond to a continuous period starting from January 1990 to December 2013 obtained from QD gauging station. However, the available rainfall data cover only the period starting from July 2009 till December 2013 which were retrieved from Chtoura weather station. All the gathered data are split into training and validation sets according to Table 2.6.

Correlation Analysis

Using the training data set Γ_1, the PACF suggests a significant correlation at 99% confidence level up to 4 days of river flow lag (Figure 2.15). Whereas, utilizing the data of Γ_2 and Γ_3, the CCF reveals a significant relationship at 99% confidence level up to 2 days of rainfall lag (Figure 2.16).

Table 2.6 Hydrological and Meteorological data distribution.

Group	Type	Date	Set	Notation
1	Streamflow	Jan. 1990 - Dec. 2012	Training	Γ_1
2	Streamflow	Jan. 2013 - Dec. 2013	Validation	Λ_1
3	Rainfall	Jul. 2009 - Dec. 2012	Training	Γ_2
4	Rainfall	Jan. 2013 - Dec. 2013	Validation	Λ_2
5	Streamflow	Jul. 2009 - Dec. 2012	Training	Γ_3

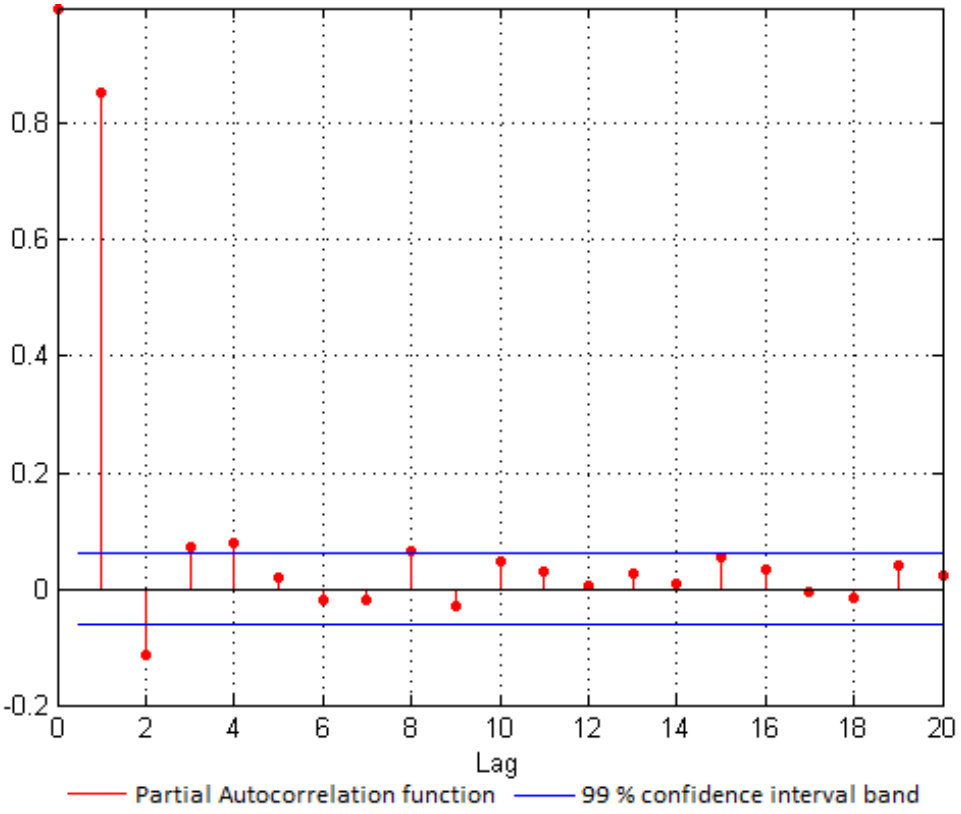


Figure 2.15 Partial Autocorrelation Function of Litani flow.

Input Selection

The river flow here is modeled using only recorded rainfall and streamflow data. Based on the graphical representation of the PACF and CCF given in Figures 2.15 and 2.16, two different input combinations were constructed for the models C-FSM1 and C-FSM2 and are listed in Table 2.7. These two models are trained later using C-FSM method.

Another essential input that needs to be identified in the Fuzzy modeling is the cluster radius. However, for the sake of being fair with all Fuzzy models and since FSM is sensitive to the number of clusters: the best radius r_a, that corresponds to the optimal performance efficiency E for 1 day lead forecast, is achieved during the validation procedure.

Table 2.8 illustrates the data used for the learning process and during validation by each model.

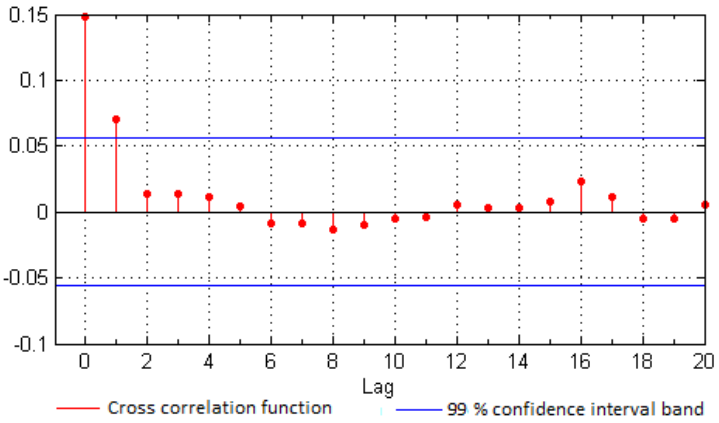


Figure 2.16 Cross Correlation Function between rainfall and Litani flow.

Table 2.7 Models' structure: input-output configuration.

Scenarios	Input structure	Output
C-FSM1	$Q_{t-1}, Q_{t-2}, \ldots, Q_{t-4}$	Q_t
C-FSM2	$Q_{t-1}, Q_{t-2}, \ldots, Q_{t-4}, P_t, P_{t-1}$	Q_t

Table 2.8 Training and validation sets assigned to different models.

Scenarios	Training sets	Validation sets
C-FSM1	Γ_1	Λ_1
C-FSM2	Γ_2 and Γ_3	Λ_1 and Λ_2
TPC-FSM	Γ_1, Γ_2 and Γ_4	Λ_1 and Λ_2

Γ_4 is a set whose elements are the residuals determined in phase one using Equation 2.15

2.7.2 TPC-FSM Implementation

First, the flow Q_t is decomposed into a linear part $L_t = L_t(Q_{t-1}, Q_{t-2}, Q_{t-3}, \ldots)$ and a nonlinear part N_t tightly related to rainfall, where the subscript t represents the time step (in days). Afterward, the reader proceeds according to the following steps:

Step 1. The input pattern for L_t is determined using the PACF on the training data set Γ_1 (Table 2.6)

Step 2. Parameters of $\widehat{L}_t$ are obtained by training over Γ_1 using C-FSM. However, to guarantee its linearity, the number of clusters is set to one using an appropriate cluster radius.

Step 3. $\widehat{L}_t$ is estimated for one day ahead over the period defined in Γ_3.

Step 4. The residual d_t is computed by subtracting the predicted $\widehat{L}_t$ from the actual data Q_t as in the equation:

$$d_t = Q_t - \widehat{L}_t \tag{2.32}$$

The obtained data set for d_t covers the period between July 2009 and December 2012 and it was denoted by Γ_4.

Step 5. The residual d_t can be written in the form:

$$d_t = f(d_{t-1}, d_{t-2}, ..., P_t, P_{t-1}, ...) + e_t \tag{2.33}$$

such that e_t is a random error and f is a nonlinear function given by:

$$f(d_{t-1}, d_{t-2}, ..., P_t, P_{t-1}, ...) = \sum_{i=1}^{M} g_i^t \cdot f_i^t$$

where $f_i^t = \theta_0 + \theta_i^{(1)} \cdot d_{t-1} + \theta_i^{(2)} \cdot d_{t-2} + ... + \gamma_i^{(0)} \cdot P_t + \gamma_i^{(1)} \cdot P_{t-1} + ...$ is defined over the i^{th} region, θ_0, $\theta_i^{(1)}$, $\theta_i^{(2)}$, ..., $\gamma_i^{(0)}$, $\gamma_i^{(1)}$, ... are parameters and g_i^t is the membership function.

Step 6. The input pattern for d_t is specified using the PACF and the CCF on Γ_4 and on Γ_4 with Γ_2 respectively. The goal is determining the residual lags along with the rainfall lags.

Step 7. The C-FSM algorithm is used to tune the parameters of $\widehat{N}_t$ the estimate of d_t using data of the sets Γ_4 and Γ_2 (Table 2.6).

Step 8. $\widehat{N}_t$ is determined using data in the set Λ_2 for different lead days. By the way, $\widehat{L}_t$ is also computed for same lead days over the validation period January 2013 - December 2013.

Step 9. The outcome obtained from step 8 will yield to the forecasting of the nonlinear $(\widehat{N}_t)$ and linear $(\widehat{L}_t)$ component of the time series. Then the forecasted values of the river flow time series are obtained as follows:

$$\widehat{Q_t} = \widehat{L}_t + \widehat{N}_t \tag{2.34}$$

The PACF of the time series was employed to examine the effect of preceding flows on L_t. The structure of L_t appeared to have the form $L_t = L_t(Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4})$. Then, the model is trained using C-FSM algorithm over Γ_1. During the training process, the cluster radius is chosen such that the number of clusters is equal to one in order to guarantee the linearity of $\widehat{L}_t$. L_t is forecasted for one day ahead using $\widehat{L}_t$ over the period July 2009 - December 2012. Then d_t is computed over the same period. The PACF, CCF and the corresponding 99% confidence bands from lag 0 to lag 20 were respectively estimated for both d_t and (d_t, P_t) data as presented in Figures 2.17 and 2.18.

Based on the displayed correlation analysis, the triplet (d_{t-1}, P_t, P_{t-1}) was considered as an input for d_t. For the proposed combination, the time series model $d_t = f(d_{t-1}, P_t, P_{t-1}) + e_t$ is trained using also the C-FSM algorithm.

2.7.3 Results and Discussion

The coefficient of determination $R^2 = 0.124$ indicates a weak relationship between rainfall and the river flow. Furthermore, 12.4% of the variability in river flow is explained by that of the rainfall. However,

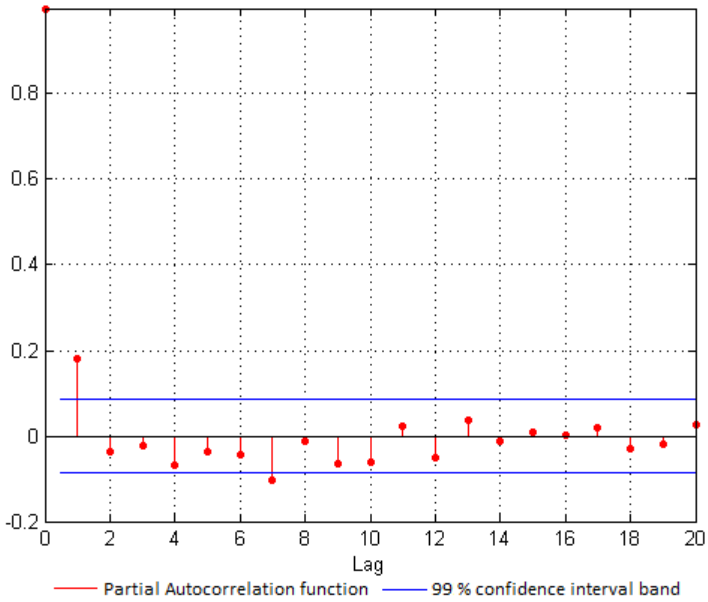


Figure 2.17 Partial Autocorrelation Function for the residual d_t.

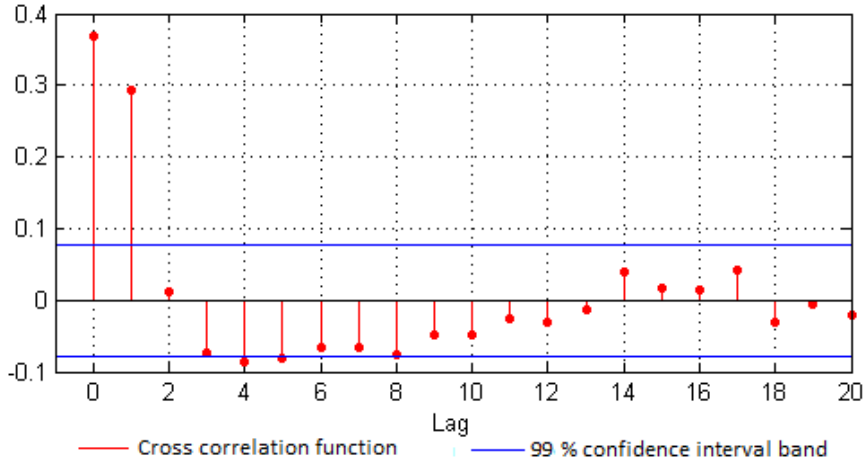
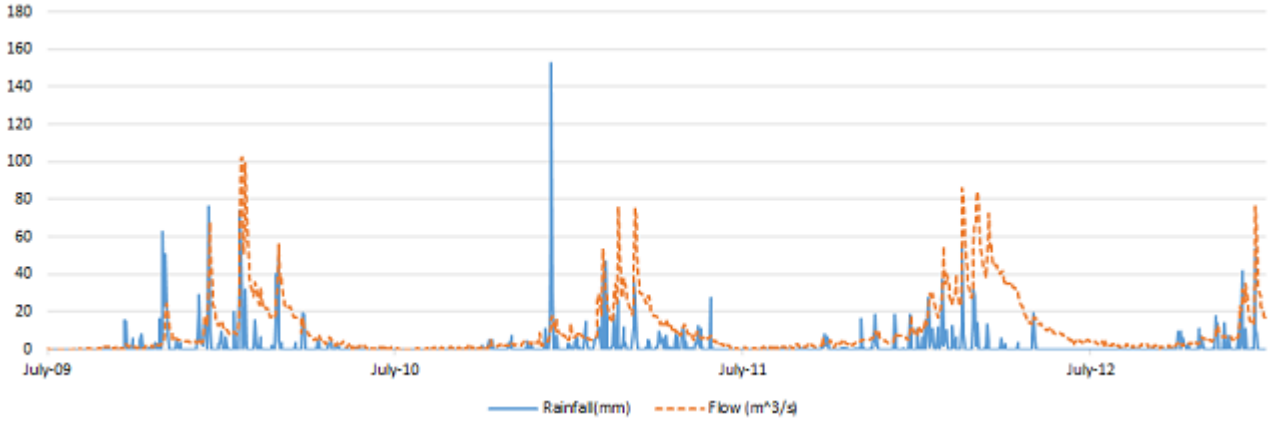


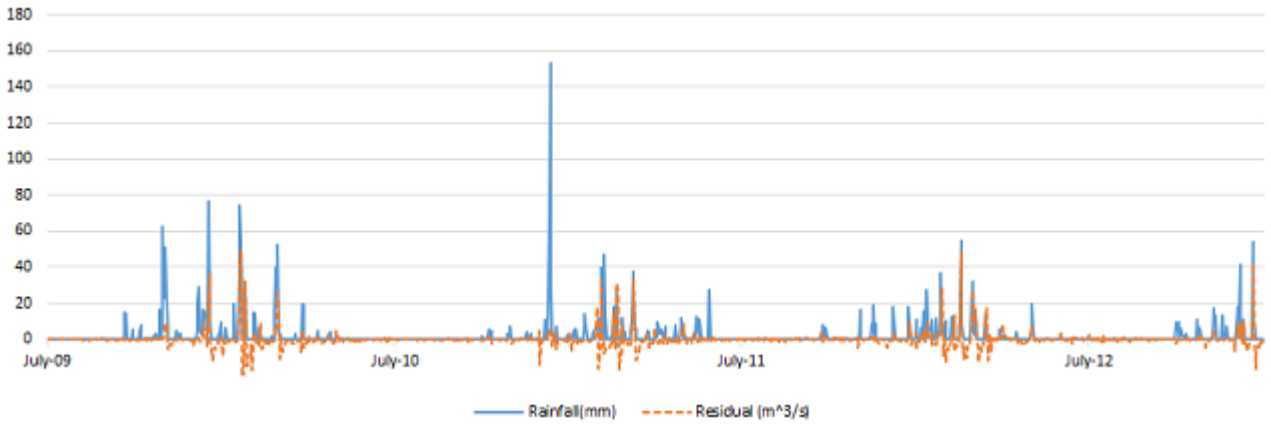
Figure 2.18 Cross Correlation Function for the residual d_t and rainfall P_t.

the rainfall-residual relationship has shown a better R^2 of value 0.521. Thus, the explained variability has increased from 12.4% to 52.1% with an improvement of almost 40%. This can be interpreted by the fact that, once the component that depends on the past flows is removed from the river flow, the remaining component is affected significantly by precipitation. As a result, d_t exhibits a better correlation with rainfall P_t (Figure 2.19).

However, Table 2.9 shows the performance metrics for 1, 3 and 6 prediction horizons for the 3 different Fuzzy models. After fitting the assigned data series to the models C-FSM1, C-FSM2 and TPC-FSM, the first model shows, for all lead days, the poorest forecasting among the other models. Indeed, It is due to the fact that the river flow time series is affected by different factors other than past flows. Whereas, the performance of the C-FSM2 and TPC-FSM models accompanied with rainfall as



(a) Rainfall versus river flow.



(b) Rainfall versus residual.

Figure 2.19 Daily river flow (m^3/s), residual (m^3/s) and rainfall (mm) starting 01 Jun. 2009 till 31 Dec. 2013.

an input variable were more accurate and efficient in the streamflow prediction. As a primary result, the models efficiency (E) in explaining the hydrological process ranges between 79.1% and 93.4%.

Before going further in the results discussion, let us check the Boxplot (Figure 2.20) generated by SPSS. This diagram screens the daily streamflow historical data (January 1990 - December 2013), and identifies different outliers. In Figure 2.20, circles represent mild outliers whereas asterisks are extreme outliers; as a whole, 2.3% of the historical data series are identified as exceptional measurements. These outliers may be unexplainable or unrealistic values outside the assumed population, or could be realistic but very infrequent measurements. By this definition, outliers either result from faults (e.g., measurement or transcription error), or due to natural causes like a flood. Thus, it is important to study performance of the models in the presence and absence of these exceptional values. Recalling, TPC-FSM was trained using the data of Γ_1 while C-FSM2 was fitted utilizing data in Γ_3. Bearing in mind that, Γ_1 and Γ_3 contain respectively 95.5% and 20.3% of the existing outliers, the expected results can be summarized in the following: it is more likely that the model TPC-FSM will perform better

Table 2.9 Performance measures of forecasting daily river flow for a horizon h 1, 3 and 6.

Model	Performance index	Horizon [days] 1	3	6
C-FSM1		0.94	0.89	0.9
C-FSM2	R	0.95	0.91	0.92
TPC-FSM		**0.96**	**0.92**	**0.93**
C-FSM1		5.6	7.97	7.81
C-FSM2	$RMSE$ $[m^3/s]$	5.23	7.06	6.81
TPC-FSM		**4.47**	**6.92**	**6.45**
C-FSM1		0.897	0.791	0.799
C-FSM2	E	0.91	0.836	0.847
TPC-FSM		**0.934**	**0.842**	**0.863**

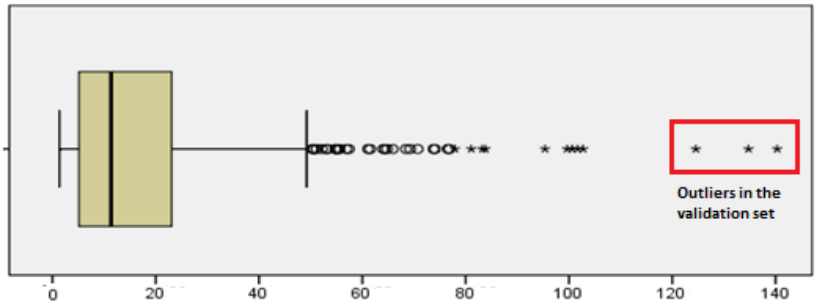


Figure 2.20 Boxplot screening the entire river flow time series (Scale: m^3/sec).

than C-FSM2 on a testing set that comprises extreme outliers, which will not be the case otherwise. In fact, the existence of significant number of outliers in the training set may sometimes disrupt the accuracy of the forecasting model in normal situation (outlier free validation period).

In the carried study, most extreme outliers (Figure 2.20) lied within the validation set Λ_2, which imposes a challenge during the validation process. Nevertheless, the two models C-FSM2 and TPC-FSM exhibit a noticeable prediction efficiency in all lead days in terms of *RMSE* and E compared with C-FSM1. But the remarkable performance is achieved by TPC-FSM. It shows the lowest *RMSE* and the highest E. Figure 2.21 describes the representative details of the hydro-graphs for 1, 3 and 6 days ahead associated with the TPC-FSM model. One can notice for different lead days, how the model dealt with the sharp spike associated with the river flow of 140 m^3/s that occurred around January 9 2013. This peak flow is a result of heavy rain for three consecutive days (total accumulation 107 mm).

Concerning variability, the presence of extreme values will induce a higher variance in the stream-flow time series. For the longest lead period, the models C-FSM1, C-FSM2 and TPC-FSM were able to explain respectively 81%, 84.6% and 86.5% of the actual river flow variability. Results indicate that the developed model reproduces variations in the observed flows very well. The lower R^2 produced by the C-FSM1 model is primarily due to the existence of other factors that promote variability rather

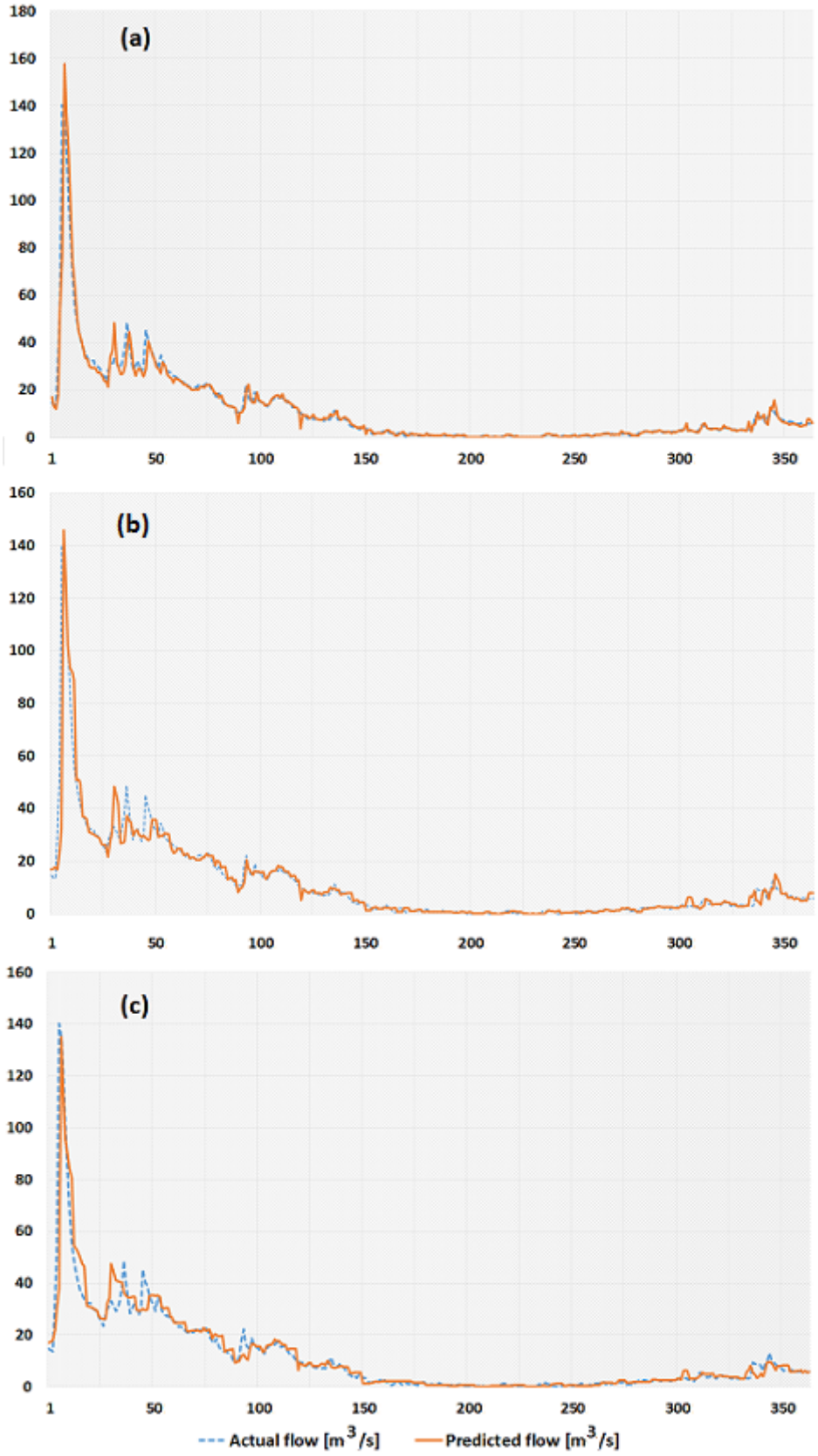


Figure 2.21 TPC-FSM river flow estimates for (a) 1 day (b) 3 days (c) 6 days ahead along with the observed flow for Litani river over the testing period (1 January - 31 December, 2013).

than river flow itself (example rainfall, temperature, soil wetness,...etc.). Indeed, TPC-FSM model managed to capture extreme values as well as variations of the original data set in a better way and consequently can lead to superior performance.

Removing outliers is still a controversial issue for scientist since some outliers are of primary interest in modeling rare events (Nisbet et al., 2009). However, the performance of the models was studied after eliminating the effect induced by the extreme outliers in the validation set Λ_2 (enclosed within the red box in Figure 2.20). Results in Table 2.10 indicate that the models C-FSM2 and TPC-FSM were capable of producing more appreciable outcomes in all lead-days forecasts. As expected, the C-FSM2 model showed better accuracy when the effect of the extreme values was canceled.

Table 2.10 Performance measures of forecasting daily river flow when outliers are excluded.

		Horizon [days]		
Model	Performance index	1	3	6
C-FSM2	$RMSE$ $[m^3/s]$	2.39	3.37	3.22
TPC-FSM		2.6	3.84	3.57
C-FSM2	E	0.968	0.937	0.94
TPC-FSM		0.963	0.92	0.93

With the purpose of verifying the advantage of the TPC-FSM scenario, Nisbet claimed that outliers inject noise in the data set; this noise will reduce the predictability of the forecasting model (Nisbet et al., 2009). On the other hand, leaving outliers contained within the data set, the model will be capable of capturing extreme events. In this case, TPC-FSM outperforms the C-FSM2 paradigm. However, since the training set for C-FSM2 encompassed fewer outliers than that of TPC-FSM, C-FSM2 shows better accuracy in normal situations.

Concerning the best line fit between observed and simulated flow values, the coefficients of correlations R of C-FSM1, C-FSM2 and TPC-FSM are 0.9, 0.92 and 0.93 respectively for 6 days ahead forecast. As a result, the TPC-FSM model is more competent in capturing the linearity and nonlinearity in river flow signal among the other models.

Furthermore, Figure 2.22 shows the scatter plot of both the observed flow and the predicted flow obtained by using the TPC-FSM model on the testing period for 6 days lead. The line $y = x$ represents the perfect fit case when the predicted and the observed river flows are equal. In fact, the reader can notice that, along the line $y = x$, a tight dispersion for flows ranging between 0-40 m^3/s and a wider one for the higher flows. Thus, based on the data distribution for high, moderate and low flows, the forecasting model showed a very good prediction accuracy for the low to moderate values of the flow, where most of the observations concentrate, but it shows less accuracy for higher values.

As a final attempt to examine the effectiveness of the TPC-FSM scenario, the number of parameters is checked for each paradigm. Both C-FSM2 and TPC-FSM encompass 33 parameters while C-FSM1 comprise only 18 parameters. Since C-FSM2 and TPC-FSM consider rainfall as input other than the

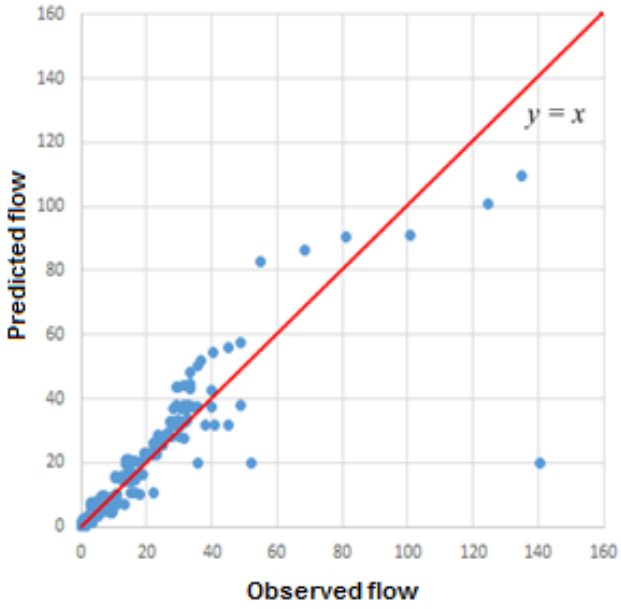


Figure 2.22 Observed versus predicted river flow $[m^3/s]$ for the testing period: 01 Jan. till 31 Dec. 2013.

past river flows, both models outperform C-FSM1. On the other hand, Alessio claims that the larger number of the estimated parameters, the larger the estimation errors and in this case the over fitted model has poor predictive performance Alessio, 2016). However, based on the achieved results, TPC-FSM exhibits the best performance related to river flow variability, similarity with rainfall time series, efficiency and the capture of linearities and nonlinearities in the river flow signal. The proceeding may explain why C-FSM2 and TPC-FSM encompass the same number of parameters, otherwise considering more parameters will degrade the forecasting capabilities of TPC-FSM, which here is not the case.

Overall, TPC-FSM model performance in reproducing and inferring river flow was quite well, taking into account 2.3% of the historical observations were classified as outliers. It was able to reproduce high flows reasonably in comparison with the C-FSM2 model. Though data scarcity manifested in C-FSM2 model did not impose a problem in reconstructing the river flow in a normal situation, it did when dealing with peak flows. However, backing up the 3.5 years of rainfall-streamflow daily measurement data with an extra of 19.5 years of river flow, enabled the TPC-FSM model to exhibit more efficiency. It provided plausible results and a reasonable agreement with the observed peak streamflow. Thus, TPC-FSM was robust enough to be used in situations where data possess within a certain number of extreme measurements.

2.8 Estimating Daily Evaporation from Poorly – Monitored Lakes

Estimation of water losses, due to evaporation, is an integral component to ensure optimal short and medium term hydropower operation. In a way that, taking into account evaporation losses from the exposed water surface of the reservoir can improve the accuracy and precision of the hydropower-irrigation model.

According to Penman, the most widely recognized and imperative factors influencing evaporation are the following: sunshine, air, soil temperature, relative humidity, vapor-pressure deficit, atmospheric pressure, and wind speed (Penman, 1948). However, many studies require the consideration of the evaporation factor, but suffer from the scarcity of the required measurements. Sometimes, the only available measurements are: Temperature (T), Relative Humidity (RH), and Dew Point (D). To deal with such situation, a two-step approach is suggested to estimate evaporation rates.

At step one, using Levenberg-Marquardt (LM) algorithm, a Multi-variate Nonlinear Regression (MNR) model $D = f(T, RH)$ (f is a function based on Magnus formula) is trained and tested to estimate the dew point. Consequently, the Solar Radiation (R_S) can be approximated using a special function that relates R_S to D. At step two, considering the three inputs T, RH, and R_S, the simplified version of Penman formula provides an estimate of the lake evaporation rate (Valiantzas, 2006). Therefore, the proposed method provides a ready solution in poorly-monitored lakes. In fact, the routinely forecasted temperature and humidity will be sufficient to produce plausible evaporation prediction.

The model was applied to QD (Figure 2.23) in order to approximate evaporation losses. Up to date, there is no precise estimation of water losses in QD due to evaporation. There is only rough approximation provided by ECODIT led Consortium (ECODIT-led Consortium, 2015). It claims that Dam evaporation losses, in the case of Lebanon and for an average rainfall year, may reach 2 % of the delivered water per year. This study provides a better estimate for the water losses due to evaporation in the Dam. In fact, accurate estimations will improve and ensure effective hydropower-irrigation management.

2.8.1 Suggested Method for Estimating Evaporation

The proposed approach offers, on a daily basis, a close insight about how much evaporation is occurring in a given weakly monitored lake. Bearing in mind, the scarcity of the of required measurements in this region, a data driven model is trained using the nonlinear least square method for forecasting dew point taking temperature and relative humidity as inputs. Then, the estimated dew point along with temperature are fed into a simplified version of Penman formula to estimate the evaporation rate at the surface of the lake.

Figure 2.23 Qaraoun Dam (QD) - Lebanon (LRA, 2016).

Multi-variate Nonlinear Least Square Implementation

The main effort in this modeling process is searching for the unknown parameters of the nonlinear mathematical problem to predict the Dew point (D_i) in terms of the two independent variables: Temperature (T_i) and Relative Humidity (RH_i). A special mathematical nonlinear function f is used to model the Multi-variate Nonlinear Regression (MNR) problem:

$$D_i = f(T_i, RH_i, \theta) + e_i \tag{2.35}$$

The unstructured deviations from the function f are described via random errors e_i. The random errors are assumed to follow a Normal distribution, i.e.:

$$e_i \sim N(0, \sigma)$$

Parameters are determined based on minimizing the sum of squared errors:

$$\min_{\theta} \; S(\theta) = \sum_{i=1}^{n} [D_i - f(T_i, RH_i, \theta)]^2 \tag{2.36}$$

Now, the attention is turned to choosing the suitable nonlinear regression function f that could be a tricky step. However, the dew point is related to the ambient temperature (T), and the relative humidity (RH), by the well-known Magnus formula (Magnus, 1844) which is given in Equation 2.37:

$$D = B \frac{\ln\left(\frac{RH}{100}\right) + \frac{A \cdot T}{B+T}}{A - \ln\left(\frac{RH}{100}\right) - \frac{A \cdot T}{B+T}} \tag{2.37}$$

Several distinct values of A and B were determined in the literature (Bolton, 1980; Buck, 1981; SENSIRON company, 2006). As a consequence, Magnus Model is a data driven model. It is affected

by the collected data (temperature, relative humidity and dew-point). Therefore, it is wise to estimate the values of A and B based on data retrieved from the site of the case study.

Since useful nonlinear regression functions are often derived from the theory of the application area, the regression function that will be adopted, in this case, has the same form as Magnus formula:

$$\widehat{f}(T, RH, \theta) = \theta_2 \frac{\ln\left(\frac{RH}{100}\right) + \frac{\theta_1 \cdot T}{\theta_2 + T}}{\theta_1 - \ln\left(\frac{RH}{100}\right) - \frac{\theta_1 \cdot T}{\theta_2 + T}} \tag{2.38}$$

where $\theta = (\theta_1, \theta_2)$ is a parameter that will be estimated using LM algorithm

2.8.2 Lake Evaporation

During hydropower and irrigation operational planning, knowing the evaporation rate from a reservoir is vital in circumstances where direct measurements are not possible. In that context, the difficulty with Penman formula (Penman, 1948) is the need to have values of the four climatic elements: net-radiation intensity, atmospheric humidity, wind speed and temperature. Not all of them are commonly available in any study region. In practice, the only accessible data, and in the best cases, are about temperature and relative humidity. Based on the literature review (Valiantzas, 2006), it is better to develop an approximation to Penman formula for open water which relies solely on temperature measurements. The outcome may resemble yet another empirical formula both in appearance and in the simplicity of use, nevertheless, it should still have much of the generality of the main formula of Penman, with sufficient enough accuracy for the practical problems.
The simplified Penman formula for the evaporation rate from a lake is given in Equation 2.39 (Valiantzas, 2006):

$$E_0 \approx 0.047 R_s \sqrt{T' + 9.5} - 2.4 \left(\frac{R_s}{R_A}\right)^2 + 0.09(T' + 20)\left(1 - \frac{RH}{100}\right) \tag{2.39}$$

E_0: Evaporation rate (mm/day), T' : Daily mean temperature (^{o}C), RH : Relative Humidity (%), R_s : Solar Radiation (MJ/m^2/day) and R_A : Extraterrestrial Radiation (MJ/m^2/day).

Although temperature and humidity are routinely measured, the solar radiation can be estimated with sufficient accuracy. Valiantzas had suggested the following radiation empirical formula to estimate of solar radiation (Valiantzas, 2013):

$$R_S \approx k_{R_S} \cdot R_A \sqrt{(T_{\max} - D)} \tag{2.40}$$

$k_{R_s}$is the empirical radiation adjustment coefficient that generally, differs depending on location, from 0.12 to 0.25 (default value 0.17), $T_{\max}$: Maximum temperature (^{o}C), D : Dew point (^{o}C). Figure 2.24 gives a better insight about evaporation estimation approach and the way it works.

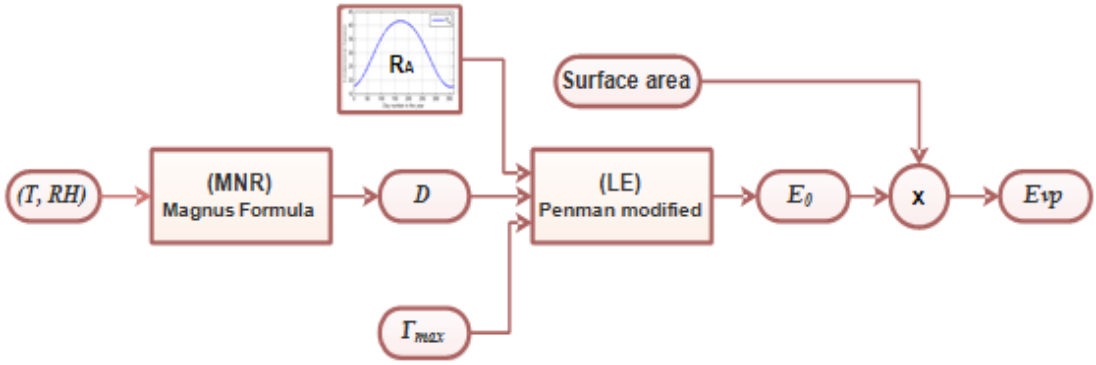


Figure 2.24 Evaporation calculation scheme.

Both estimates given in Equations 2.39 and 2.40 were discussed extensively in the work of Valiantzas (Valiantzas, 2006; Valiantzas, 2013). They were proved to be reliable approximations. Therefore, the suggested approach for estimating evaporation should produce plausible results as well.

2.8.3 Simulations and Results

The first part of this work focused on estimating the dew point. In that regards, daily climate data from 2009 till 2014 were used to train the Multivariate Nonlinear Regression model (MNR). The remaining data (year 2015) are utilized for testing the model performance.

The surface fit was obtained using MATLAB's Curve Fitting Toolbox with the Custom Equation fit type option. The custom equation is fed with the nonlinear function described in Equation 2.38. Using the LM algorithm, the fitting surface is shown in Figure 2.25.

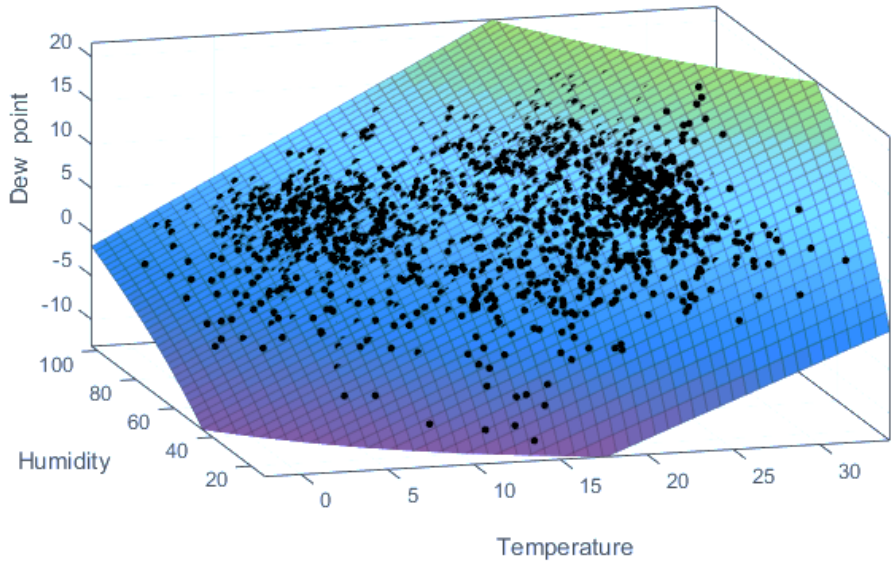


Figure 2.25 Curve fitting of the training set.

Regarding performance metrics, it is important to validate the effectiveness of the proposed NLS method applied on the MNR model. Performance was evaluated using two statistical indicators: Root Mean Square Error (*RMSE*) and Mass Curve Coefficient (*E*).

The unknown coefficients (θ_1, θ_2) of the MNR model presented in Equation 2.38 were determined based on the LM algorithm and using the available training data. The proposed method successfully determined the coefficients at 95% confidence bounds:

$$\theta_1 = 15.59 \in (14.87,\ 16.34)$$

$$\theta_2 = 230.7 \in (217.5,\ 242.9)$$

with goodness of fit: $R^2 = 0.992$ and $RMSE = 0.44$.

It is observed that dew point predictions using the trained MNR model had a very good agreement with recorded dew point (Figure 2.26) over the testing period ($R^2 = 0.997$).

Furthermore, the performance metrics given in Table 2.11 are considered for the testing period starting January till the end of December 2015 for the MNR model. It is obvious that the outcome shows a remarkable fitness. Therefore, the LM algorithm manages to give a good calibration of the nonlinear mathematical function that relates dew point, temperature and relative humidity.

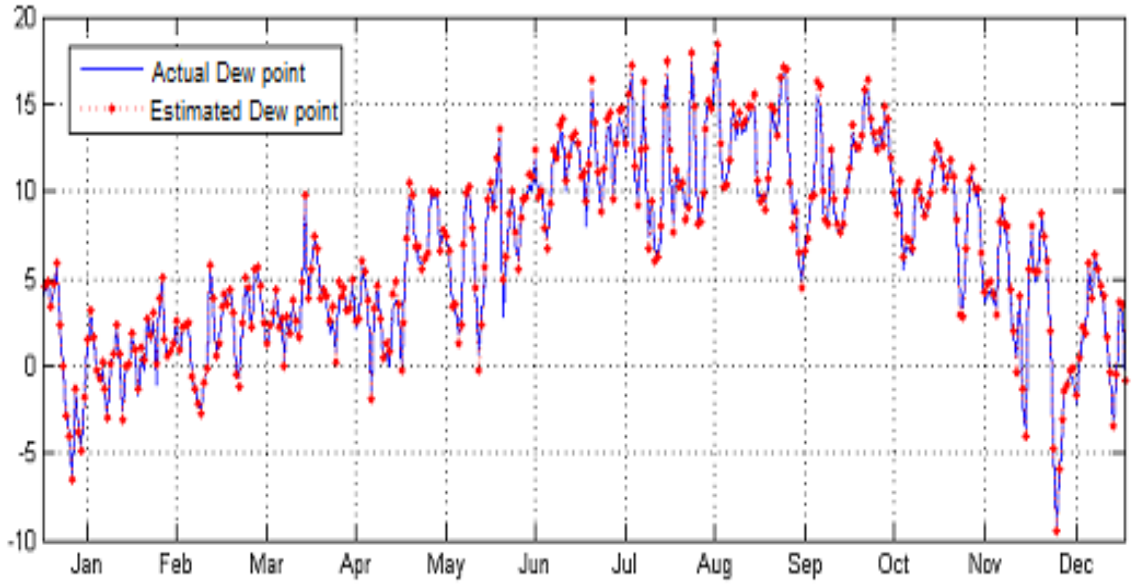


Figure 2.26 Actual versus estimated dew point over the testing period (January – December 2015).

Overall, the achieved results are accurate estimation with good fitness. This prediction accuracy is manifested in small *RMSE* and a value of *E* near 1.

In addition, Table 2.11 and Figure 2.27 confirm the assumption that the distribution of errors, joining the measured and predicted dew point values, is approximately Normal. In fact, the drawn

Table 2.11 Performance statistics.

Model	Data	Statistics		Error	
		RMSE	*E*	Skewness	Kurtosis
MNR	Testing	0.4794	0.992	-0.432	1.63

conclusion concerning Normality is based on the small values for the Skewness and Kurtosis, and on the linearity of points in the Normal Q-Q plot given in Figure 2.27.

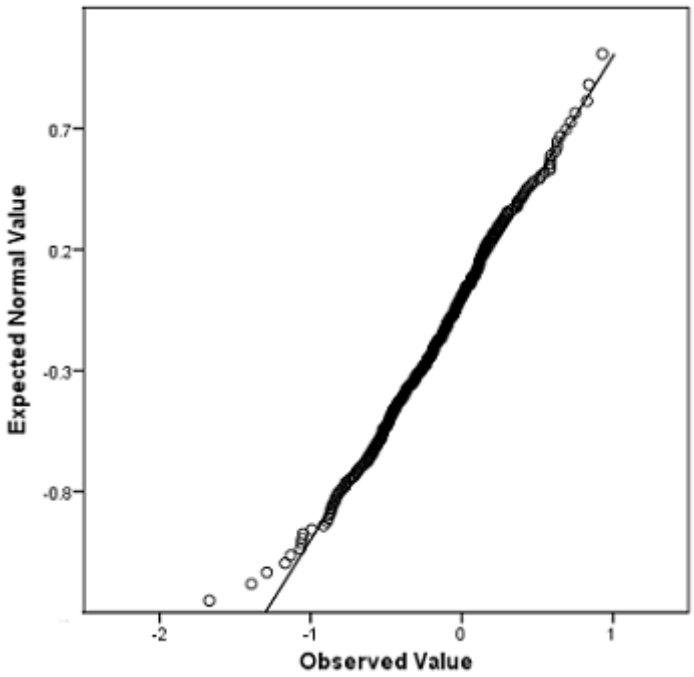


Figure 2.27 Normal Q-Q plot of error.

The remaining step is estimating the surface evaporation rate at the reservoir. But first, there is another input parameter of vital importance in the application, which is the Extraterrestrial Radiation R_A. Extraterrestrial radiation is the intensity (power) of the sun at the top of the Earth's atmosphere. It is expressed as a function of latitude, date and time of day. Well, in the studied case, the daily values of R_A throughout the year for Qaraoun region were calculated. Their plot is shown in Figure 2.28. For further information on R_A calculations, the reader may check (FAO, 2017).

During the bibliographic review, studies carried on surface evaporation on QD or nearby regions were extremely hard to come across. However, to check the plausibility of the obtained results, it is useful to seek a benchmark. Indeed, this task is challenging, especially when data concerning surface evaporation from the Dam are not available for the time being. Fortunately, there is another weather station, called Tal Amara (Figure 2.3), located in the Litani basin. It has a long-run recorded measurements for the years (1954-2002) of pan "class A" evaporation that may be very helpful in this

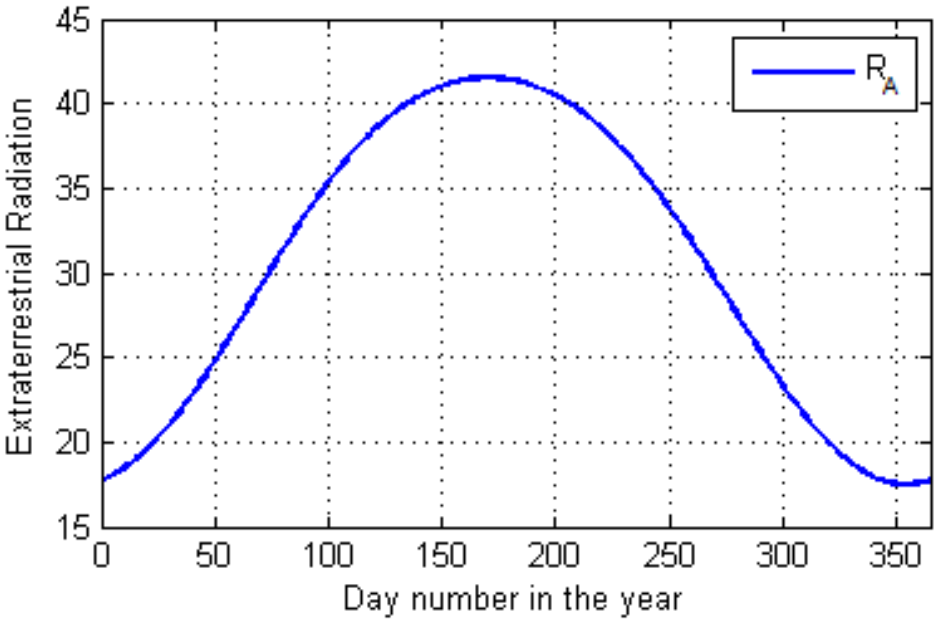


Figure 2.28 Extraterrestrial radiation at Qaraoun region.

situation. In fact, Tal Amara was considered because it shares almost the same aspects concerning location and climate with Machghara weather station as indicated in Table 2.1.

Evaporation from a water pan is usually at a higher rate because the body of water have metal sides that get hot with the sun. In addition, while light penetration in a pan is essentially uniform, light penetration in natural open bodies of water will decrease as depth increases. Therefore, Lake and pan evaporation should be treated differently. However, evaporation measurements from a Class A pan, when combined with an appropriate pan coefficient can be considered to be open-water evaporation.

According to Deodhar, both lake and pan evaporation are correlated (Deodhar, 2008), and the relationship between them can be described in the following linear equation:

$$E_0 = k \cdot E_p \tag{2.41}$$

where E_p is pan evaporation, and k is a pan coefficient. Many references suggested that k is approximately equal to 0.8 (Deodhar, 2008). Therefore, and for comparison sake, it is necessary to convert the Dam evaporation into pan evaporation using Equation 2.41. Obtained results from Qaraoun lake along with data recorded at Tal Amara are presented in Table 2.12.

Table 2.12 Pan evaporation (*mm*/month): Qaraoun versus Tal Amara.

Location	Month					
	May.	Jun.	Jul.	Aug.	Sep.	Oct.
Tal Amara	248	290	334	288	214	165
Qaraoun	290	275	327	314	264	169

Estimated evaporation values were based on the testing data of the year 2015 provided by Machghara weather measuring station. These estimations were compared with the monthly mean calculated from the long-run recorded data (1954-2002) retrieved from Tal Amara (Karam et al., 2003). Results have shown a strong correlation equal to 0.89. Table 2.12 is depicted in bar graph form in the Figure 2.29.

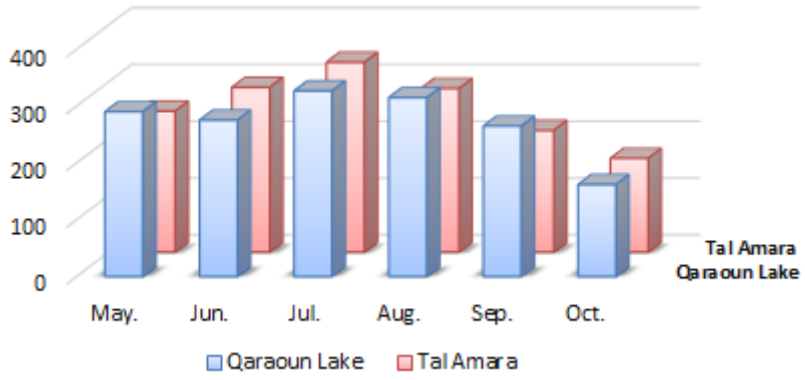


Figure 2.29 Pan evaporation (*mm*/month) starting from May till October at Qaraoun region and Tal Amara station.

The deviation from the average, between the observed and the estimated evaporation rates, for different months, are due to the inter-annual climate variability arising from day-to-day weather events. Upon looking on the obtained results, and taking into account the climate annual variations (Temperature, Humidity, Precipitation), there is apparently, a strong agreement between the predicted evaporation and the recorded ones. At stage one, the MNR model proved to be a very strong forecasting model when tested over the year 2015. Although, at stage two the evaporation data were scarce for efficient comparison with the obtained results, this issue did not impose a problem. In fact, the adopted formulas 2.39 and 2.40 in order to determine evaporation rates have been previously tested for reliability in different studies (Valiantzas, 2006; Valiantzas, 2013). Thus, the suggested approach in this work is very promising. It may provide an important tool to estimate effectively lake evaporation during hydropower planning and scheduling.

An application example

Weather data in Qaraoun region were retrieved from the Weather website (World Weather Online, 2016). It describes the climate situation in the studied region for the date of 21^{st} of July 2016 (Table 2.13).

The estimated dew point $\widehat{D}$ that corresponds to the provided information is based on the MNR model that was previously trained. It can be calculated by substituting T and RH in Equation 2.38 for $\widehat{\theta}_1 = 15.59$ and $\widehat{\theta}_2 = 230.7$,

Table 2.13 Weather data Qaraoun region.

Mean Temperature	Maximum Temperature	Relative Humidity
25.8 ^{o}C	31 ^{o}C	54.25%

$$\widehat{D} = \widehat{f}(25.8, 54.25) = 15.08\ ^{o}C$$

Using the evaporation formula in Equation 2.39, we get: $E_0 = 8.42mm/\text{day}$. In the site of QD, the water surface area $(A\left[m^2\right])$ of the reservoir can be determined in terms of the water height $(h\,[m])$ through Equation 2.42. In fact, this equation was derived by fitting the available surface-height data with a third degree polynomial:

$$A = -15.437h^3 + 2077.338h^2 + 190803.69h - 3892712 \tag{2.42}$$

If the water level at the lake is 35 m, then the estimated surface area is 4,668,297 m^2. Thus, water loss due to evaporation is approximately equal to the following quantity: $4668297 \cdot E_0/1000 = 39,307\,m^3$.

In a word, the suggested approach offers a cheap method in order to estimate evaporation, rather than installing expensive, high-tech data acquisition systems for that purpose. The section provide a tool that may help engineers and researchers to carry first a preliminary investigation before going into a full-scale study.

2.9 Conclusion

In this chapter, two types of hydrological models had been introduced: the first type is responsible for forecasting daily river flow, while the second one aims to estimate the daily evaporation rates from open lakes. The suggested models with different input variables were trained and tested using data retrieved from the Litani river project. Then the outcomes were evaluated by using different statistical indicators, namely $RMSE$, MAE, R^2 and E.

Section 2.6 discussed comprehensively the use of Fuzzy inference for multi-step-ahead daily river flow time series forecasting. Despite the scarcity, heterogeneity and non-Normality of meteorological-hydrological data aside with uncertainties inherited from illegal activities reported along the Litani river, due to factors like urbanization and industrialization, the outcomes of the C-FSM and C-FSM_MA models came very reasonable. Furthermore, the Moving Average filter has provided a supportive tool during Fuzzy modeling. It didn't just reduce the noise inherent within rainfall data, but it has also preserved the streamflow variability due to rainfall. Overall, the analysis presented in this section provides that, a variant of the C-FSM_MA model had shown a better accuracy over the rest of the models in the river flow forecasting based on the mentioned statistical indicators.

Motivated by the successful implementation of Fuzzy inference on river flow prediction, Section 2.7 adopted the same approach to deal with another delicate situation, besides heterogeneity and

non-Normality of Litani river data. It is recognized as forecasting streamflow in case of meteorological-hydrological data asymmetry. In fact, the two models C-FSM1 and C-FSM2 were trained over re-sampled data while the TPC-FSM was trained over the available asymmetric data set. Then, after validating the models, the results were compared and evaluated. Regardless the existence of extreme outliers in the hydrological data, the models TPC-FSM, C-FSM1 and C-FSM2 showed very satisfactory outcomes. But, the TPC-FSM model outperforms the other two models. However, by canceling the effect of the existing extreme values, the performance of the C-FSM2 model was quite good and even better than that of the TPC-FSM model despite hydro-meteorological data shortage.

However, in Section 2.8, a methodology was presented to estimate water losses due to evaporation using only commonly measured climate data (temperature and relative humidity). To illustrate the effectiveness and capabilities of the suggested approach, Qaraoun lake was selected as a case study. As a first step, a Multi-variate Nonlinear Regression (MNR) model was trained using the least square method. It was able to accurately predict the dew point with $R^2 = 0.99$. Afterward, the output of the model is utilized as an input for the simplified version of Penman equation. The outcome is an estimate for the surface evaporation. In fact, obtained values were compared with the monthly evaporation average retrieved from a near region to the lake known as Tal Amara. Considering the limited amount of data, estimated rates came reliable with a correlation coefficient equal to 0.89. Overall, the achieved results were accurate enough to carry out, later in the thesis, further assessment of the economic impact of evaporation losses from Qaraoun reservoir on hydropower generation and irrigation. At this level, knowing the associated financial losses will motivate engineers to submit economically feasible solutions to increase revenues by reducing the evaporation volume.

Although the available data suffer from different types of drawbacks, the data driven models based on Two-Phase Constructive Fuzzy System Modeling and Nonlinear Least Square method were successfully applied. TPC-FSM and NLS aim to establish river flow and evaporation, respectively, at QD. Their promising results were a motivation to adopt the TPC-FSM model for generating future streamflow and the NLS model to estimate evaporation water losses as a part of the water balance equation in order to improve the hydropower operation scheduling. The forecasts will help in finding an accurate optimal operation policy, at different stages, through the right discharge decisions.

Chapter 3

Irrigation and Agricultural Modeling

> The discovery of agriculture was the first big step toward a civilized life.
>
> *Arthur Keith*

3.1 Introduction

Demand on water is increasing exponentially due to extraordinary population and industrial growth. The supply is, therefore, far less than the actual demand and further its existence is being threatened by the adverse effects of climate change. Water resources management in the next decade is inevitable and should be every nation's primary objective. In fact, there is a growing interest to develop advanced management methods to prevent wasting water in the course of satisfying human needs, protecting health, ensuring food production, restoring of ecosystems, as well as for social economic evolvement and for sustainable development. Programs are launched by the European Union through the Common Agricultural Policy (CAP) or internationally through the Consortium of International Agricultural Research Centers (CGIAR) in order to overcome these crucial issues.

Nowadays, there is a great urge to present new irrigation technologies in agricultural research. So multiple optimization methods are suggested to: find the right crop selection, implement crop rotation and schedule precise irrigation. Traditionally, agricultural models primarily focused on maximizing the yield and the economic return per unit area by allocating water to different crops according to their water needs (Afshar and Mariño, 1989; Onta et al., 1995; Garg and Ali, 1998). With time, the studies have switched to deficit irrigation and its impact on crop yield production. The objective was regulating deficit irrigation in a way to save water by subjecting crops to periods of moisture stress with minimal effects on yields. Within this approach and based on FAO reports, the reduction in the yield may be little, compared with the benefits obtained through diverting the saved water to cover wider cropped area (FAO - Water Report, 2002). Garga claims that optimal irrigation is useful in increasing the crop production, the irrigated area and the net economic returns (Garga and Dadhich, 2014).

In this chapter two complementary mathematical models were developed for agricultural planning. The aim is to provide answers for the following questions:

1. What is the optimal water allocation, land resources and crop pattern taking into account water availability, crop production, and cropping area constraints?
2. What is the maximum attainable profit?
3. How can the obtained profit be distributed among active farmers?
4. How to optimize the distribution of crops over farmlands bearing in mind crop rotation?

In fact, the provided models were formulated taking into consideration the addressed questions. The first model is Multi-Crop Planning (MCP) implemented under deficit irrigation. The aim is divided into two sub-objectives: 1- maximizing profit of the system using the available water resources and by suggesting an optimal multi-cropping pattern; 2- generating optimal irrigation profile that can be used during medium term hydropower planning. The second model is the Profit Distribution (PD) paradigm. It aims to distribute the obtained profit by the MCP model among active farmers. Figure 3.1 describes the framework of both models.

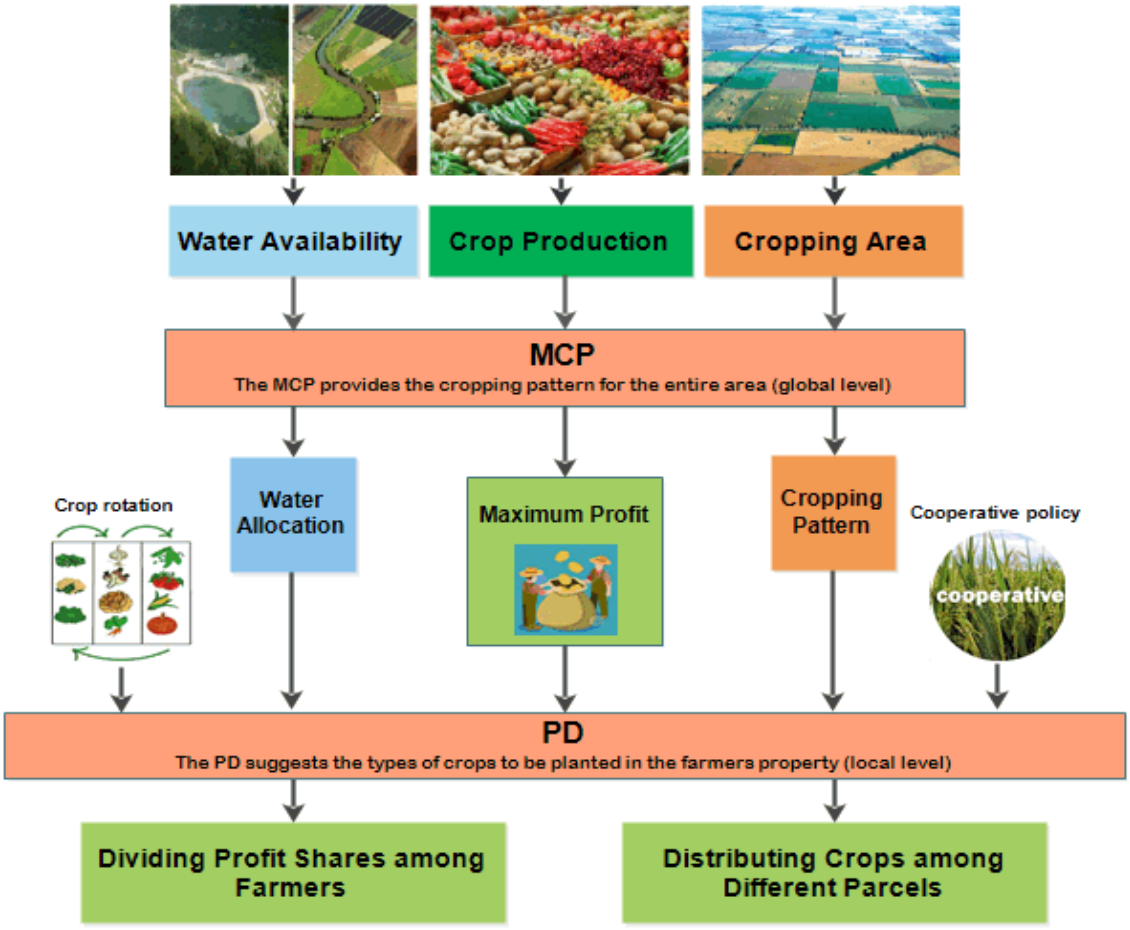


Figure 3.1 The two complementary models: MCP and PD.

List of Symbols

Notation	Description	Unit
Parameters		
l	Number of stages	Unitless
n	Number of crops	Unitless
m	Number of farmlands	Unitless
A_{total}	Total area of agricultural region	ha
A_k	Area of the farmland k	ha
r_j	Available amount of water at stage j	m^3
c_i	Production bound for crop i	Kg
p_i	Selling price of crop i	€/ Kg
pw	Cost of water	$€/m^3$
Ym_i	Potential amount of crop i	Kg/ha
Etm_{ij}	Potential evapotranspiration of crop i at stage j	mm
WR_{ij}	Required water amount for crop i at stage j	m^3/ha
ky_{ij}	Yield response factor of crop i at stage j	Unitless
λ_{ij}	Index of sensitivity of crop i at stage j	Unitless
β_{ik}	Proportion of crop i of parcel k grown in the previous season	Unitless
C_i	Miscellaneous costs to plant crop i	$€/ha$
s_k	Maximum number of crops that parcel k can accommodate	Unitless
Decision variables		
X_i	Area of the land planted with crop i	ha
ETa_{ij}	Actual evapotranspiration of crop i at stage j	mm
WA_{ij}	Applied water amount to crop i at stage j	m^3/ha
α_{ik}	Proportion ratio of crop i planted in parcel k	Unitless
y_{ik}	$y_{ik} = 1$ if crop i is assigned to parcel k and $y_{ik} = 0$ otherwise	Unitless
Dependent variables		
Ya_i	Produced amount of crop i	Kg/ha
P_i	Profit obtained when crop i is assigned to area X_i	€
B_i	Cost of used water by crop i	$€/ha$
Q_k	Financial benefits acquired from planting parcel k	€
R_k	Profit share of farmer k	€
a_i	Financial gain per hectare from planting crop i	$€/ha$

3.2 Multi-Crop Planning Implemented Under Deficit Irrigation

In this section, a Decision Support Tool (DST) based on a Non-Linear Programming (NLP) model for optimal multi-crop planning is proposed. The aim is to maximize the net financial returns. In fact, the presented objective function in the optimization model was inspired by a relation found in the literature (Jensen, 1968; Wardlaw and Barnes, 1999). Furthermore, the model considers water limitation at each

time period (Garga and Dadhich, 2014). Another restriction ought to be considered here is the crop production quota which is important to preserve crop diversity (Greening rule - CAP). Otherwise, farmers will grow the most profitable plant, leading to agricultural surplus in some crops and shortage in others. Nevertheless, this greening rule will ensure market stability and will secure the availability of supplies. In response to the above conditions, a Multi-Crop Planning (MCP) model is presented. It is utilized to find an optimal water allocation and a desirable crop pattern to maximize financial profit.

In the past decade, comprehensive studies have been conducted on Evolutionary Algorithms (EA) for solving nonlinear programming problems concerning optimal crop planning and irrigation water allocation. For example, Genetic Algorithm (GA) was used to solve such problems in the literature (Álvarez et al., 2004; Kumar et al., 2006), while the literature (Georgiou et al., 2006) search for the optimal irrigation reservoir operation using Simulated Annealing (SA). On the other hand, Particle Swarm Optimization (PSO) algorithm was applied to find the optimal reservoir operation for the irrigation of multiple crops (Kumar and Reddy, 2007; Noory et al., 2012).

However, the problem-solution approach presented in this investigation is as follows: preliminary mathematical tools are addressed that involves solution existence, linear model extraction and a relaxation formulation of NLP. Afterward, two meta-heuristic algorithms SA and PSO are implemented as a numerical technique for solving the MCP problem. The approach trait is using the solution of the linear problem as an initial guess for the SA, while for PSO the particle swarm is initiated in the neighborhood of that solution, rather than generating them randomly (Georgiou et al., 2006; Kumar and Reddy, 2007). The effectiveness of the suggested models is tested and evaluated using experimental and real data.

3.2.1 MCP Model Formulation

Objective Function

The objective function of the MCP model is the net profit from crop production, which is calculated by subtracting the total cost (manual labor, seeds, fertilizers, water used,...etc.) from the market value of the yield. However, as a first step, one should establish the water-crop relationship which contains timing, quantity of water applied and the effects of crop to water stress on yield at different growth stages. A widely used relation is presented by Jensen (Jensen, 1968) and it is expressed in the following formula:

$$\frac{Ya_i}{Ym_i} = \prod_{j=1}^{l} \left(\frac{ETa_{ij}}{ETm_{ij}} \right)^{\lambda_{ij}} \tag{3.1}$$

whereas Noory presents a linear relationship between relative yield and relative evapotranspiration (Noory et al., 2012). It empirically derives yield response factors[1] (k_y) for individual growth stages (i.e. establishment, vegetative, flowering, yield formation and ripening). However, Jensen's model

[1] According to FAO, yield response factor (k_y) represents the effect of the reduction in evapotranspiration on yield losses.

(Equation 3.1) can be applied at time steps smaller than the growth stages. Its sensitivity indices λ_{ij} are related to the yield response factors ((ky_{ij})) (Georgiou et al., 2006) by the following polynomial:

$$\lambda_{ij} = 0.2418(ky_{ij})^3 - 0.1768(ky_{ij})^2 + 0.9464(ky_{ij}) - 0.0177 \tag{3.2}$$

The polynomial is obtained for a coefficient of determination $R^2 = 0.999$.
In this work, the water balance in soil is not considered. According to Wardlaw, it can be assumed that the ratio of the actual crop evapotranspiration to potential crop evapotranspiration is the same as the ratio of irrigation supply to demand (Wardlaw and Barnes, 1999), that is:

$$\frac{ETa_{ij}}{ETm_{ij}} = \frac{WA_{ij}}{WR_{ij}} \tag{3.3}$$

Then, combining Equations 3.1 and 3.3, the outcome is a crop-water relationship:

$$Ya_i = Ym_i \prod_{j=1}^{l} \left(\frac{WA_{ij}}{WR_{ij}} \right)^{\lambda_{ij}} \tag{3.4}$$

The profit obtained from planting crop i can be determined using the formula:

$$P_i = [p_i Ya_i - (B_i + C_i)] X_i \tag{3.5}$$

Therefore, writing P_i in terms of the variables X_i and WA_{ij}, Equation 3.5 becomes:

$$P_i = \left[p_i Ym_i \prod_{j=1}^{l} \left(\frac{WA_{ij}}{WR_{ij}} \right)^{\lambda_{ij}} - (pw \sum_{j=1}^{l} WA_{ij} + C_i) \right] X_i \tag{3.6}$$

Hence, the objective function is now given by:

$$F = \sum_{i=1}^{n} P_i \tag{3.7}$$

Constraints

The objective function is bounded by a set of constraints: water limitations at each stage, crop water requirement, crop-water relationship, crop production quota and total agricultural area.

1. If the amount of water available at each time step or stage is limited to a fixed quantity r_j, for $j = 1, \ldots, l$, then it is important to consider water limit constraint:

$$\sum_{i=1}^{n} X_i \cdot WA_{ij} \leq r_j, \text{ for all } j = 1, \ldots, l \tag{3.8}$$

However, in regions with abundant water resources, the availability of water is not a problem. In this case, water is provided with no limits, so this constraint is no more restrictive and could be omitted.

2. Under deficit irrigation, the applied water can not exceed the required amount:

$$WA_{ij} \leq WR_{ij}, \text{ for all } i = 1,...,n \text{ and } j = 1,...,l \tag{3.9}$$

however, in a state of full irrigation equality is assumed.

3. Crop-water relationship:

$$Ya_i = Ym_i \prod_{j=1}^{l} \left(\frac{WA_{ij}}{WR_{ij}} \right)^{\lambda_{ij}}, \text{ for all } i = 1,...,n \tag{3.10}$$

4. The produced quantity of crop i can not be greater than a fixed quantity c_i:

$$X_i Ya_i \leq c_i, \text{ for all } i = 1,...,n \tag{3.11}$$

This condition is vital for two main reasons: it maintains diversity of crops and keeps the market values of crops stable. In fact, any overproduction of a certain crop can cause a decrease in price which is here not the case.

5. Area constraint:

$$\sum_{i=1}^{n} X_i \leq A_{total} \tag{3.12}$$

6. Non-negativity constraints:

$$WA_{ij} \geq 0, \text{ for } i = 1,...,n \text{ and } j = 1,...,l \tag{3.13}$$

$$X_i \geq 0, \text{ for } i = 1,...,n \tag{3.14}$$

The Nonlinear Optimization Problem

In the foregoing, the decision variables are the planted areas $X_i's$ and the applied water at each stage WA_{ij}. Let us denoted by $S \subset \mathbb{R}^{n+l \times n}$, the set of all points that satisfy constraints 3.8–3.14, then the Nonlinear Programming NLP problem becomes:

$$\max_{(X,WA) \in S} F(X,WA)$$

where $X = (X_1, X_2, ..., X_n)$ and $WA = vec(WA_{ij})$. Recall *vec* of a matrix is a linear mapping which converts the matrix into a column vector.

3.2.2 Mathematical Approach

In the following and before proceeding to the numerical part, a preliminary theoretical study is carried out. It involves existence, benchmark linear models and a relaxation formulation of the main NLP problem.

Proposition 1 *(Existence) The* NLP *problem admits a solution in S.*

Proof. Since $\lambda_{ij} > 0$, the function F defined from $\mathbb{R}^{n+l\times n}$ into $\mathbb{R}$ is a continuous function. Moreover, it is obvious that the set S is closed and bounded, thus S is a compact set. By Weierstrass theorem, F attains its global maximum in S. ■

Proposition 2 *In case of no water limits and full irrigation, the* NLP *problem is transformed into a Linear Programming problem (*LP1*) and has the form:*

$$(\text{LP1}) \begin{cases} \max\limits_{X} f^T X \\ \sum_{i=1}^{n} X_i \leq A_{total} \\ 0 \leq X_i \leq c_i/Ym_i, \text{ for all } i = 1, ..., n \end{cases}$$

where f is a vector in $\mathbb{R}^n$ such that for every $i = 1, ..., n$, $f_i = p_i Ym_i - (pw \sum\limits_{j=1}^{l} WR_{ij} + C_i)$, $X = (X_1, X_2, ..., X_n)$.

Proof. No water limits with full irrigation means that constraint 3.8 is omitted and further, for all $i = 1, ..., n$ and $j = 1, ..., l$, we have $WA_{ij} = WR_{ij}$. So, $\frac{WA_{ij}}{WR_{ij}} = 1$, Ya_i=Ym_i. In this case, the profit from crop i becomes $f_i = p_i Ym_i - (pw \sum\limits_{j=1}^{l} WR_{ij} + C_i)$. As a result, the objective function and the nonlinear constraints become linear, and the problem is transformed into the form:

$$\begin{cases} \max\limits_{X} \sum\limits_{i=1}^{n} f_i X_i \\ X_i Ym_i \leq c_i, \forall i = 1, ..., n \\ \sum\limits_{i=1}^{n} X_i \leq A_{total} \\ X_i \geq 0, \text{ for all } i = 1, ..., n \end{cases}$$

By setting $f = (f_1, f_2, ..., f_n)$ and $X = (X_1, X_2, ..., X_n)$, the above linear problem is reshaped into LP1. ■

Proposition 3 *Suppose full irrigation is considered within a limited amount of available water. The* NLP *problem is transformed into a Linear Programming model (*LP2*) and it has the form:*

$$(\text{LP2}) \begin{cases} \max\limits_{X} f^T X \\ \sum\limits_{i=1}^{n} X_i \cdot WR_{ij} \leq r_j, \ \forall \ \ j = 1, ..., l \\ \sum\limits_{i=1}^{n} X_i \leq A_{total} \\ 0 \leq X_i \leq c_i / Ym_i, \ \forall i = 1, ..., n \end{cases}$$

Proof. The proof is carried in same way as that of proposition 2, but constraint 3.8 should be put back due limited amount of water. ■

The solution of the LP1 and that of LP2 model presents the optimal crop distribution with respect to two different scenarios: the first one, full irrigation with no water limits while the second case is full irrigation with water amount limitations.

Remark 1 *These two models are intended to provide a reference for the* NLP *problem. In fact,* LP2 *gives a lower bound for the* NLP *objective function, that is:*

$$\max_{(X,WA) \in S} F(X, WA) \geq \max_{X \in S'} f^T X$$

where S' is the set of all points that satisfy constraints of LP2.

Recall 1: Denote by E the objective function to be maximized over a domain D. A relaxation of a maximization problem:

$$z = \max \{E(x); \ x \in D \subset \mathbb{R}^n\}$$

is another maximization problem of the form:

$$z_R = \max \{E_R(x); \ x \in D_R \subset \mathbb{R}^n\}$$

with the following properties: $D_R \supseteq D$ and $E_R(x) \geq E(x)$ for all $x \in D$.

Let us rearrange F :

$$F = \sum_{i=1}^{n} \left[p_i Ya_i - (pw \sum_{j=1}^{l} WA_{ij} + C_i) \right] X_i$$

$$F = \sum_{i=1}^{n} p_i Ya_i X_i - \sum_{i=1}^{n} (pw \sum_{j=1}^{l} WA_{ij} + C_i) X_i$$

Now define the new objective function:

$$F_R = \sum_{i=1}^{n} p_i \min(Ya_i X_i, c_i) - \sum_{i=1}^{n} (pw \sum_{j=1}^{l} WA_{ij} + C_i) X_i \tag{3.15}$$

Proposition 4 *The nonlinear programming problem defined as:*

$$\max_{(X,WA)\in S_R} F_R(X,WA)$$

is the relaxed version of the NLP *problem and it is denoted by* RNLP*, where* $S_R \subset \mathbb{R}^{n+l\times n}$ *is the set of points that satisfy constraints 3.8–3.10, 3.12–3.14.*

Proof. It is clear that $S \subset S_R$ and for all $(X,WA) \in S$, we have $F_R(X,WA) = F(X,WA)$. Thus, RNLP is the relaxation of the NLP problem. ■

Proposition 5 *The* NLP *and* RNLP *problems admit the same global solution.*

Proof. In this proposition, it is enough to prove that the RNLP problem attains its maximum in S. So, let assume that (X^*,WA^*) is the global solution of the RNLP problem such that $X^*_{i_0} \cdot Y^* a_{i_0} > c_{i_0}$ for some i_0 (*i.e.* $(X^*,WA^*) \in S_R - S$), here $Y^* a_{i_0} = Ym_{i_0} \prod_{j=1}^{l} \left(WA^*_{i_0 j}/WR_{i_0 j}\right)^{\lambda_{i_0 j}}$.
Define $\Delta X'_{i_0}$ such that $Y^* a_{i_0} \Delta X'_{i_0} := X^*_{i_0} \cdot Y^* a_{i_0} - c_{i_0} > 0$. Divide by $Y^* a_{i_0}$, we get:

$$\Delta X'_{i_0} = X^*_{i_0} - \frac{c_{i_0}}{Y^* a_{i_0}} < X^*_{i_0} \tag{3.16}$$

Set $X' := (X^*_1, X^*_2, ..., X^*_{i_0-1}, X'_{i_0}, X^*_{i_0+1} ..., X^*_n)$ such that $X'_{i_0} = X^*_{i_0} - \Delta X'_{i_0}$. It is obvious that:

$$X'_{i_0} < X^*_{i_0} \tag{3.17}$$

.
Moreover,

$$X'_{i_0} \cdot Y^* a_{i_0} = (X^*_{i_0} - \Delta X'_{i_0}) \cdot Y^* a_{i_0} = X^*_{i_0} \cdot Y^* a_{i_0} - \left(X^*_{i_0} \cdot Y^* a_{i_0} - c_{i_0}\right) = c_{i_0} \tag{3.18}$$

Well, the reader can easily check that $(X',WA^*) \in S_R$.
Now using Equation 3.18, we get:

$$\sum_{i=1}^{n} p_i \min(Ya_i X'_i, c_i) = \sum_{i=1}^{n} p_i \min(Ya_i X^*_i, c_i) \tag{3.19}$$

while by using Inequation 3.17, we have:

$$\sum_{i=1}^{n} (pw \sum_{j=1}^{l} WA^*_{ij} + C_i) X'_i < \sum_{i=1}^{n} (pw \sum_{j=1}^{l} WA^*_{ij} + C_i) X^*_i \tag{3.20}$$

Therefore,

$$F_R(X', WA^*) > F_R(X^*, WA^*) \tag{3.21}$$

which contradicts our assumption. So, any global solution of the RNLP problem must satisfy constraint 3.11. Thus, the RNLP problem attains its maximum in S. However, NLP and RNLP problems are identical on the set S. As a result, both problems admit the same global solution. ■

3.2.3 Meta-heuristic Methods

The optimization of the objective functions addressed in both NLP and RNLP is a problem without obvious analytical solution and perhaps with multiple local optimum. In the recent years, Evolutionary Algorithms (EA) have become popular tools for nonlinear optimization problems. In fact, two Evolutionary algorithms are used in this work: the first is the Particle Swarm Optimization (PSO) and the second is the Simulated Annealing (SA) algorithm.

Particle Swarm Optimization

Particle Swarm Optimization was first presented by Kennedy and Eberhart in 1995. The PSO search procedures are based on the swarm concept (inspired by social behaviors of bird flocking or fish schooling), which is a group of individuals that are able to optimize certain fitness function. Every individual can transmit information to another and ultimately allow the entire group to move towards the same object or in the same direction. It is a method to simulate the behavior of individuals of the species who work for the benefit of the whole group.

PSO is initialized with a population of random solutions creating a particle swarm and searches for optima by updating generations. Each particle keeps track of its coordinates in the search space which are related to the best solution (fitness) it has reached so far. This value is referred to *Pbest*. Another "best" value that is tracked by the particle swarm optimizer is the best value obtained so far in all of the particle swarm. This best value is a global best and is named *Gbest*. To find the optimal solution, each particle moves in the direction of its *Pbest* and *Gbest*. After continuous iterations, the particle swarm will gravitate towards the optimum solution. The parts of PSO are given below:

(1) *Velocity update*

$$v_i^{k+1} = \omega v_i^k + c_1 \cdot \text{rand} \cdot (Pbest_i^k - x_i^k) + c_2 \cdot \text{rand} \cdot (Gbest^k - x_i^k)$$

- c_1 and c_2 are learning factors of PSO
- rand is a random number uniformly distributed
- $Pbest_i^k$ individual best optima for particle i after k iterations

- $Gbest^k$ group optima after k iterations
- ω inertia weight
- v_i^{k+1} velocity of particle i in iteration $k+1$
- x_i^k position of particle i in iteration k

(2) *Position update*

- $$x_i^{k+1} = x_i^k + v_i^{k+1}$$

(3) *Weighting*

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{Iter_{\max}} \cdot Iter$$

- $\omega_{\max}$ largest weight
- $\omega_{\min}$ smallest weight
- $Iter$ iterative times
- $Iter_{\max}$ maximum iterative times for PSO

Simulated Annealing

The basic concept of SA was inspired from statistical thermodynamics by American physicist Metropolis in 1953. In 1983, Kirkpatrick suggested using this concept for finding solutions for optimization problems, and was the first literature to successfully utilize SA in combinatorial optimized problems. Theoretically, the SA algorithm guarantees the convergence to a global optimal solution under certain assumptions and given infinite execution time. In practice, however, globally optimal or near-optimal solutions can be obtained in a large yet finite number of iterations (Georgiou et al., 2006).
Before describing the SA algorithm, let us denote by $E(x)$ the function being minimized, where x is the vector of decision variables of dimension d. Then, the basic steps of the algorithm are the following:

1. Set the step number $k = 0$. Choose the initial temperature T_0 and an initial state x_0. Afterward, calculate its energy E_0.
2. Find a feasible candidate state $x_{k+1} = x_k + \triangle x$, $\triangle x$ is randomly generated from a normal distribution of mean 0 and variance T_k
3. Calculate E_{k+1} and the energy difference $\triangle E = E_{k+1} - E_k$

4. If $\triangle E < 0$, state improvement, update solution; otherwise accept it, if $random(0,1) \leq e^{(-\triangle E/T_k)}$
5. Lower the temperature with the geometric cooling scheme proposed by (Kirkpatrick et al., 1983), $T_{k+1} = \alpha T_k$, $\alpha \in (0,1)$
6. Increment k and repeat steps 2–6 until k satisfies some specified stopping criterion.

An essential parameter that should be considered during the implementation of SA is the initial temperature T_0. If it is set too low, the randomly generated candidate states will never be far from the initial state and the search space will not be properly explored. In this case, the algorithm is trapped into local minima, we term it as premature convergence. On the other hand, if it is set too high, the vast majority of candidate states will be rejected as infeasible and the generation procedure will be very inefficient. Kirkpatrick suggested that a suitable initial temperature is one that results in an average increase acceptance probability of about 0.8. The value of T_0 will clearly depend on the scaling of cost function. Hence the determination of T_0 is problem-specific.

3.2.4 Implementation and Approach

The solution approach is carried through the following steps:

- Solve LP2. Denote by $X^* = (X_1^*, ..., X_n^*)$ solution of LP2 and by $WX^* = vec(\chi_i \cdot WR_{ij})$, where $\chi_i = \begin{cases} 0 & \text{if} \quad X_i^* = 0 \\ 1 & \text{if} \quad X_i^* \neq 0 \end{cases}$ and *vec* of a matrix is a linear mapping which converts the matrix into a column vector.
- Solve the NLP problem using PSO by initiating a particle swarm in the neighborhood of (X^*, WX^*).
- Resolve NLP and RNLP using SA with (X^*, WX^*) as initial guess.
- Compare the different approaches, then choose the best one.

3.2.5 Simulations and Discussions

Numerical Example

In this part, an experimental evaluation is performed for each algorithm in order to compute the optimal cropping pattern and irrigation scheduling for a lot of six crops. However, optimal solution determination requires knowledge about the area of agricultural region A_{total} , availability of water r_j at each stage $j = 1,2,3,4$ during the irrigation season and of course the crop characteristics k_y.

Under the given conditions concerning the available water resources and crop production, the models were tested over a spread acreage covering 322 acres, whereas the available amount of water during the four stages is 245000 m^3.

The programs were coded in MATLAB language and ran on Intel Core i7-5500U CPU @ 2.40 GHZ, 12.0 GB RAM. Furthermore, the main parameters needed for the optimization procedure for the models NLP and RNLP are obtained after some experimentation. For example, the initial temperature parameter for the SA was set to 23 based on Kirkpatrick suggestion. This is assumed to be high enough to avoid getting stuck in a local maxima and to allow the initial exploration of the solution space without generating excessive numbers of infeasible candidate states.

Here, the initial guess for SA algorithm is either: 1- randomly generated in the search space (NLP-SA; RNLP-SA); or 2- the solution of the LP2 problem (NLP-SA1; RNLP-SA1). For the PSO algorithm, $\omega_{\max} = \omega_{\min} = 0.5$ and the number of particles was set to 1000, where the particles position were initialized either: 1- randomly over the whole search space (NLP-PSO); or 2- with a uniformly distributed random vector in the neighborhood of the solution obtained by the LP2 model (NLP-PSO1). Each scenario (Figure 3.2), for the nonlinear models, is run for 15 times. The top results are presented in Table 3.1.

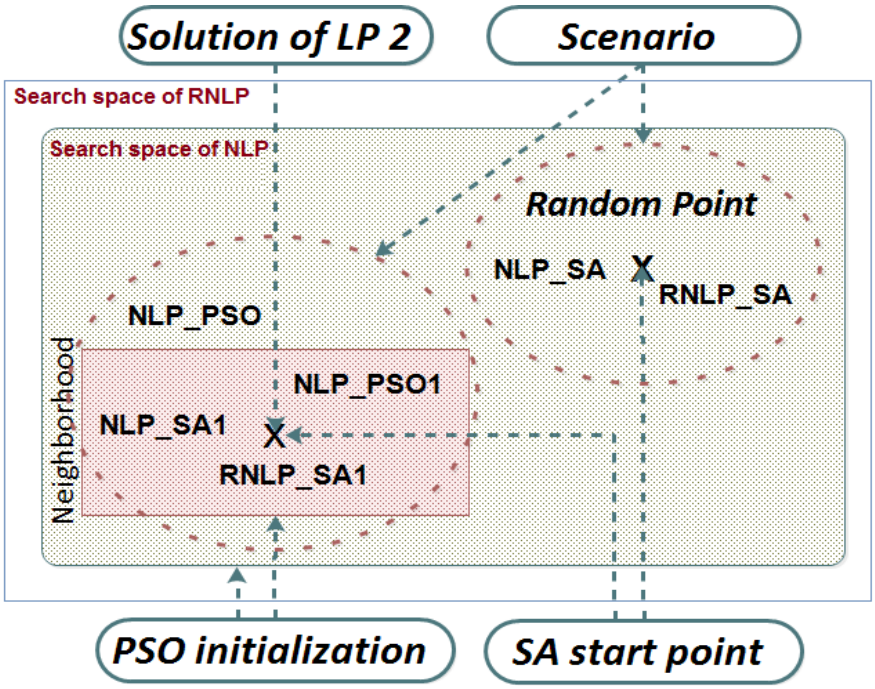


Figure 3.2 Algorithms initialization scenarios.

Results and Discussions

Table 3.1 shows the recommended optimum crop pattern found by each model and the consumed amount of water. In case of water abundance, full irrigation is applied and the net profit gained is equal to €626240. However, implementing the same strategy (full irrigation) in case of limited water resources, the optimal net profit has decreased to €345430 and the crop spread area has reduced to 44.3%.

Table 3.1 Experimental results.

Scenerio Name	**Problem**	**Algorithm Used**	**Initialization** Starting point	**Area in Acres of each crop i** 1	2	3	4	5	6	**Irrigated Area (ac)**	**Consumed Water** (m^3)	**Profit** (€)	**Elapsed Time (sec)**
LP1	LP1	Interior-Point	Random	0	0	0	166.7	114.0	41.0	322	418600	626240	0.02
LP2	LP2	Interior-Point	Random	0	0	0	28.6	114.0	0	142.6	185710	345430	0.03
NLP_SA	NLP	Simulated Annealing	Random	12.7	0.4	0.8	83.4	108.5	10.1	215.9	242700	400297	817
RNLP_SA	RNLP	Simulated Annealing	Random	51.5	0.5	0.5	77.4	117.1	1.2	248.2	241740	402555	2052
NLP_PSO	NLP	Particle Swarm Optimization	Random	20.8	2.6	5.2	125.3	91.4	46.9	292.2	245000	346512	854
NLP_SA1	NLP	Simulated Annealing	Sol. LP2	9.2	1.3	3.1	82.3	121.2	0.2	217.4	241420	405356	759
RNLP_SA1	RNLP	Simulated Annealing	Sol. LP2	6.1	1.2	0.5	85.1	115.2	2.5	210.5	241500	408199	1008
NLP_PSO1	NLP	Particle Swarm Optimization	Sol. LP2	3.6	1.3	1.3	70.9	116.4	1.8	195.3	232230	401373	498

Well, considering the water limitation (245000 m^3), deficit irrigation and crop pattern re-arrangement is the alternative option. Despite the water shortage, the optimal net financial returns obtained from the nonlinear programming models has reached 65.2% with respect to the profit obtained by LP1 (Figure 3.3). Regarding the crop spread area, it managed to cover 65.4 % of the total area. Moreover, it is found that crops 4 and 5 have the highest plantation areas according to the solutions of LP1 and LP2. In addition, crops 1, 2 and 3 were avoided because they produce less profit. Therefore, it leads us to consider crops 4 and 5 as dominating crops. Motivated by these dominating crops, the SA and PSO algorithms were launched with the aid of the solution of LP2 as a starting point for the schemes NLP_SA1, RNLP_SA1 and NLP_PSO1. In fact, this guarantees that the achieved profit in the worst case scenario is at least that of LP2 and not below. Since the water supplied to the farmers is less than

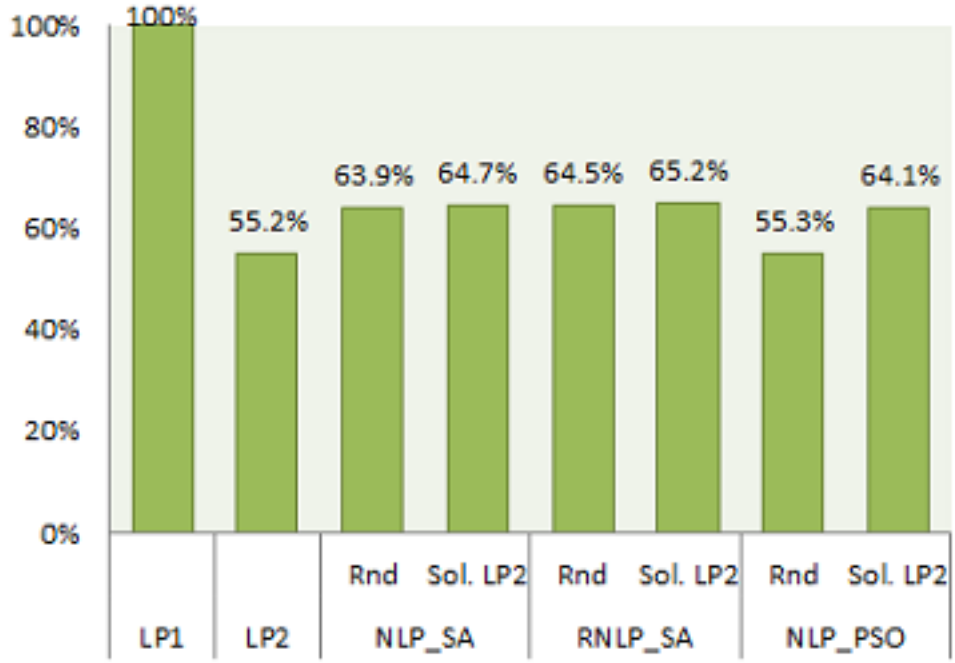


Figure 3.3 Profit function comparison between different models.

desired, the solution that could be adopted by a farmer is the one that provides the greatest revenue. In this case and based on the obtained results, the RNLP_SA1 scheme is recommended.

In what follows, the performance measure for both algorithms SA and PSO is examined for different initialization methods using the Coefficient of Variation (CV) that is given in Equation 3.22:

$$CV = \frac{\sigma}{\mu} \tag{3.22}$$

where μ is the mean and σ standard deviation. In fact, CV exhibits the extent of variability in relation to the mean.

Table 3.2 and Figures 3.4, 3.5 show, after running each scheme 15 times, the mean of the maximized objective functions, the average execution time and the corresponding CV. It is obvious that when

the algorithm was initialized with the aid of the solution of LP2 problem, the maximized objective functions exhibit smaller deviation from the mean. However, when the algorithm was initialized with a random starting point, it was less reliable and in most of the times it had converged to a local maximum.

Table 3.2 Performance of each scheme.

Model	Mean Objective Function	Coefficient of Variation	Mean Elapsed Time	Coefficient of Variation
NLP_SA	383206	0.048	461	0.556
RNLP_SA	367594	0.139	1985	0.778
NLP_PSO	305785	0.119	983	0.187
NLP_SA1	396460	0.016	561	0.493
RNLP_SA1	404858	0.009	1269	0.388
NLP_PSO1	367388	0.075	483	0.020

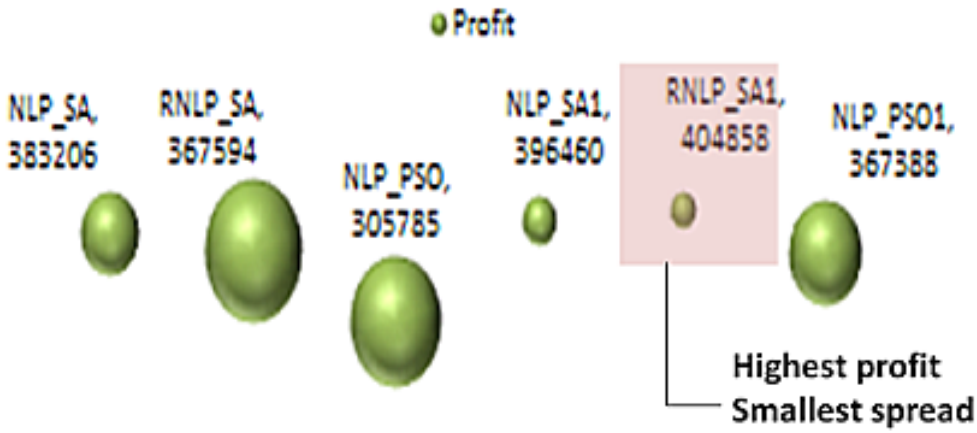


Figure 3.4 Mean net profit in euros; the size of the ball represents the standard deviation of the profit.

According to Tables 3.1 and 3.2, the profit maximization and solution variability of RNLP_SA1 model came very promising when compared with the rest of scenarios. The RNLP_SA1 possesses the highest profit with the lowest variance at a reasonable execution time. However, if one is seeking a faster running time in the nonlinear problem category, the NLP_PSO1 may be the next preferred choice. The execution time, on average, is reduced by 50.9% relative to NLP_PSO and by 62% relative to RNLP_SA1. In contrast, the top maximized profit obtained by NLP_PSO1 is €401373 with a mean value of €367387 and $CV = 0.075$. This small average profit in comparison with the other ones may be explained by the fact that, basic PSO can be easily trapped into a local optimum. Another competing alternative regarding runtime is LP2. Its solution is obtained almost instantly and the achieved profit is

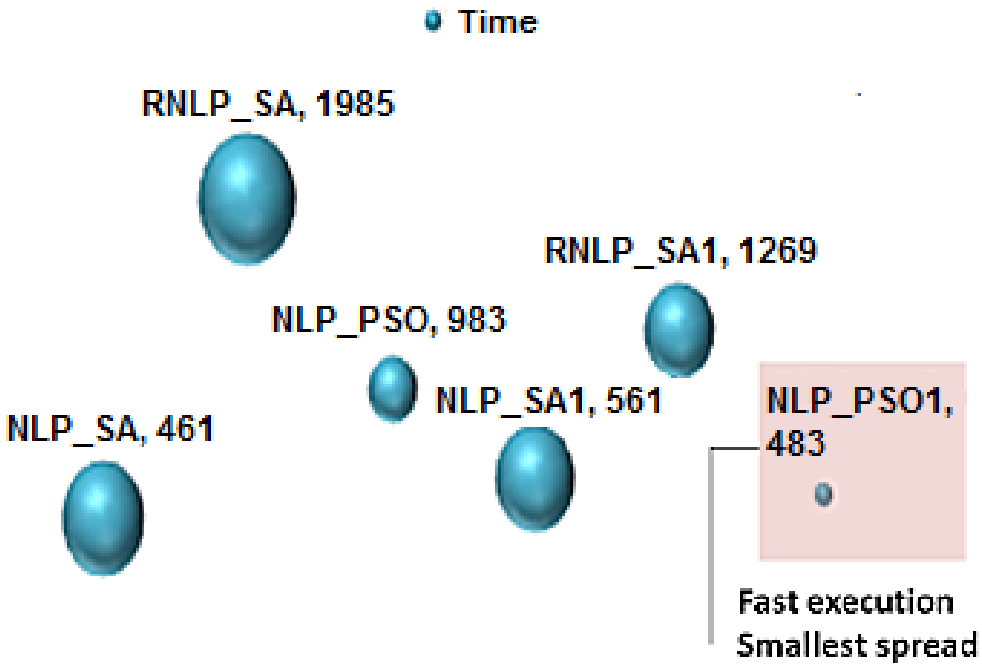


Figure 3.5 Mean elapse time (sec).

just 10% behind that of RNLP_SA1 (Figure 3.3). LP2 gives us a chance to gain a closer insight about the dominating crops, especially in case when decisions are to be made in a very short time period.

In a word, RNLP_SA1 is an excellent tool for solving the MCP problem, it achieves high accuracy with moderate execution time. However, according to the literature review increasing the number of variables the SA will converge rather slowly in order to provide sufficient moves carried out in every variable direction (Onbaşoglu and Ozdamar, 2001) . So NLP_PSO1 may provide an extra tool, but with a trade off between accuracy and runtime.

3.2.6 Real Case Study: Bekaa Valley

In order to analyze the expected profitability of cultivation of various crops (potato, maize, silage and tomato), the proposed linear LP2 and relaxed RNLP formulations of the MCP problem are applied in Bekaa Valley agricultural project (total area: 2000 *ha*) during the spring-summer production cycle. This specific time is the driest period of the year, it is irrigated by diverting water from QD through a network of canals. In fact, Qaraoun reservoir is the only source of water for the agricultural project. Furthermore, the model's decision interval for irrigation was set to one month period. Table 3.3 represents water availability at each month starting May till September.

Water requirements for different crops are given in Table 3.4, agricultural costs are summarized in Table 3.5, whereas the potential yield, selling price and production bounds of potato, maize, silage and tomato are presented in Table 3.6.

Table 3.3 Water availability in *MCM* during months: May, June, July, August and September.

Month	May	June	July	August	September	Total
Water availability (r_j) [*MCM*]	1.12	1.76	2.12	1.84	1.52	8.36

Table 3.4 Monthly water requirements for each crop in MCM/ha.

Crop	Monthly water requirements MCM/ha				
	May	June	July	August	September
Potato	0.37	1.87	2.84	2.76	0.99
Maize	0	1.18	2.46	3.05	0.69
Silage	0	1.18	2.46	2.80	0.41
Tomato	0.37	1.87	2.84	2.76	0.99

Table 3.5 Operation, irrigation and other costs used in the model.

Miscellaneous Items	Cost	Unit
Operation cost		
Machinery	100	€/*ha*
Labor	50	€/*ha*
seeds	100	€/*ha*
Fertilizers	300	€/*ha*
Harvest	100	€/*ha*
Irrigation costs		
Water	0.05	€/m^3

Table 3.6 Potential yield, selling price and production bounds.

Crop	Potential yield (ton/ha)	Price (€/*ton*)	Production bounds (ton)
Potato	15	900	5000
Maize	14	610	4000
Silage	24	800	12000
Tomato	28	715	5000

Table 3.7 Simulation results.

Scenario	Problem	Algorithm Used	Area in Hectares of each crop i				Irrigated Area (*ha*)	Consumed Water (*MCM*)	Profit (10^6 €)
			Potato	Maize	Silage	Tomato			
1	LP2	Interior-Point	0	0	480.14	178.57	658.71	4.871	12.122
2	RNLP	Simulated Annealing	94.00	0.00	586.86	194.50	875.36	5.815	13.073

Using the data provided by Tables 3.3, 3.5 and 3.6, the optimum crop pattern obtained by LP2 and RNLP models along with the consumed water amount and profit are presented in Table 3.7. As expected, a better solution was provided by the relaxed formulation of the MCP problem with a net profit gained equal to 13.073 million euros. The RNLP model-based on deficit irrigation scheduling has also determined the optimal cropped area which equal to 875.36 *ha* . Nevertheless, implementing the same strategy in case of full irrigation with limited water resources, the optimal net profit has decreased to 12.122 million euros and the crop spread area has reduced to 658.71 *ha*. Using the RNLP model, it can be noticed that farmers profit rose by at least 7.85% and the cultivated area increased by 10.85% when compared with the LP2 model. On the other hand, LP2 model managed, also in the real case, to detect the dominating crops (Silage and Tomato) where they re-appear in the outcome of the RNLP model. Furthermore, the consumed amount of water is 5.815 *MCM*, while the available amount is 8.356 *MCM*. The reader might wonder why the model did not utilize the whole available amount of water and therefore cropping a wider area. In fact, the answer is provided in Tables 3.3, 3.8.

Table 3.8 The used amount of water under deficit irrigation (RNLP model).

Month	May	June	July	August	September	Total
Used water [*MCM*]	0.107	1.226	2.121	1.834	0.527	5.815

Upon comparing both tables, one may notice that only 9.56 % and 34.7 % of the accessible water was used for irrigation during the months May and September respectively. This can be interpreted by that fact that during May the crops are in the establishment stage, while in September they are in ripening stage, and in both stages, the water consumption by the plants is at a low level. This leads the reader to ask another crucial question “Why LRA diverts water for irrigation to Bekaa Valley without being completely utilized?”. This is due to data scarcity. There are no complete records about the types of crops (or trees) being irrigated, precise cropping area or the practiced irrigation method (full or deficit). The work here is carried only on a sample of crops (potato, maize, silage and tomato) while the Bekaa Valley is planted with a wide variety of crops (or trees) using the same irrigation profile. Perhaps, in this situation, the diverted water is fully utilized. In all cases, the provided irrigation profile is not optimal since it is tightly linked to crop types, cropping area and irrigation method. To resolve this problem, an adjusted formulation of the MCP model (AMCP) is introduced. It has the ability to rearrange irrigation water releases R_j starting with an initial water volume V according to the planted crops.

The AMCP problem satisfies the constraints 3.9, 3.10, 3.12, 3.13 and in addition to the constraints that are given in 3.23, 3.24.

$$R_j = \sum_{i=1}^{n} X_i \cdot WA_{ij} \tag{3.23}$$

$$\sum_{i=1}^{l} R_j \leq V \tag{3.24}$$

where R_j is the released water at stage j and V is the assigned water volume for irrigation. Therefore, the AMCP problem becomes:

$$(AMCP) \begin{cases} \max F_R \\ 0 \leq WA_{ij} \leq WR_{ij}, \forall i = 1,...,n \text{ and } \forall j = 1,...,l \\ Ya_i = Ym_i \prod_{j=1}^{l} \left(\frac{WA_{ij}}{WR_{ij}}\right)^{\lambda_{ij}} \\ R_j = \sum_{i=1}^{n} X_i \cdot WA_{ij} \\ \sum_{i=1}^{l} R_j \leq V \\ \sum_{i=1}^{n} X_i \leq A_{total} \\ X_i \geq 0, \ \forall i = 1,...,n \end{cases}$$

Using $V = 8.36$ and the data given in Tables 3.4, 3.5 and 3.6, the obtained solution of the AMCP is as follows: 1- the acquired profit is 17.112 million euros; 2- the cropped area is 1068.25 *ha*; 3- the assigned amount of water is now fully utilized according to Table 3.9.

Table 3.9 The used amount of water under deficit irrigation (AMCP model).

Month	May	June	July	August	September	Total
Water released (R_j) $[MCM]$	0.187	1.611	2.822	2.984	0.756	8.36

According to the obtained results, the AMCP model managed to increase the profit by 31% and the cropped area by 22% by simply rearrange water releases. Therefore, as a part of the hydropower-irrigation management plan, the optimal irrigation profile generated by AMCP model can be used instead of the profile that is based on the operator's experience. By this method, the risk that reservoir's water release for irrigation is not fully utilized is reduced.

In general, the study aims, using real data, to verify what had been previously proved and discussed at the beginning of the section. Overall, the results came as expected. It encouraged us to perform further on site investigations to gather more precise information as a part of a future work. Integrating the collected information into the AMCP model will enable planners to set accurate and exact agricultural schemes.

3.3 Profit Distribution Model

In the previous section, a Multi-Crop Planning (MCP) optimization model was introduced for cropping pattern and water allocation as a nonlinear programming problem. Its solution had promoted an efficient use of water with a flexibility to keep the chosen crops under either full or deficit irrigation throughout different stages. Based on the provided numerical example, best results were obtained by solving the relaxed formulation of the MCP model using SA algorithm. Afterward, the relaxed formulation of the MCP problem was resolved using real data obtained from the Bekaa Valley.

However, in this part, the main interest is in the method of distributing financial profit among active farmers in that Valley on the local level. According to Staatz, agricultural cooperatives strive to maximize the benefits they generate for their members (which usually involves zero-profit operation) and the profit is distributed amongst farmer members as dividend payout which is a fixed amount of cash proportionate to the farmer's share (Staatz, 1987). In fact, the Profit Distribution (PD) model adopts the cooperative measure to identify the cropping pattern at the farmland level. The basic idea of this PD problem is a re-initialization of the mixed integer linear programming model presented in the work of Pap (Pap, 2008) but by putting it in a better context and bearing in mind two main things: 1- an efficient crop distribution over farmlands; 2- crop rotation. Here, one may claim that the crop rotation mode can be established in the MCP problem. In fact, this procedure will increase the complexity of the problem and will make it harder to solve. Therefore, it was more convenient to introduce the crop rotation constraint in the less complex PD problem.

Here, in this section, the suitable crop allocation among parcels is determined not by maximizing profit (Pap, 2008), but by using the optimal profit obtained by the MCP model and distributing it according to a defined cooperative profit policy through minimizing the sum of absolute profit deviations within farmers' parcels.

3.3.1 Models Description

PD model is motivated by agriculture cooperative approach in profit distribution. In a sense that, cooperatives provide a method for farmers to join together in an association, through which a group of farmers can acquire a better income (typically financial) than by going alone. This approach can be related to a form of economic synergy, where "two or more agents work together to produce a result not obtainable by any of the agents independently" (Staatz, 1987), whereas the profit reimbursement (either through the dividend payout or rebate) is shared only amongst the farmer members. In this research, the maximum profit obtained from the MCP problem is distributed among farmers according to each one share without actual existence of an agriculture cooperative. The objective is to minimize the sum of profit absolute deviation in the course of distributing crops over farmlands or parcels.

Objective Function

It is known that economic benefit is a strong driver in forming a multi-cropping system, while managing fields as zones helps to reduce input costs (Moore and Wolcott, 2000). Here, the distribution of crops is achieved on farmer's activity area level by optimizing a suitable objective function, which is described in the following.

Certain agriculture cooperatives distribute profit to farmers like a corporation that distributes profit through dividends to shareholders. In this research, the share will be proportional to the activity area of the farmer and based on that, the cut of farmer k is given by:

$$R_k = \frac{A_k}{A_{total}} \max F_R \tag{3.25}$$

Recall here, $A_{total} = A_1 + A_2 + ... + A_m$, where A_k is the area of parcel k (Figure 3.6) and F_R is the objective function of the relaxed formulation of the MCP problem given in Equation 3.15.

Furthermore, on implementation, this requires to plant all the n crops in every parcel k proportional to the area A_k, which is very hard task to fulfill due to the physical nature of the farmland (field capacity, structure,...etc.) and to the agricultural management practices (fertilizer distribution, harvesting,...etc.). To solve this issue, a new variable Q_k is introduced, such that for all $k = 1, ..., m$:

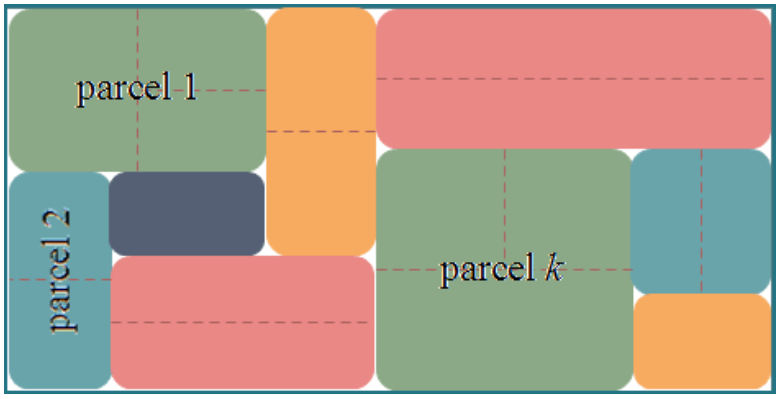


Figure 3.6 The studied agricultural region consists of several farmlands or parcels.

$$Q_k = \sum_{i=1}^{n} a_i \cdot \alpha_{ik} \cdot A_k \cdot y_{ik} \tag{3.26}$$

In fact, Q_k represents the financial benefits acquired when i various crops (not necessarily all the n crops) are planted in the parcel k (Figure 3.6) according to a proportion ratio equal to α_{ik}. Furthermore, y_{ik} is a binary variable which is defined as: $y_{ik} = \begin{cases} 1 & \text{if crop } i \text{ is assigned to parcel } k \\ 0 & \text{otherwise} \end{cases}$

It is noted here, for all $i = 1,...,n$, a_i can be calculated according to the following steps:

$$(X^*, WA^*) = \underset{(X,WA)\in S_R}{\arg\max}\ F_R \tag{3.27}$$

$$P_i^* = p_i \min(Ya_i X_i^*, c_i) - (pw \sum_{j=1}^{l} WA_{ij}^* + C_i) \cdot X_i^*$$

For $i = 1,...,n$,

$$a_i = \begin{cases} 0 & \text{if} \quad X_i^* = 0 \\ \frac{P_i^*}{X_i^*} & \text{if} \quad X_i^* \neq 0 \end{cases}$$

At this point, the aim is to find α_{ik} and y_{ik} for all $i = 1,...,n$ and $k = 1,...,m$, such that the objective function defined as:

$$F' = \sum_{k=1}^{m} |Q_k - R_k| \tag{3.28}$$

is minimum within a set of different constraints that are given below.

Constraints

The objective function is assumed to be bounded by a set of constraints described in the following:

1. The ratio of cropped area of parcel k can't exceed one:

$$\sum_{i=1}^{n} \alpha_{ik} \cdot y_{ik} \leq 1,\ \forall\ k = 1,...,m \tag{3.29}$$

2. The sum of all acreage planted with crop i should be equal to X_i^*:

$$\sum_{k=1}^{m} \alpha_{ik} \cdot y_{ik} \cdot A_k = X_i^*,\ \forall\ i = 1,...,n \tag{3.30}$$

3. As mentioned before, farmlands physical nature and agricultural practices had to be truly considered, so that the multi-cropping system does not invoke a disadvantage, especially with small farms. Therefore, parcels are categorized according to the maximum number of crops that they can accommodate:

$$\text{parcel } k \text{ can accommodate } s_k \text{ crops i.e.} \sum_{i=1}^{n} y_{ik} \leq s_k\ ;\ s_k \in \{1,..,n\} \tag{3.31}$$

4. It is known that agricultural productivity is sustained through crop rotation. Henkel claims that growing the same crop in the same place for many years in a row disproportionately depletes the

soil of certain nutrients (Henkel, 2015). With rotation, a crop that leaches the soil of one kind of nutrient is followed during the next growing season by a dissimilar crop that returns that nutrient to the soil or draws a different ratio of nutrients.

$$\alpha_{ik} \cdot y_{ik} \leq 1 - \beta_{ik} \tag{3.32}$$

Mixed Integer Non-linear Programming Problem

The PD model is expressed as a Mixed Integer Non-Linear Programming (MINLP) problem with two decision variables α_{ik} and y_{ik}, and it is represented in following:

$$(\text{MINLP}_{\beta}) \begin{cases} \min F' \\ \text{subjected to:} \\ \sum_{i=1}^{n} \alpha_{ik} y_{ik} \leq 1, \ \forall\, k = 1, ..., m \\ \sum_{k=1}^{m} \alpha_{ik} y_{ik} A_k = X_i, \ \forall\, i = 1, ..., n \\ \sum_{i=1}^{n} y_{ik} \leq s_k, \ s_k = 1, .., n \\ \alpha_{ik} y_{ik} \leq 1 - \beta_{ik}, \forall\, i = 1, ..., n; \forall\, k = 1, ..., m \\ 0 \leq \alpha_{ik} \leq 1, \forall\, i = 1, ..., n; \forall\, k = 1, ..., m \\ y_{ik} \in \{0, 1\}, \forall\, i = 1, ..., n; \forall\, k = 1, ..., m \end{cases}$$

The next subsection presents a solution technique based on two steps linearization approach in order to solve the MINLP problem.

3.3.2 Solution Technique for the PD Model

The PD model, as formulated, is a Mixed Integer Non-Linear Programming (MINLP) optimization problem. The goal is to convert it into a Mixed Integer Linear Programming (MILP) that can be easily solved using common optimization packages such as LINGO, CPLEX or MATLAB.
As a first step, choose a new variable Y_{ik} such that for all $i = 1, ..., n$ and for all $k = 1, ..., m$:

$$Y_{ik} = \alpha_{ik} y_{ik} = \begin{cases} \alpha_{ik} & \text{if } y_{ik} = 1 \\ 0 & \text{if } y_{ik} = 0 \end{cases}$$

It is obvious that for all $i = 1, ..., n$ and for all $k = 1, ..., m$:

$$Y_{ik} = \alpha_{ik} y_{ik} \Leftrightarrow \begin{cases} 0 \leq Y_{ik} \leq y_{ik} \\ \alpha_{ik} + y_{ik} - 1 \leq Y_{ik} \leq \alpha_{ik} \end{cases}$$

then for $k = 1, ..., m,\ Q_k$ becomes:

$$Q_k = \sum_{i=1}^{n} a_i \cdot Y_{ik} \cdot A_k \tag{3.33}$$

Therefore, constraints 3.29, 3.30 and 3.32 are now linear. However, in order to deal with the absolute value in the expression of F', consider the second change of variable as in Ferguson lecture notes (Ferguson, 2010):

$$Z_k \geq |Q_k - R_k|\,,\ \forall\, k$$

and then two new constraints are introduced:

$$Z_k \geq Q_k - R_k,\ \forall\, k = 1, ..., m$$

$$Z_k \geq R_k - Q_k,\ \forall\, k = 1, ..., m$$

The objective function of the PD problem is reduced to the following expression:

$$F' = \sum_{k=1}^{m} Z_k$$

Now, the MINLP problem has been linearized into MILP with a decision variable $w = (y, \alpha)$ and of the form:

$$(\text{MILP}_\beta) \begin{cases} \min\ F' = \sum\limits_{k=1}^{m} Z_k \\ \text{subjected to:} \\ Y_{ik} - y_{ik} \leq 0,\ \forall\, i = 1, ..., n;\ \forall\, k = 1, ..., m \\ -Y_{ik} + \alpha_{ik} + y_{ik} \leq 1,\ \forall\, i = 1, ..., n;\ \forall\, k = 1, ..., m \\ Y_{ik} - \alpha_{ik} \leq 0,\ \forall\, i = 1, ..., n;\ \forall\, k = 1, ..., m \\ \sum\limits_{i=1}^{n} Y_{ik} \leq 1,\ \forall\, k = 1, ..., m \\ \sum\limits_{k=1}^{m} Y_{ik} A_k = X_i,\ \forall\, i = 1, ..., n \\ \sum\limits_{i=1}^{n} y_{ik} \leq s_k \\ Y_{ik} \leq 1 - \beta_{ik},\ \forall\, i = 1, ..., n;\ \forall\, k = 1, ..., m \\ \sum\limits_{i=1}^{n} a_i \cdot Y_{ik} \cdot A_k - Z_k \leq R_k,\ \forall\, k = 1, ..., m \\ -\sum\limits_{i=1}^{n} a_i \cdot Y_{ik} \cdot A_k - Z_k \leq -R_k,\ \forall\, k = 1, ..., m \\ Y_{ik} \geq 0,\ \forall\, i = 1, ..., n;\ \forall\, k = 1, ..., m \\ Z_k \geq 0,\ \forall\, k = 1, ..., m \\ 0 \leq \alpha_{ik} \leq 1,\ \forall\, i = 1, ..., n;\ \forall\, k = 1, ..., m \\ y_{ik} \in \{0, 1\},\ \forall\, i = 1, ..., n;\ \forall\, k = 1, ..., m \end{cases}$$

3.3.3 Results and Discussions

Here, the mathematical model for PD is solved using the integer linear programming package from LINGO 16.0. The coded model is ran on Intel Core i7-5500U CPU @ 2.40 GHZ, 12.0 GB RAM.

In the absence of detailed information at farm level of the Bekaa Valley and without loss of generality, the suggested PD problem approach is illustrated over seven parcels of the respective areas (*ha*): 500, 200, 100, 1000, 100, 90 and 10. It is useful to recall that the total area of parcels is 2000 *ha*, where the planted area should be 875.36 *ha* (solution of RNLP problem). The goal to distribute the 875.36 *ha* with the suggested crop pattern over the parcels proportional to farmer's activity area. However, before the execution of the MILP model, the maximum number of crops that any parcel can accommodate is set to two. Then, the MILP is ran for two seasons. In Season One, crop rotation is disregarded ($\beta_{ik} = 0$, constraint 3.32 was removed), while in Season Two β_{ik} is set to equal to α_{ik} that was determined by the model in the first season. During model's execution, the MILP formulation of the PD problem has in Season One 28 binary variables, 64 continuous variables and 209 constraints while in Season Two, the number of constraints increased to 237. In both seasons, the optimal solution was reached in less than 1 second.

Tables 3.10, 3.11 show the simulated crop patterns for parcels 1 till 7 over 2 years based on the cropping profile suggested by RNLP model in the real case study (Subsection 3.2.6).

Table 3.10 Simulation results - Season One.

Farmer's parcel	Activity Area	Percentage (alpha)				Planted Area (*ha*)				Profit millions €
		Potato	Maize	Silage	Tomato	Potato	Maize	Silage	Tomato	
1	500	0%	0%	42.43%	0%	0	0	212.15	0	3.268
2	200	0%	0%	0%	37.79%	0	0	0	75	1.307
3	100	92.07%	0%	0%	0%	92.07	0	0	0	0.654
4	1000	0%	0%	37.47%	4.41%	0	0	375	44.1	6.537
5	100	1.93%	0%	0%	37%	2	0	0	37	0.654
6	90	0%	0%	0%	37.79%	0	0	0	34.01	0.588
7	10	0%	0%	0%	37.79%	0	0	0	3.78	0.065

Table 3.11 Simulation results - Season Two.

Farmer's parcel	Activity Area	Percentage (alpha)				Planted Area (*ha*)				Profit millions €
		Potato	Maize	Silage	Tomato	Potato	Maize	Silage	Tomato	
1	500	0%	0%	42.43%	0%	0	0	212.14	0	3.268
2	200	47%	0%	20.77%	0%	94	0	41.54	0	1.307
3	100	0%	0%	0%	37.79%	0	0	0	37.79	0.654
4	1000	0%	0%	29.08%	11.89%	0	0	290.75	118.92	6.537
5	100	0%	0%	42.43%	0%	0	0	42.43	0	0.654
6	90	0%	0%	0%	37.79%	0	0	0	34.01	0.584
7	10	0%	0%	0%	37.79%	0	0	0	3.78	0.065

In the presented simulations, the best agriculture practice is achieved with a sum of the absolute profit deviation equal to zero. In this case, the cropping scheme represents the optimal pattern based on the defined cooperative policy. The tables show the percentage of each crop and the corresponding planted areas in each parcel for two consecutive years. The importance of the model being applied on two successive seasons is to explore how MILP model respond to the crop rotation constraint. It is obvious that the model managed to re-arrange crops among parcels without affecting the profit distribution among farmers in both seasons.
In that context, several observations are made:

1. In Season One, parcel 1 is planted with 42.43% silage of its total area, which was planted again in Season Two with the same percentage. Nevertheless, the solution is still feasible since the cropped part, in both seasons, is less than 50%. The situation is the same for parcels 6 and 7.

2. The case is not the same for the parcels 2, 3, 4 and 5, a complete new crop arrangement (Figure 3.7) is made by the PD model.

3. The MILP model succeeded to distribute profit among farmers. For example, 50%, 25% and 0.5 % of the global profit area possessed by the owners of parcel 4, 1 and 7 respectively. In fact, these outcomes are expected since parcels 4, 1 and 7 represent respectively 50%, 25% and 0.5% of the entire area. It is clear that the largest parcel achieved the highest profit. Furthermore, partitions of the same area achieve the same revenue, as in parcels 3 and 5.

In general, the results obtained from the PD model suggest that the proposed approach is effective in finding the optimal solution for both crop distribution over parcels and rotation scheduling for multiple farmers and crops.

Despite the obtained good results, some people may claim that the farmer who possesses the majority of the farmlands will receive the highest income. This allows landlords to accumulate capital at an accelerated speed. Consequently, in case of large variation among parcels' areas, a profit gap will emerge between small farmers and major farmers. However, in the future work, this issue will be resolved by inducing a weight factor ω_k to Equation 3.25 according to the following formula:

$$R_k = \omega_k \max F_R \tag{3.34}$$

such that $\sum_{k=1}^{m} \omega_k = 1$.

Here ω_k does not only depends on the farmers activity area, but it depends, for example, on other factors that offer opportunities to strengthen family farming, beginner farmers, young people involved in agricultural youth organizations and farmers who need financial assistance. This measure will reduce the profit gap between smallholders and landlords. As a result, smallholders can sustain their small agricultural businesses.

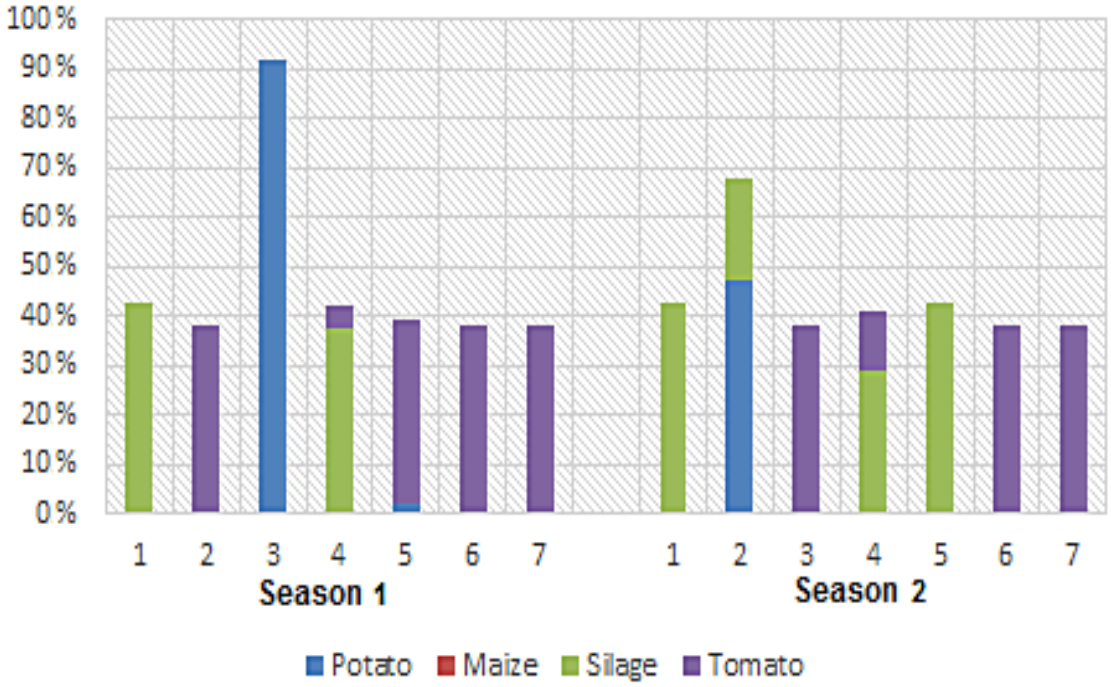


Figure 3.7 Crops distribution over the seven parcels (Seasons One and Two).

3.4 Conclusion

In this chapter, two-stage mathematical programming models for optimal multi-crop planning and profit distribution are presented. The aim is to assist farmers to make best decisions over a given planning horizon. Specifically, these decisions produce an optimal cropping plan which maximizes farmer' profit on both global and local level while satisfying different agricultural constraints.

Section 3.2 presents some mathematical tools and a numerical approach for resolving the Multi-Crop Planning (MCP) problem. Firstly, a NLP model is established to describe the MCP problem. Afterward, two linear formulations and a relaxed version are extracted from the NLP model. Based on the provided numerical example, results obtained by Simulated Annealing and Particle Swarm Optimization algorithms that were initiated near the solution of a specific linear problem revealed a significant decrease in algorithm's execution time and an increase in the cropped area and the total farming financial income under deficit irrigation. The computational results lead us to consider a real case situation to study the real capabilities of the suggested approach. It was illustrated through its implementation using real data obtained from the Bekaa-Valley region near Qaraoun reservoir whereas the availability of water at each stage is tightly linked to hydropower production. However, the achieved irrigation schedule did not utilize all the available amount of water. To overcome this issue, an adjusted version (AMCP) of the MCP model was introduced. The AMCP model managed to boost revenues by re-arranging irrigation profile. In fact, the AMCP model can play an important

role in the next chapter. It aims to provide the medium term hydropower scheduling model with the optimal irrigation profile.

Overall, the results of the real case came as expected during the experimental tests. It encouraged us to integrate the obtained outcomes into the Profit Distribution (PD) model to effectively distribute revenues, on the local level, among active farmers.

Section 3.3 exhibits a method based on a defined cooperative policy to distribute the global profit achieved by MCP model between farmers. The capabilities of the suggested PD model were illustrated through its implementation on Bekaa Valley over two consecutive seasons. The MILP formulation of the PD problem managed to maintain the same profit distribution scheme in the presence/absence crop rotation policy. This technique is effective to deal with the complexity of managing profit among several land owners involved in the same agricultural project. One of the potential drawbacks is manifested in the countries with un-free economy whereas this approach grants the highest profit to the farmer with the largest activity area. However, as a part of the future work, the resolution of this issue could be achieved by replacing the activity area to total area ratio with a weight factor. This weight factor is proportional not only to the activity area, but it is related to financial situation of the farmer. In this case, beginner farmers can improve their holdings and sustain their agricultural businesses.

Chapter 4

Optimal Operation of Cascade Hydropower-Irrigation Plants

> Energy cannot be created or destroyed, it can only be changed from one form to another
>
> *Albert Einstein*

4.1 Introduction

Nowadays, hydropower is the leading renewable source for electricity generation, it supplies, according to World Energy Council, around 71% of all renewable electricity. In the year 2016, it reached 1,064 *GW* of the installed capacity (675 *GW* in year 2000), whereas it generated 16.4% of the world's electricity from all sources (World Energy Council, 2016). The drivers for the upsurge in hydropower development are the increased demand for electric power, energy storage, flexibility of generation, hydroelectricity's low cost and freshwater management. The awareness of all these factors still motivates engineers to present new implementations to optimize hydropower operation. Despite being a mature technology, hydropower has still a significant potential mostly in cascade systems. It is common to notice several hydropower plants constructed on the same river and sharing a common water lake (Yildiran et al., 2015). Therefore, on the same river, new plants can be sub-joined and old ones can be upgraded in order to increase electricity and agricultural productions. These improvements are achieved only through intelligent operations and coordination. Otherwise, spillages may occur and water may be wasted without being utilized neither for electric generation nor for irrigation.

This chapter suggests detailed models in order to enhance medium and short term hydropower-irrigation operation. To achieve the desired goal, it is important to consider the following key components:

- Inputs: load demands; hydrological estimations (discussed extensively in Chapter Two); irrigation requirements (discussed in Chapter Three); initial reservoir storage; electricity prices; maintenance costs.
- Constraints: electrical (plants' power capacity, power transition,...etc.); hydraulic (pipelines, spillage system, storage,...etc.) safety (maintenance, water hammer effect, storage,...etc.); irrigation agreements; environmental.
- Objectives: power balancing, satisfying irrigation demands, minimizing financial losses of the hydropower operation, efficient load distribution, maintaining high turbine efficiency.

The medium-term scheduling model is typically used to determine the optimal strategy for hydropower-irrigation operation over one year (with time increment: one month). The aim is to efficiently manage water resources during hydropower production and irrigation. Moreover, the medium-term model, in case of similar topology description as the short-term model, plays an important support in supplying suitable boundary conditions. In fact, after solving, the role of the hydropower operator is to transform results from the medium-term scheduling process into a form suitable in the short-term scheduling process.

The short-term model is implemented within a time horizon of a day to a week and uses hourly time increments. It supports a more detailed system model than in the medium-term model. Its main task is to suggest the daily generation schedule for hydropower-irrigation plants. It involves different sub-problems: unit commitment, economic load dispatch, efficiency curve, maintenance tasks,...etc.

In this chapter the objective is to suggest a new approach which provides the best possible planning and operating decisions in the medium and short run. In fact, in this chapter several topics (Figure 4.1) are treated that covers: 1- the financial impact of evaporation on hydropower and irrigation sectors; 2- electric load distribution based on efficiency curve; 3- Medium Term Hydro Generation-Irrigation Scheduling (MTHGIS) modeling; 4- Short Term Hydro Generation Scheduling (STHGS) modeling. Afterward, all the suggested models are tested in Litani Project - Lebanon.

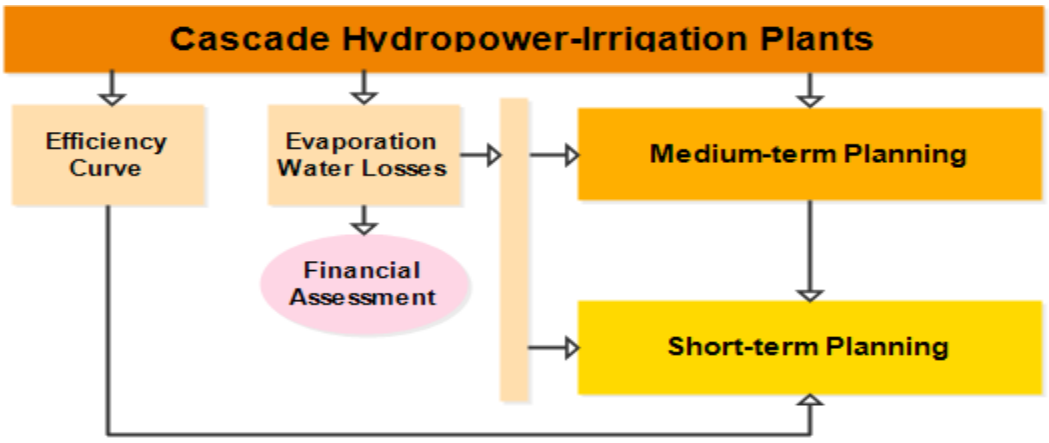


Figure 4.1 Framework of the modeling process.

List of Symbols

Notation	Description	Unit	
Indices		Medium term	Short term
t	Index of time periods,$t = 1,...,T$. T is the planning time horizon	$months$	$hours$
i	Index of the hydropower plant, $i = 1,...,m$. m is the total number of cascade plants	Unitless	Unitless
n_i	Number of generating units in plant i	Unitless	Unitless
j	Index of the generating unit	Unitless	Unitless
Decision variables			
$x_{ij}(t)$	$x_{ij}(t) = 0$ if unit j of plant i at time t is off and $x_{ij}(t) = 1$ otherwise	-	Unitless
$z_{ij}(t)$	$z_{ij}(t) = 0$ if unit j of plant i at time t is not under maintenance and $z_{ij}(t) = 1$ otherwise	-	Unitless
$q_{ij}(t)$	Water flow into turbine j of plant i at time t	-	m^3/s
pw_i	Power generated by any working unit at plant i	-	MW
$p_{ij}(t)$	Power produced by unit j of plant i at time t	-	MW
$P_i(t)$	Total power produced by plant i at time t	Gwh	MW
$R_i(t)$	Water released at plant i at time t	MCM	m^3
$Sp_i(t)$	Spillage at plant i at time t	MCM	m^3
$Ir_i(t)$	Irrigation release from reservoir i	MCM	m^3
$S_i(t)$	Water storage in reservoir i at time t	MCM	m^3
$A_i(t)$	Water surface of reservoir i at time t		m^2
$h_i(t)$	Water head in reservoir i at time t	m	m
Miscellaneous parameters			
$Evp_i(t)$	Surface evaporation at reservoir i at time t	MCM	m^3
$I_i(t)$	Water inflow into reservoir i at time t	MCM	m^3
$Ir_i^{\max}$	Maximum irrigation release from reservoir i	MCM	MCM
$MA_i(t)$	Water intake from reservoir i at time t for municipal use	MCM	m^3
$D(t)$	Power demand at time t	GWh	MW
$S_i^{\min}$	Minimum storage in reservoir i	MCM	MCM
$S_i^{\max}$	Maximum storage in reservoir i	MCM	MCM
$R_i^{\min}$	Minimum release from reservoir i	MCM	m^3
$R_i^{\max}$	Maximum release from reservoir i	MCM	m^3
$Sp_i^{\min}$	Minimum spill from reservoir i	MCM	m^3
$Sp_i^{\max}$	Maximum spill from reservoir i	MCM	m^3
ID_i	Irrigation demand over the scheduling period	MCM	MCM
$q_i^{\min}$	Minimum release through any turbine in plant i	-	m^3/s
$q_i^{\max}$	Maximum release through any turbine in plant i	-	m^3/s
$p_i^{\min}$	Minimum power output by any unit in plant i	-	MW
$p_i^{\max}$	Maximum power output by any unit in plant i	-	MW
τ_i	Time required for the discharged water from reservoir i to reach reservoir $i+1$	-	hr
c_i	Generation coefficient of any unit in plant i	-	$MW.s/m^3$
η_i	Efficiency of a given unit at plant i	Unitless	Unitless
T_{down}	Minimum down time		h
T_{up}	Minimum up time		h
$\Delta_{ij}(t)$	Maximum threshold of the degradation function	-	Unitless
RUL_{ij}	Remaining useful time for the unit j of plant i	-	h
W_i	Maintenance window	-	h
$pr(t)$	Selling price of one MWh	-	€
$mc_{ij}(t)$	Maintenance cost per hour of unit j of plant i at time t	-	€

Note: All generation units (turbines) in any given plant are assumed to have the same specifications.

4.2 Case Study: Actual Operation

The Litani Project is a typical example of a hydropower-irrigation cascade system that is poorly monitored and operated. Such situation is common in most developing countries. In fact, advanced monitoring and operating systems are rarely applied in these countries because of their high implementation cost and their need of constant technological follow-up. As a result, the available water in the watersheds is not fully exploited for irrigation, hydropower, fishery, recreation, ...etc. Thus, the sustainable growth of the country remains elusive. This work presents cheap methods for a proper water resource management. In this chapter, the presented work is not confined only to medium term and short term scheduling but also covers the economic impact of evaporation losses on hydropower and irrigation sectors, best units' efficiency and maintenance tasks within the hydropower plants.

In order to suggest any DST, it is important to understand the limitations of the current process so that any new operational plan can perform better during reservoir operations. The actual operation of the current decision practice in the Litani Project can be summarized in the following:

1. No studies concerning the impact of evaporation on the hydropower and irrigation sectors.
2. Operation scheduling is based on the professional experience of the operators. Thus, optimal scheduling is not guaranteed in the absence of automated procedures.
3. The hydropower planning does not encompass any river forecasting models. As a result, the process is subject to large errors given the substantial large variability in stream flow.
4. Risk of spillage or flood. Water is lost without being utilized for electricity generation or irrigation.
5. No clear framework to facilitate maintenance of the hydropower plants in order to increase generating units life time.
6. Risks concerning structural safety and recreation.

Figure 4.2 presents a brief description of the current operation in the Litani Project.

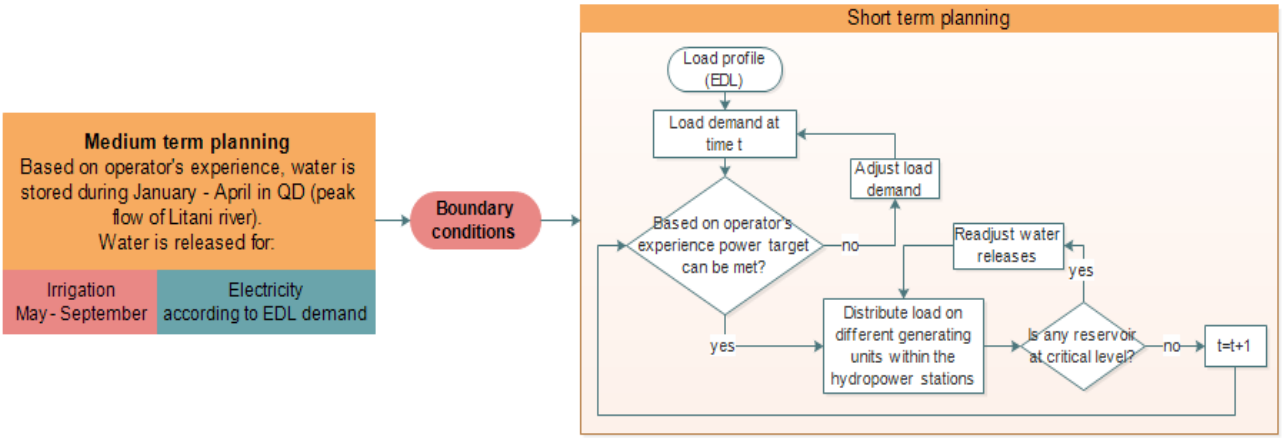


Figure 4.2 Bloc diagram reflecting the actual Litani Project operational procedure.

4.3 Multi-Reservoir System Description

4.3.1 General Profile

A multi-reservoir system consists of several hydropower stations interconnecting different lakes through a network of tunnels. However, a system is known as a cascade hydropower system (Figure 4.3) when two or more hydropower plants are constructed in series such that the runoff discharge of one hydropower plant is used as the inflow of the next hydropower plant and so on.

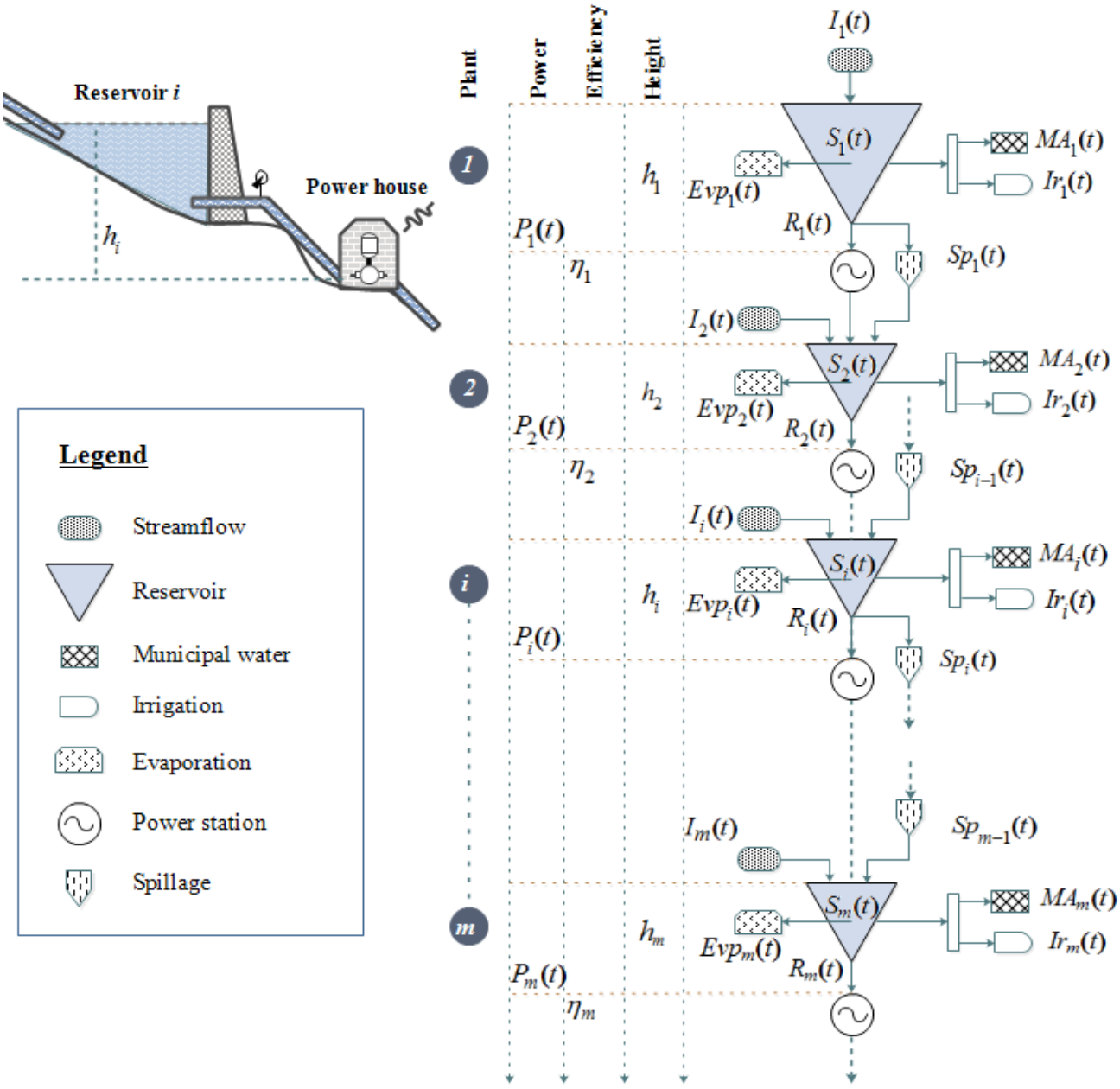


Figure 4.3 A typical cascade hydropower system.

In fact, Figure 4.3 represents an $m-$reservoir cascade hydropower system that consists of m reservoirs and m hydropower plants connected in series. The water balance equation of each reservoir i is affected at any time t by: river inflow $I_i(t)$, water losses due to evaporation $Evp_i(t)$, discharge from the preceding reservoir $R_{i-1}(t)$, spillage from the preceding reservoir $S_{i-1}(t)$ and water releases for miscellaneous use (spillage $S_i(t)$, municipal $MA_i(t)$, irrigation $Ir_i(t)$, power generation $R_i(t)$). The generated power $P_i(t)$ by plant i is proportional to head h_i, plant efficiency η_i, and the inflow $R_i(t)$.

4.3.2 Case Study Profile - Litani Project

All the suggested models in this chapter are applied in the Litani Project that was described in Chapter One. As a reminder, this project involves a main reservoir QD with a storage capacity of 220 *MCM* and two secondary smaller reservoirs known as Anan and Joun. The project diverts the Litani river flow through a system of tunnels inter-connecting three hydropower plants: Markaba, Awali and Charles Helou (Table 4.1). Further specifications associated with the Litani Project are addressed in Table 4.2 and Figure 4.4 (LRA, 2016).

Table 4.1 Data associated with each plant (Source: LRA records).

Plant	Turbine no/type	Installed capacity (MW)	Discharge capacity (m^3/s)	Water fixed head (m)	Plant efficiency
Markaba	2xFrancis	34	22	200	0.7
Awali	3xPelton	108	33	400	0.8
Charles-Helou	2xFrancis	48	30	200	0.8

Table 4.2 Case study notations - Litani project (Source: LRA records).

no	Name	Notations	Description
1	Litani river flow	$I_1(t)$	Main driver of the system
2	Qaraoun lake storage	$S_1(t)$	Main reservoir (dead storage 40 *MCM*, active storage 170 *MCM*, maximum capacity 220 *MCM*)
3	Canal 900 withdrawal	$Ir_1(t)$	Water is released from Qaraoun reservoir for Bekaa irrigation (8-10 *MCM*/yr)
4	Qaraoun spillage	$Sp_1(t)$	Qaraoun reservoir released through spillway
5	Qaraoun hydro-release	$R_1(t)$	Water is released into Markaba plant
6	Ain Zarqa spring	$I_2(t)$	Spring inflow to Markaba pond (not used for storage)
7	Qasimieh project	$Ir_2(t)$	Water is diverted for irrigation of Qasimieh irrigation project from the Markaba Reservoir (30-40 *MCM*/yr)
8	Markaba discharge	$R_2(t)$	Water outflow from Markaba pond
9	Jezzine Spring	$I_3(t)$	The ponds Markaba and Anan are connected by the Awali tunnel, where the Jezzine spring flow joins the running water
10	Azour overflow	$Sp_2(t)$	the Awali Tunnel is not meant to be pressurized or run full. When upstream capacities are greater than the capacity of the tunnel or of Anan reservoir, the excess flow is diverted to the Awali river, feeding Joun lake
11	Anan inflow	$I_4(t)$	The water inflow into Anan lake comes from Awali Tunnel
12	Lebaa Irrigation	$Ir_3(t)$	Water is released for irrigation of Lebaa from Anan reservoir (8-10 *MCM*/yr)
13	Anan spillage	$Sp_3(t)$	Anan reservoir released through spillway into Azour overflow
14	Anan hydro-release	$R_3(t)$	Water is released into Awali plant
15	Awali River	$I_5(t)$	Awali inflow into Joun lake
16	Joun spillage	$Sp_4(t)$	Water is spilled back into the river
17	Joun hydro-release	$R_4(t)$	Water is released into Charles Helou plant
18	Anan lake	$S_2(t)$	storage of Anan lake (maximum capacity 170, 000 m^3)
19	Joun lake	$S_3(t)$	Storage of Joun lake (maximum capacity 300, 000 m^3)
20	Lake evaporation	$Evp_1(t)$	Water evaporation from QD

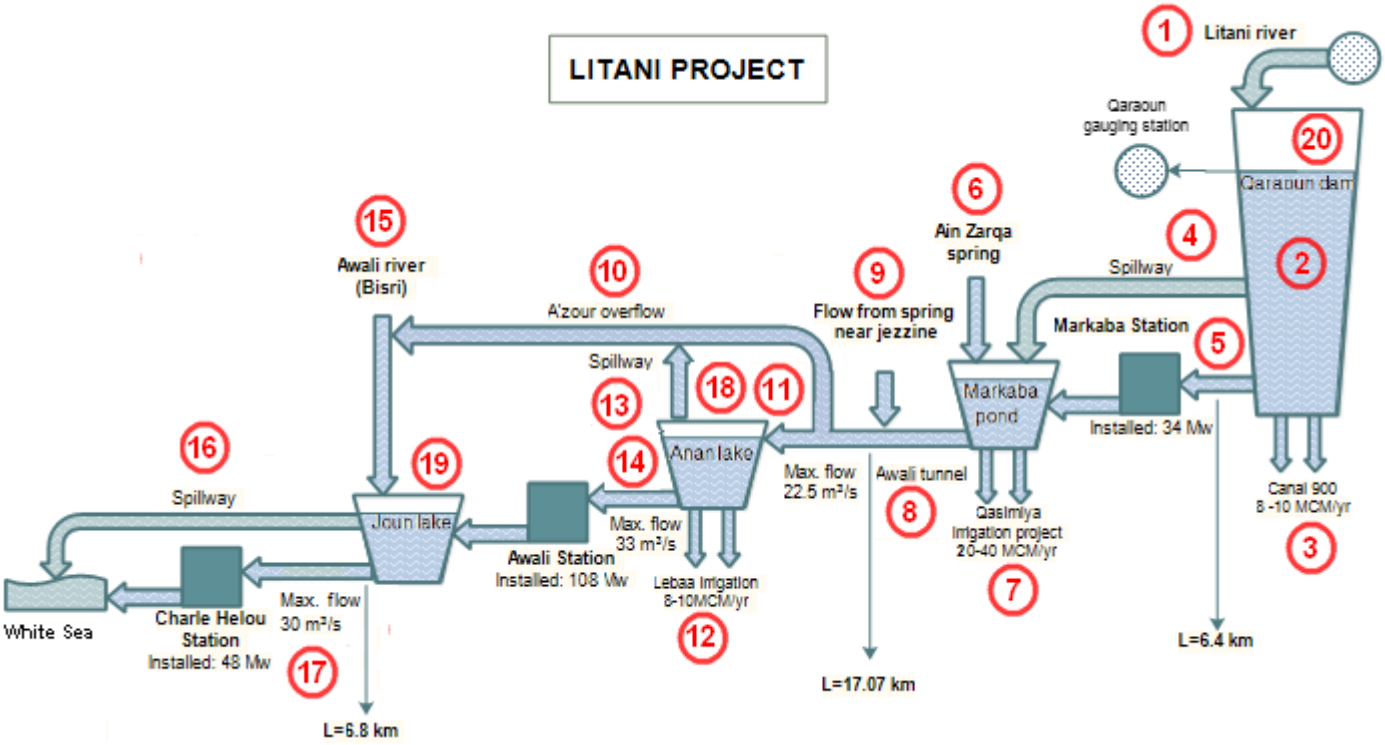


Figure 4.4 Litani Project configuration.

4.4 Financial Assessment of Evaporation on Hydropower and Irrigation in the Litani Project

In warm countries, storage of water behind dams leads to consumptive water loss through evaporation on the water surface. This section aims to assess the financial impact of water evaporation on the Litani Project cascade hydropower plant. Based on the achieved results, we are able to identify whether to include evaporation losses or not during the modeling process.

Increased evaporation reduces hydroelectric generation for all types of dams, but these effects will be most drastic for those with reservoirs with large surface areas. It is known that, due to the direct relationship between the surface area of the body of water and its rate of evaporation, the geometry of reservoirs determines how susceptible they are to evaporation (McJannet et al., 2008). Reservoirs with higher surface area to volume ratios are more vulnerable to losing water from evaporation, which reduces the facility's power production capacity. Retrofitting reservoirs to make them deeper with a smaller surface area would reduce evaporation, however it is very expensive (Mcjannet et al., 2008). The following presents a brief description of the impact of evaporation on the hydropower and irrigation sectors in the Litani Project.

4.4.1 Hydropower Sector

The aim of this part is to provide a method to estimate the hydroelectricity financial losses related to evaporation. In case of QD (Figure 2.23), the water spread area is around 11.9 km^2. However, this area varies throughout the year due to several reasons such as power generation, irrigation and evaporation. To get a reliable approximation of the monthly water losses due to evaporation, the formula exhibited in Equation 4.1 is used:

$$Evp_1(t) = E_0(t) \cdot \frac{A_1(t+1) + A_1(t)}{2} \tag{4.1}$$

where t is the time in months, $Evp_1(t)$ is the water loss from QD dam $[MCM]$, $E_0(t)$ is the evaporation rate $[m]$ and $A_1(t)$ is the water spread area of QD $[km^2]$. Here, E_0 is calculated using the methodology suggested in Chapter Two (Section 2.8) and the obtained results are presented in Figure 4.5.

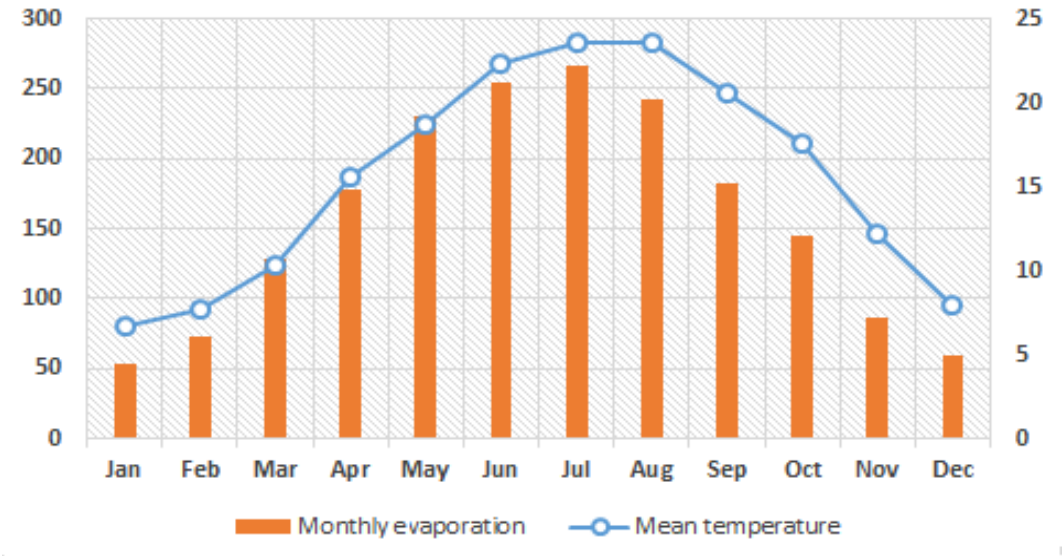


Figure 4.5 Monthly evaporation $[mm]$ and mean temperature $[^oC]$ at QD region.

Afterward, the lost power P $[MWh]$ due to evaporation can be easily calculated using the power formula (Loucks et al., 1981) given in Equation 4.2:

$$P_1(t) = 2725 \cdot \eta_1 \cdot Evp_1(t) \cdot h_1(t) \tag{4.2}$$

Table 4.3 shows the recorded surface areas and the average water heights at QD during the year 2013. Using these measurements, the suggested approach is successfully implemented whereas the results are exhibited in Table 4.4. It is important to note here that the exposed surfaces of the other lakes (Anan and Joun) are too small to provide an evaporation source, therefore they were not included in the financial assessment.

In Lebanon, "Electricité du Liban" (EDL) is the sole electricity buyer and seller. It pays €24 for each one MWh (EDL records). Therefore, if the water loss due to evaporation is used for hydropower

Table 4.3 Water spread area and average water elevation at QD during the year 2013.

Qaraoun Dam	Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Surface area [Km^2]	9	10.75	10.94	11.15	10.82	10.32	9.73	8.9	7.99	6.99	6.22	5.61
Avg. water head [m]	250.31	256.54	257.23	257.98	256.80	255	252.90	249.98	246.75	243.24	240.5	238.36

Table 4.4 Water and power losses due to evaporation.

Qaraoun Dam	Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Surface area [Km^2]	9.00	10.75	10.94	11.15	10.82	10.32	9.73	8.90	7.99	6.99	6.22	5.61
Water losses [MCM]	0.487	0.777	1.410	1.985	2.494	2.629	2.586	2.162	1.458	1.017	0.540	0.333
Avg. water elevation [m]	50.31	56.54	57.23	57.98	56.80	55.00	52.90	49.98	46.75	43.24	40.50	38.36
Energy losses [MWh]												
Markaba	232.47	380.17	691.83	976.62	1221.76	1278.99	1247.40	1031.08	686.21	472.04	247.62	151.35
Awali	424.56	677.44	1229.53	1730.57	2174.94	2292.86	2254.80	1885.57	1271.31	887.14	470.68	290.28
Charles Helou	205.91	328.56	596.32	839.33	1054.84	1112.04	1093.58	914.50	616.59	430.26	228.28	140.78

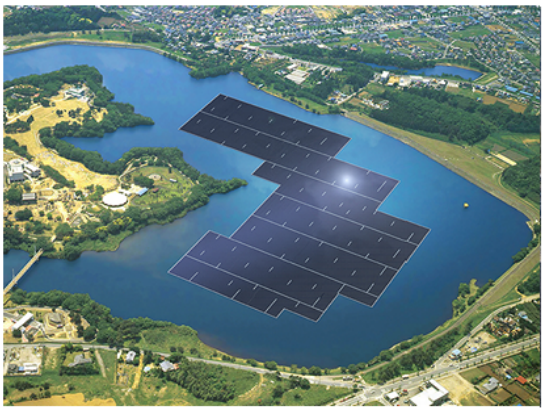
production (without irrigation consideration), the potential generated energy is approximately 31768 MWh. In this case, the revenue loss is around €762864.

On average, the annual generated power from Litani project is approximately 625 GWh. In this case, the produced income is 15 million euros. By simple calculation, it is estimated that the financial losses due to evaporation is 5.1% of the total achieved profit.

4.4.2 Irrigation Sector

The following represents a rough estimation of the agricultural economical losses due to evaporation in the Litani Project. It has been seen in Chapter Three, a complete discussion about deficit/non-deficit irrigation models with the objective of maximizing profit. In fact, under certain settings, the RNLP model managed by allocating 5.82 MCM of water to irrigate an agricultural zone of area 875.36 ha. In this case, the achieved profit was 13.07 million euros of agricultural products. Here, based on Table 4.4, the approximate total water loss due to evaporation is 17.88 MCM. This water volume is enough to irrigate three zones of the same size that was suggested by the RNLP model. Thus, using the same setting given in Section 3.2.6, the evaporated water is estimated to produce a financial loss around 39.22 million euros.

On both sectors, evaporation imposes not a small economical wastage. Thereby, water losses due to evaporation cannot be simply neglected and it will be considered in both schedule models, the medium-term as well as in the short-term. Overall, knowing the associated financial losses motivates engineers to submit feasible solutions (Figure 4.6: floating solar panels or shade balls) to increase hydropower-irrigation revenues by reducing the evaporation volume.



(a) Floating solar panels deployed on Yamakura Dam Tokyo, Japan (Science Alert, 2016).

(b) Hetch Hetchy Dam covered with floating black balls in California, USA (California WaterBlog, 2016).

Figure 4.6 Potential solutions to reduce the effect of reservoir surface evaporation.

4.5 Electric Load Distribution for Optimal Efficiency

This section introduces a method to improve the solution of unit commitment and economic load dispatch problems[1] by taking two important measures. First, a method is suggested to maintain, during the hydropower operation, a high turbine efficiency similar to the approach presented in the work of Lu (Lu et al., 2015). He claims that the efficiency curve of the turbine unit does not present the tendency to be an increasing function versus rated flow. In fact, turbine units retain their high efficiencies when running below the design flow. Figure 4.7 shows the nonlinear function that links efficiency with various rated discharges of different turbines.

For example, Francis turbine efficiency falls sharply as it is operated below half of its design flow, but Pelton and Kaplan turbines preserve their higher efficiency even at very low flow rate. Therefore, in order to maintain a near optimal turbine efficiency, releases need to be adjusted according to turbine type:

- Francis turbine

$$0.68q_i^{\max} \leq q_{ij}(t) \leq q_i^{\max} \tag{4.3}$$

- Pelton turbine

$$0.35q_i^{\max} \leq q_{ij}(t) \leq q_i^{\max} \tag{4.4}$$

- Kaplan turbine

$$0.35q_i^{\max} \leq q_{ij}(t) \leq q_i^{\max} \tag{4.5}$$

Here, it is assumed that all turbines in the same plants have similar specifications. Based on the new flow settings, the efficiency is now almost linear. For Francis turbine it ranges between 0.84-0.85

[1] Unit commitment and economic load dispatch are problems respectively concerned with generating units scheduling and load distribution within the hydropower plant in order to optimize operation economical losses.

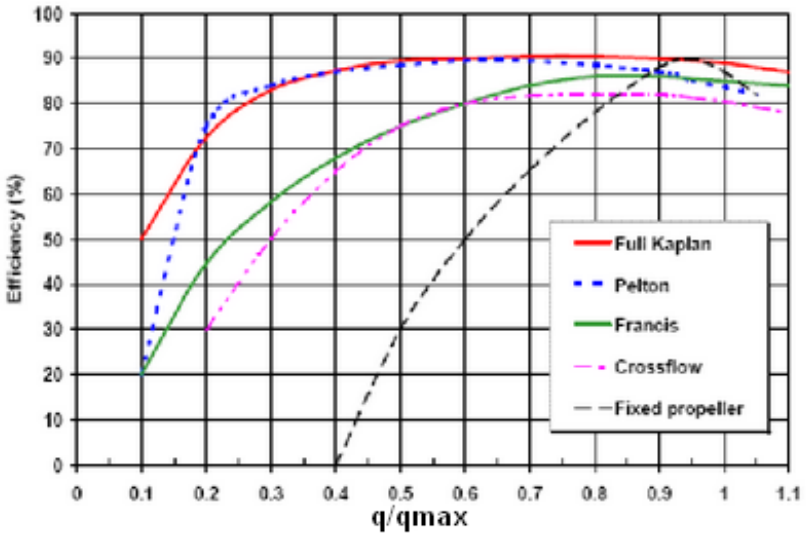


Figure 4.7 Efficiencies comparison of different turbines (Kaldellis et al., 2005).

(average: 0.845), while the efficiency of the Pelton and Kaplan turbine lies between 0.85-0.9 (average: 0.875). Thus, by imposing the new bounds on water releases, the improved solution of the unit commitment problem is manifested by keeping the unit running at high efficiency online, otherwise, the unit is offline.

Second, it is often encountered a situation where several turbine units are running simultaneously and each one is operating at different load. The next paragraph provides a mathematical explanation justifying the importance of distributing the load equally among running units in terms of efficiency.

As a remark, the function f that relates unit efficiency η with the output power p ($\eta = f(p)$) share a similar profile as the function that links efficiency with turbine flow concerning concavity (Kaldellis et al., 2005). In fact, this note is important for resolving the next problem.

In a certain hydropower plant, suppose N identical units all are running in the same time, whereas the target total power is P_T. The goal is to maximize the unit's efficiency η_k, $k = 1,..,N$; given that the sum of p_k's (p_k : the power produced by kth unit) is equal to P_T. The result is a multi-objective programming problem:

$$\begin{aligned} &\max(\eta_1, \eta_2, ..., \eta_N) \\ &\eta_k = f(p_k),\ k = 1,..,N \\ &\sum_{k=1}^{N} p_k = P_T \end{aligned} \tag{4.6}$$

All units have the same priority level, then by using the weighted sum method, the problem given in 4.6 becomes single objective:

$$\max \sum_{k=1}^{N} \frac{f(p_k)}{N} \qquad (4.7)$$
$$\sum_{k=1}^{N} p_k = P_T$$

The obtained form is a simple concave programming problem that can be solved by the method of Lagrange multipliers. The Lagrange function is defined as:

$$\mathscr{L}(p_1, p_2, ..., p_N, \lambda) = \sum_{k=1}^{N} \frac{f(p_k)}{N} + \lambda \left(\sum_{k=1}^{N} p_k - P_T \right)$$

Calculating the gradient of $\mathscr{L}$ and equating it to zero, $\nabla \mathscr{L} = 0$. We get the system of equations:

$$\begin{cases} \frac{f'(p_1)}{N} + \lambda = 0 \\ \frac{f'(p_2)}{N} + \lambda = 0 \\ \vdots \\ \frac{f'(p_k)}{N} + \lambda = 0 \\ \sum_{k=1}^{N} p_k - P_T = 0 \end{cases}$$

Therefore, $f'(p_1) = f'(p_2) = ... = f'(p_N)$ and $\sum_{k=1}^{N} p_k = P_T$. Since f' is a strictly decreasing function, $p_1 = p_2 = ... = p_N$ (unique solution). Thus, the solution of the problem given in 4.7 is $p_1 = p_2 = ... = p_N = \frac{P_T}{N}$. Thereby, in a situation where more than one generating units are running, it is recommended, for optimal efficiency, to divide the load equally between them. As a result, economic load dispatch is better realized by the equal load distribution.

4.6 Medium-term Scheduling of CHIP

Many studies dealt with medium-term scheduling of multi-purpose cascade reservoir system. Apparently, the operational planning involves interactions and trade-off between various objectives which may be occasionally complementary but often are competitive and conflicting. For instance, there are basically conflicts of interest among: 1- upstream-downstream municipal water share; 2- water supply to different sectors (hydropower, irrigation, recreation); 3- structural safety of different reservoirs; 4- environmental issues. Any mis-coordination between the cascade reservoirs, water may be wasted without being utilized efficiently. Thereby, for researchers, maximizing the benefit from the simultaneous operation of cascade hydropower plants imposes a great challenge.

In the work of Bai, the multi-objective operation of cascade reservoirs is transformed into single objective: maximizing power generation, whereas the other sub-goals are turned into constraints (Bai et al., 2015). However, Reddy considered a bi-objective problem: minimizing irrigation deficit while maximizing power generation (Reddy and Kumar, 2006). However, others maximize both reservoir storage and the difference between the generated power and pumping-back power (Yang et al., 2015). In Lebanon, the main objective of the Litani Project is to secure both a stable source of electric power to meet domestic load demand and a firm source of water for irrigation while meeting the system's physical and operational constraints. To achieve the desired goal, first a general nonlinear multi-objective Medium-Term Hydro Generation - Irrigation Scheduling (MTHGIS) model is presented to minimize power, irrigation and municipal water deficits. Afterward, the suggested model is implemented in Litani Project. The major driving force in making operation decision is to ensure the availability of sufficient amount of water to meet the power and irrigation demands during the planning horizon. Unfortunately, Litani river, the main driver of the system, follows a pluvial regime that tapers down during the dry season. Thus, there is an urge to preserve some of the water for that crucial time in order to protect consumers and farmers from electricity deficits and water shortage respectively. On the other hand, when water is in abundance, system operation should focus on making the best use of available resources and on protecting the system from the risk of unnecessary spillage.

Most of the solution approaches of multi-criteria optimal operation of the cascade system use posterior methods based on evolutionary algorithms (Reddy and Kumar, 2006; Afshara et al., 2007; Yang et al., 2015). In this work, we had switched to single-objective mathematical programming technique that is easy to implement: Weighted Sum (WS) Method. The main motivation, behind this choice, is two reasons: 1- the solution of the suggested mathematical programming problem is Pareto optimal; 2- the wide availability of single-objective optimizers like LINGO, BARON, GAMS and AIMMS. However, two drawbacks had faced us with WS method. First, a uniform distribution of the weights among objective functions does not always result in an even distribution of solutions on the Pareto front. Second, the approach cannot find solutions on non-convex parts of the Pareto front, although Pareto optimal solutions do often exist. Nevertheless, this issue can be resolved by considering the Adaptive Weighted Sum (AWS) method which will be our major concern in the future work. Another complicating factor in meeting the desired aim is the fact that the electricity demand and water inflows are both uncertain, as both primarily depends on weather conditions. However, in this section, average values are considered for power demand, irrigation requirements and inflows (normal conditions). In fact, the main concern in this research is: 1- providing the hydropower operator a general overview of the medium-term water management plan under normal conditions; 2- supplying flexible boundary conditions for short-term planning problem. From this perspective, the aim of the developed MTHGIS framework is to assist the operation engineers in improving HIP operation and to make optimal operational and trade-off decisions while meeting various hydraulic and electrical constraints.

4.6.1 Mathematical Formulation of the MTHGIS Model

Medium-term planning of water resource projects is usually scheduled on monthly basis. The time step in this section is one month and the scheduling horizon is set to one year ($t = 1, ..., 12$).

Multi-Objective Function

Water demand is classified according to the usefulness of its application. In this part, the main interest is the water allocation to the three different sectors: agriculture, hydropower and municipal water. The aim is to minimize the multi-objective functional presented in Equation 4.8 by optimally allocating water to each of the three mentioned sectors.

$$\min \left(f_{irr}, f_{hydro}, f_{muni}\right) \tag{4.8}$$

where f_{irr}, f_{hydro} and f_{muni} are the deficit functions for irrigation, hydropower and municipal water respectively.

1. **Agricultural Objective Function**

   The reservoir releases are pushed to fit irrigation demand through the minimization of the objective function given in Equation 4.9 :

$$f_{irr} = \sum_{t=1}^{T} \sum_{i=1}^{N} \left(ID_i(t) - Ir_i(t)\right)^2 \tag{4.9}$$

2. **Hydropower Objective Function**

   The goal is to balance power demand with supply. It can be realized by minimizing the objective function given in Equation 4.10:

$$f_{hydro} = \sum_{t=1}^{T} \left(D(t) - \sum_{i=1}^{N} P_i(t) \right)^2 \tag{4.10}$$

3. **Water for Municipalities Objective Function**

   This component encompasses all the water uses for both households and small industries within the region i. Based on the national regulations for planning water resources, the drinking water and the home-use water should be fully met even though it may have less economic returns compared with the other sectors. If the municipal i water demand is $MD_i(t)$, then the objective function will be defined as follows:

$$f_{muni} = \sum_{t=1}^{T} \sum_{i=1}^{N} \left(MD_i(t) - MA_i(t)\right)^2 \tag{4.11}$$

However, due to that fact that domestic water should be always fully satisfied, then we shall consider for anytime t, $MA_i(t) = MD_i(t)$. It will be integrated into the objective space as a constraint.

Thus, the multi-goal optimization problem is now transformed in to a bi-objective optimization problem which is described below:

$$\min \left(f_{irr}, f_{hydro}\right)$$

Constraints

1. According to Loucks , the power supply $[MWh]$ by plant i is given by the formula (Loucks et al., 1981):
$$P_i(t) = 2725 \cdot \eta_i \cdot R_i(t) \cdot h_i(t) \tag{4.12}$$

2. Water balance equations:
$$S_1(t+1) = S_1(t) + I_1(t) - Ir_1(t) - R_1(t) - Sp_1(t) - Evp_1(t) - MD_1(t) \tag{4.13}$$
$$S_i(t+1) = S_i(t) + I_i(t) - Ir_i(t) + R_{i-1}(t) + Sp_{i-1}(t) - R_i(t) - Sp_i(t) - Evp_i(t) - MD_i(t) \tag{4.14}$$
where $i = 2, ..., m$.

3. Reservoir release constraint:
$$R_i^{\min} \leq R_i(t) \leq R_i^{\max} \tag{4.15}$$

4. Reservoir storage limits:
$$S_i^{\min} \leq S_i(t) \leq S_i^{\max} \tag{4.16}$$

5. Spillage constraints:
$$Sp_i^{\min} \leq Sp_i(t) \leq Sp_i^{\max} \tag{4.17}$$
Further, to guarantee that water is not spilled unless the upper limit of the release is reached, the following constraint is imposed:
$$Sp_i(t) = \frac{[sgn(R_i(t) - R_i^{\max}) + 1]}{2} Sp_i(t) \tag{4.18}$$
where sgn is the sign function. However, in certain situations such as irrigation shortage, removing this constraint may reduce the irrigation deficit on the account of power production. In this case, DM may evaluate the MTHGIS model with and without constraint 4.18. Afterward, the best scenario can be adopted. However, since the study is carried under normal conditions, whereas spillage is not recommended (Zhang et al., 2016), Equation 4.18 is set active all the time. In this case, the released water has dual benefits power production and irrigation.

6. Reservoir water head:

$$h_i(t) = a \cdot S_i(t)^2 + b \cdot S_i(t) + c \tag{4.19}$$

7. Reservoir water spread area:

$$A_i(t) = a' \cdot S_i(t)^2 + b' \cdot S_i(t) + c' \tag{4.20}$$

8. Reservoir surface evaporation:

$$Evp_i(t) = \frac{A_i(t) + A_i(t+1)}{2} E_0(t) \tag{4.21}$$

9. Irrigation demand and bounds:

$$0 \leq Ir_i(t) \leq Ir_i^{\max} \tag{4.22}$$

10. Downstream and upstream quota for irrigation:

$$I_i^{\min} \leq \sum_{t=1}^{T} Ir_i(t) \leq I_i^{\max} \tag{4.23}$$

4.6.2 Solution Technique

The optimization model is solved using a prior method known as Weighted Sum (WS). The methodology of the WS method involves selecting scalar weights w_i such that w_i lies in the interval $[0,1]$ with $\sum_{i=1}^{2} w_i = 1$ and minimizing the following composite objective function:

$$F = w_1 \overline{f}_{irr} + w_2 \overline{f}_{hydro} \tag{4.24}$$

The normalized objective functions are determined according to the method suggested in the literature (Kim and de Weck, 2006). The method is described in the following:

$$\overline{f}_{irr} = \frac{f_{irr} - f_{irr}^U}{f_{irr}^N - f_{irr}^U} \text{ and } \overline{f}_{hydro} = \frac{f_{hydro} - f_{hydro}^U}{f_{hydro}^N - f_{hydro}^U}$$

Utopia point:

$$f_{irr}^U = f_{irr}(x_{irr}^*) \text{ and } f_{hydro}^U = f_{hydro}(x_{hydro}^*)$$

Nadir point:

$$f_{irr}^N = \max\left\{ f_{irr}(x_{irr}^*),\ f_{irr}(x_{hydro}^*) \right\} \text{ and } f_{hydro}^N = \max\left\{ f_{hydro}(x_{irr}^*),\ f_{hydro}(x_{hydro}^*) \right\}$$

x^*_{irr} and x^*_{hydro} are the optimal solution for the single objective optimization of f_{irr} and f_{hydro} respectively.

Such normalization is also important to ensure that the magnitudes of all performance metrics in the objective function are comparable (Marler and Arora, 2010). As a result, the new formulation of the MTHGIS is given in the following:

$$F = w_1 \overline{f}_{irr} + w_2 \overline{f}_{hydro} \tag{4.25}$$

$$w_1 + w_2 = 1 \tag{4.26}$$

where Equation 4.25 is subject to the constraints given in 4.12−4.23.

According to Zadeh, weights being positive guarantee that minimizing (4.25) provides a sufficient condition for Pareto optimality, which means that the minimum of (4.25) is always Pareto optimal (Zadeh, 1963). However, the selection of weights is based on objective preference level or specific trade-offs between power production or irrigation.

4.6.3 MTHGIS Model Implementation - Litani Project

Before implementing the suggested MTHGIS model, it should be noted here that, certain terms in the general model need to be adjusted to satisfy the profile of Litani Project. The constraints are then as follows:

1. Water balance equation - Qaraoun Lake:

$$S_1(t+1) = S_1(t) + I_1(t) - R_1(t) - Sp_1(t) - Ir_1(t) - Evp_1(t) \tag{4.27}$$

2. QD water head in terms of storage $[m]$:

$$h_1(t) = -0.0003 S_1(t)^2 + 0.2127 S_1(t) + 227.5096 \tag{4.28}$$

3. Qaraoun water spread area in terms of storage $[Km^2]$:

$$A_1(t) = -0.0001 S_1(t)^2 + 0.0686 S_1(t) + 1.5524 \tag{4.29}$$

Since Anan and Joun are very small reservoirs with relatively small heights with respect to their waterfall heights, the following assumptions are made: 1- the storage is assumed fixed since the given lakes maintain the same storage state over the planning horizon; 2- as a consequence, the waterfall head is also fixed; 3- the exposed surface for each lake is too small to provide an evaporation sources. Thus, evaporated water from the two ponds are neglected. Based on these assumptions, the associated constraints with the mentioned reservoirs are given below:

4. Water balance equation - Marakba pond (distributor):

$$R_2(t) = I_2(t) + Sp_1(t) + R_1(t) - Ir_2(t) \tag{4.30}$$

5. Hydraulic constraint Azour flow:

$$Sp_2(t) + I_4(t) = R_2(t) + I_3(t) \tag{4.31}$$

6. Water balance equation - Anan Lake:

$$I_4(t) = R_3(t) + Sp_3(t) + Ir_3(t) \tag{4.32}$$

7. Water balance equation - Joun Lake:

$$R_4(t) + Sp_4(t) = I_5(t) + Sp_2(t) + Sp_3(t) + R_3(t) \tag{4.33}$$

8. In order to prevent unnecessary spillage, the following equations are considered:

$$Sp_i(t) = \frac{[sign(R_i(t) - R_i^{max}) + 1]}{2}.Sp_i(t) \tag{4.34}$$

for $i = 1,3,4$.

$$Sp_2(t) = \frac{[sign(I_4(t) - {I_4}^{max}) + 1]}{2}.Sp_2(t) \tag{4.35}$$

where ${I_4}^{max}$ is the maximal value of Anan inflow.

In this study, averages concerning monthly power demand, monthly irrigation requirements and monthly flow data for different water sources were collected from the LRA database. The data are presented in Tables 4.5, 4.6 and Figure 4.8. It is assumed that the retrieved data are associated with a normal year.

Table 4.5 Monthly power and irrigation demands.

	Monthly power and irrigation demand											
	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Avg. power demand $\overline{D}(t)$ $[GWh]$	49.33	59.61	68.81	62.34	62.54	50.70	47.70	51.15	46.01	40.30	38.88	48.46
St. deviation $\sigma(t)$ $[GWh]$	31.54	26.83	33.92	36.44	31.63	26.18	23.27	18.56	15.81	18.27	22.16	28.48
Power demand - 2011 $[GWh]$	8.3225	28.82	43.0865	33.3345	62.183	60.4645	58.6735	68.0655	68.0655	61.5115	43.3725	50.07
Irrigation demand $[MCM]$												
Bekaa-Canal 900 project	0	0	0	0.30	1.12	1.76	2.12	1.84	1.52	0.54	0	0
Qasmieh project	0	0	0	1.18	2.64	4.15	5.02	4.34	3.59	2.08	0	0
Lebaa project	0	0	0	0.21	0.46	0.73	0.88	0.76	0.63	0.36	0	0

As an important reminder, the AMCP model presented in Chapter Three (Section 3.2) can be used to generate irrigation requirements for Bekaa, Kasmieh and Lebaa projects. Unfortunately, the exact

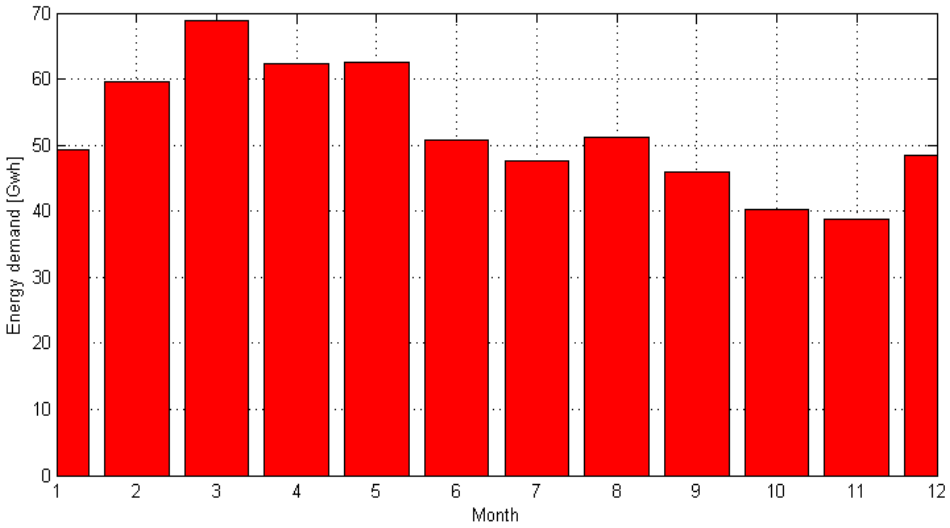


Figure 4.8 Average monthly energy demand [GWh].

crop types and the cropping area are both unknown. Therefore, it was wise to stick to the provided irrigation profile by Litani authorities in order to validate the suggested model.

Table 4.6 The average monthly streamflows in MCM: Litani river, Ain zarqa spring, Jezzine spring and Awali river.

	Average monthly inflows [MCM]											
Stream	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Litani	48.66	67.10	71.86	48.03	21.74	6.75	2.69	2.53	3.54	7.12	12.56	26.94
Ain -Zarqa	7.42	9.82	11.52	8.66	6.05	5.29	4.65	4.18	3.76	3.53	3.40	4.88
Jezzine	2.17	2.98	2.94	2.42	2.07	1.68	1.34	1.08	0.85	0.80	0.81	1.98
Awali	17.46	25.06	20.17	12.80	7.51	3.43	2.08	1.41	1.53	1.26	3.32	8.54

4.6.4 Results and Discussions

Using the available data from the Litani Project, the nonlinear formulation of the MTHGIS model is applied. Water discharges and water storage are simulated. The generated outcomes are the result of implementing WS method using LINGO 16.0 global optimizer.

Comparison results

To demonstrate the plausibility of the suggested model, a comparison is carried between the real operation implemented during the year 2011 and the simulated operation using the same power demand, irrigation requirements and river flow profile. During simulation, the priority order for water

release is agricultural, then power generation requirements. Relying on that preference, the weights are set as follows: $w_1 = 0.2$ and $w_2 = 0.8$ (the values of w_1 and w_2 can be achieved by experimentation) . The achieved results show that the MTHGIS model was able to fully satisfy power and irrigation requirements as in the actual operation over the whole planning period. Further results concerning monthly releases and monthly storage of QD are given in the Figures 4.9 and 4.10.

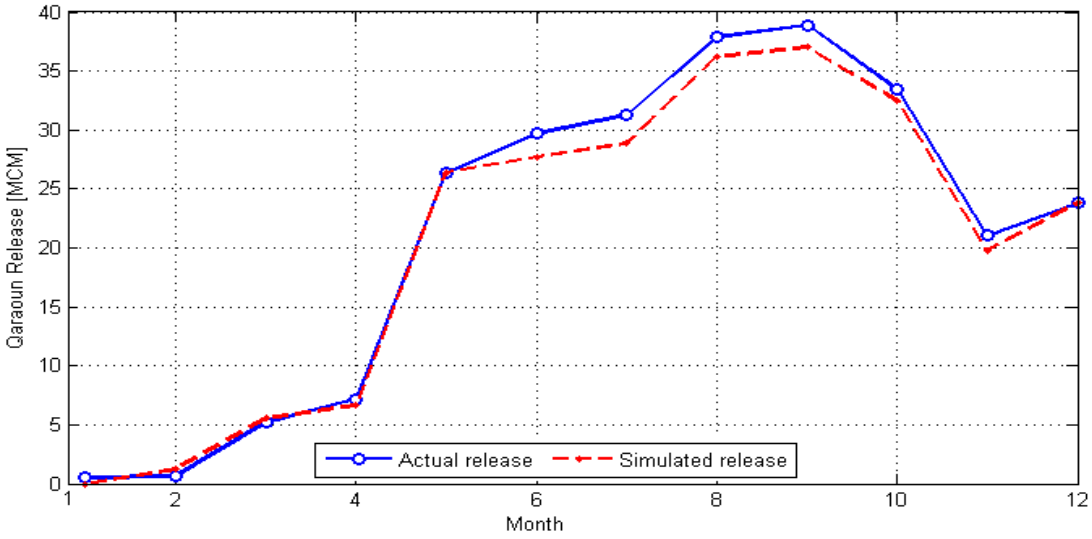


Figure 4.9 Simulated release versus actual release from QD during 2011.

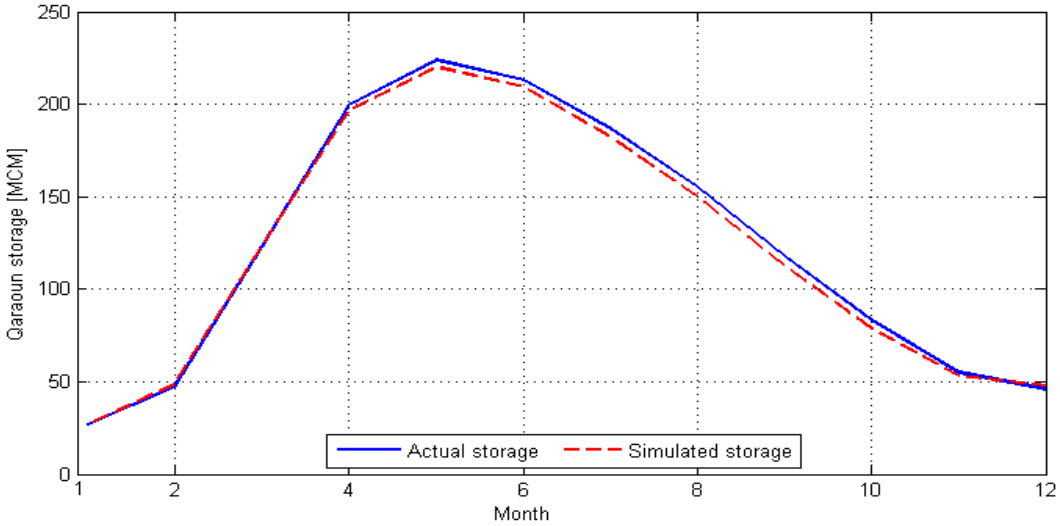


Figure 4.10 Simulated storage versus actual storage of QD during 2011.

Since the decision-making process is based on human expertise who gained the experience over the years, the outcomes of both simulated release and simulated storage were very close to the actual ones. However, it is noticeable that the simulated release is slightly lower than the actual discharge. In

fact, this is due to two main reasons: 1- the current operation is not optimal; 2- sometimes the units are running at low efficiency mode for a long time rather than running near optimal efficiency for a shorter time, which means more water discharges. Regarding the difference in storage, it is likely due to inaccuracies in reservoir storage measurements or in the river inflow data. Based on the preceding discussion, the MTHGIS model proved to be a valid and an effective tool in solving the medium term cascade hydropower-irrigation operation problem.

After the model is validated, four different simulations are considered using the data given in Tables 4.5 and 4.6. All scenarios have the same irrigation profile, but various load demand curves that are given in Table 4.7. Using the same setting for w_i's given previously, the MTHGIS model is applied

Table 4.7 Suggested scenarios according to load demand profile.

Scenario	Sim. 1	Sim. 2	Sim. 3	Sim. 4
Load curve	$\overline{D}(t)$	$\overline{D}(t)+0.1\sigma(t)$	$\overline{D}(t)+0.15\sigma(t)$	$\overline{D}(t)+0.20\sigma(t)$

using the four load curves. Figures 4.11 - 4.12 show respectively the evolution of the QD volume and the water discharge according to the operation schedule given by simulations: Sim. 1 - Sim. 4. Table 4.8 gives the shortage in each sector and the Qaraoun storage at the end of the planning horizon.

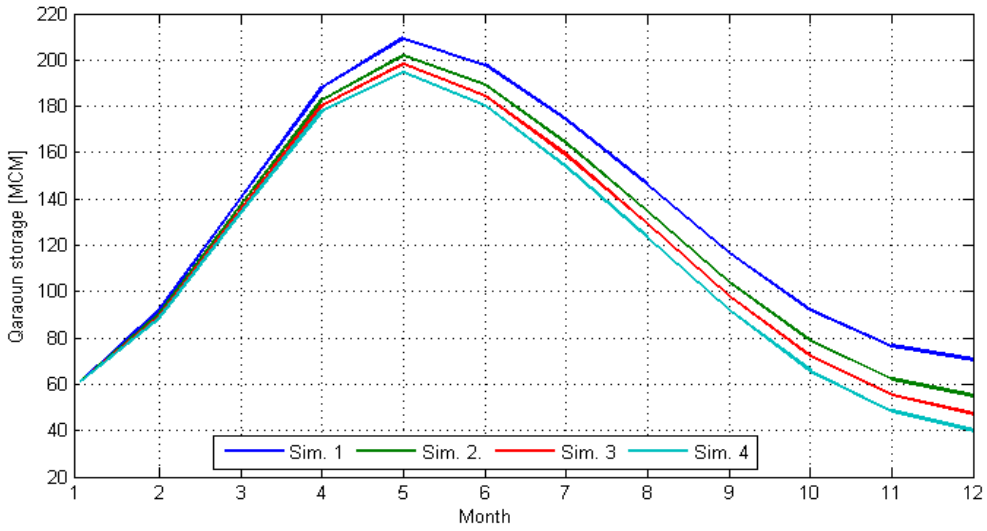


Figure 4.11 Simulated QD storage.

As it can be seen in Figure 4.11, the reservoir tries to store water during the wet season preceding peak load during Mars in order to take advantage of the higher head. The storage process stays in this way until May as QD reaches its maximum storage. Afterward, the water stock starts to descend as the Litani river starts to dry (Chapter Two - Figure 2.5).

In Lebanon, power generation is mainly dominated by fuel-gas based thermal power plants, while hydropower production does not exceed 4.5% from the total generation capacity (CDR - Lebanon).

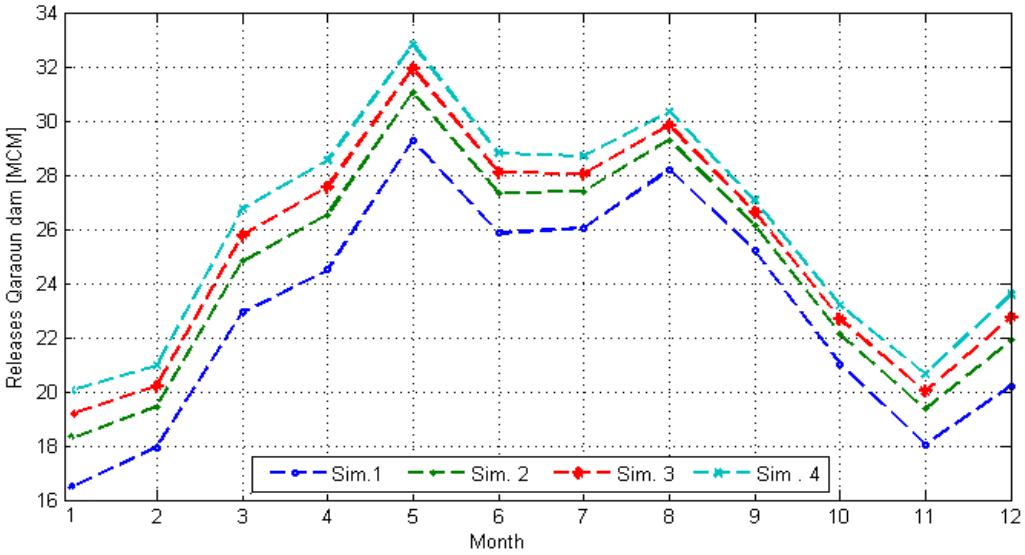


Figure 4.12 Simulated QD releases.

Hydropower stations in the country are used as load following power generation or sometimes for frequency regulation purposes of the grid. One of the main participants in the hydropower sector is the Litani Project operated by the Litani River Authority (LRA). Its main task is to meet the provided electricity demand by EDL. To test the model's interaction with load change, the average power demand is increased gradually by multiples 0.1, 0.15 and 0.2 of the standard deviation. Afterward, the models' outcomes are investigated. For Sim. 1 till Sim. 4 (Table 4.8), it obvious that energy and irrigation deficits are manifested in the last simulation. In addition, the main reservoir (QD) reaches the dead storage level at the end of the planning horizon.

Table 4.8 Summary of miscellaneous results.

Scenario	Power deficit percentage	Irrigation shortage percentage	Qaraoun final storage [*MCM*] beginning of Dec.	end of Dec.
Sim. 1	0%	0%	70.62	76.99
Sim. 2	0%	0%	55.07	59.87
Sim. 3	0%	0%	47.30	51.22
Sim. 4	**0.02%**	**1.34%**	**40.00**	**43.10**

Such situation has a severe consequence on the hydropower-irrigation cascade operation, environment and recreation. Therefore, it is recommended that the operator of the project not to increase power production till 0.2 standard deviation above average during a normal year. Alternatively, based on the present results in Table 4.8, it is suggested that the operator limit the increase in power supply by 0.15 standard deviation for two main reasons: 1- power and irrigation requirements should be fully satisfied; 2- at any time, QD storage should not be near the dead storage level, which is not a favorable condition

for effective hydropower-irrigation operation, environment and recreation. Overall, the suggested bounds for the hydropower production help system operators to manage normal fluctuations in power demand, and also it gives them more flexibility to carefully steer the reservoir storage level within the suggested specified red striped zone (Figure 4.13).

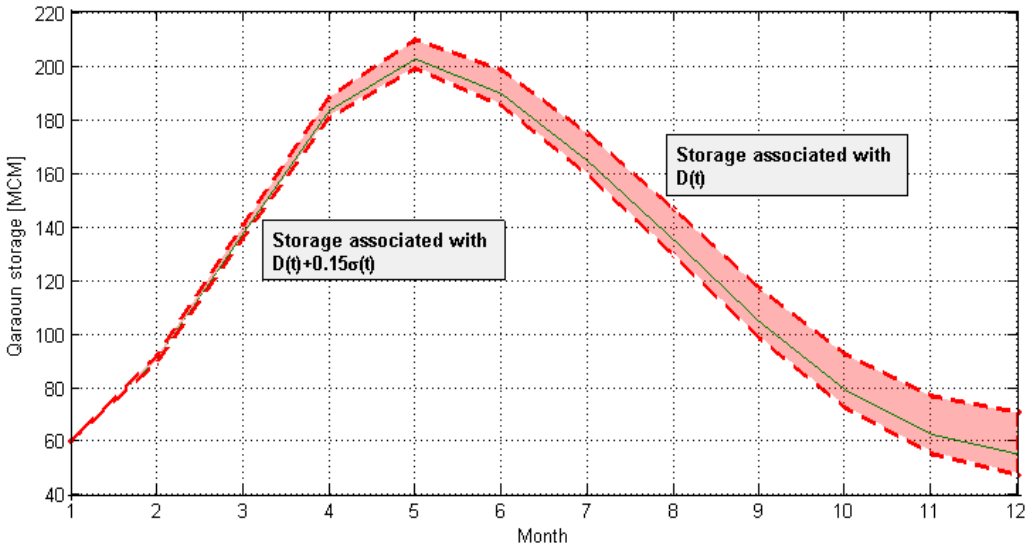


Figure 4.13 Qaraoun storage level: operation zone.

A closer look into Figure 4.13, the reader may notice that the strip is narrow during the first four months of the planning horizon while it is not the case with the remaining months. As it has been mentioned before, during the wet season and when Litani river water flow is high (Chapter Two - Figure 2.5), the reservoir tries rigidly to store water in order to take advantage of the higher head for power production. As a result, in the first four months of the year, the reservoir operation is not flexible. However, from the month of May and till the month of December, the reservoir's operating margin improved significantly, whereas Litani river water flow is at its lower levels.

Furthermore, someone may claim that, "Why the study didn't consider load profiles below average?". In fact, in Lebanon, the power sector has been notoriously characterized by a demand-supply deficit that is addressed in details in the Ministry of Energy and Water policy paper (MEW, 2010). Based on the preceding information, EDL has a high tendency to seek above average electric power from LRA in order to offset power shortage. Thereby, the consideration of other power demand cases shall not provide further important information.

Models' Running Time

For the suggested nonlinear MTHGIS optimization problem, it was solved on a 64-bits Windows 10 based computer with an Intel Core i7-5500U CPU @ 2.40 GHz, 12.0 GB RAM. The main model encompasses 217 variables and 417 equalities and inequalities constraints of which 86 are

nonlinear. The global optimizer of LINGO managed to produce optimal solutions using an optimality gap tolerance equal to 10^{-5} (stopping criterion), whereas the execution time for each simulation is exhibited in Table 4.9. The fast convergence presents an attractive choice to hydro-operators. The MTHGIS tool allows them to carry several scenarios using different load profiles (or irrigation profiles) and different river inflows in a short time period. This way, the operator will have a closer insight about the hydropower-irrgation operation from drought to floods.

Table 4.9 Execution times for different simulations.

Scenario	Sim. 1	Sim. 2	Sim. 3	Sim. 4
Execution time [s]	11	12	12	36

Medium term scheduling helps in calculating marginal water values and water share for the use in the more detailed short term scheduling model. In fact, the link between medium term and shorter term scheduling model is manifested in the following: 1- volume coupling, the medium and short term scheduling models are coupled through a fixed reservoir level at a given point in time. This level is calculated, using average values of power and irrigation demands, in the medium term model. Then the reservoir level is set as a boundary condition in the shorter term scheduling model. The presented procedure is rigid, and allows for very little flexibility in the case of deviations from the original scheduling assumptions (Doorman, 2012); 2- margin of action, the medium term scheduling model can be also executed over different load profile (or irrigation profile) scenarios. This way, several storage curves can be analyzed and evaluated. Consequently, a deviation from the fixed volume is allowed while taking into account the final reservoir storage and deficits in hydropower-irrigation sectors.

4.7 Short-term Scheduling of CHIP

In the previous section, the MTHGIS model for cascade systems has successfully provided reservoir storage bounds and a general overview of the hydropower-irrigation operation. Here in this part, the main interest is in a more detailed model concerning Short Term Hydro Generation Scheduling (STHGS) of cascade hydropower stations. However, to build a detailed realistic model, it is essential to consider all related elements to water availability, electricity production and system hydraulics. These elements are presented in Figure 4.14.

In this section, the main interest STHGS model of the CHIP. The aim is to suggest a novel model, whereas the objective is to minimize revenue losses in the hydropower operation. In fact, the presented objective function encompasses three terms related to power shortage, spillage and maintenance tasks. The power shortage term is motivated by the fact that hydropower plants are highly controllable. Therefore, while less controllable energy sources (such as nuclear, thermal,... etc.) are used as baseload power plants, hydropower plants can operate as load following and peaking power plants. In such

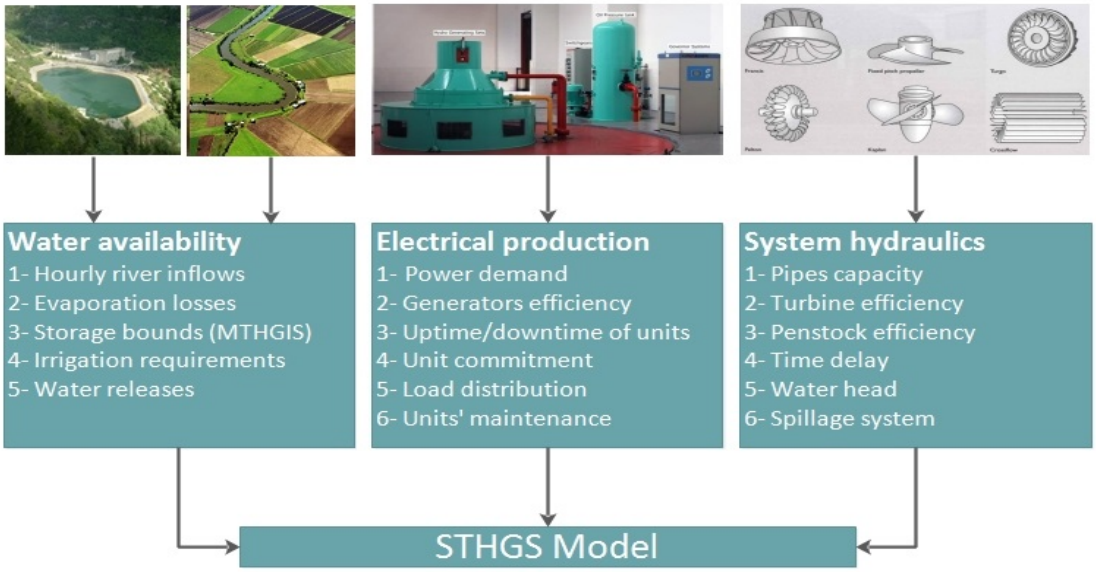


Figure 4.14 Influencing elements on STHGS model.

situation, fitting power supply with demand is recommended for hydropower generation. In fact, the minimization of energy deficit was considered in the work of Zhang, the goal is to find the optimal hourly water releases (R. Zhang and et al., 2013). However, the suggested approach may promote unnecessary spill. This problem is resolved by avoiding spill unless the turbine outflow reaches its maximum (Zhang et al., 2016) or by introducing spillage as a quadratic penalty term to the objective function (Yu et al., 2015). In this work, the spillage term is linear for more operational flexibility. Furthermore, one may claim that spillage in the upper reservoirs can help in improving the generation efficiency in the power production of lower reservoirs. This issue is resolved in Section 4.5 by: 1- maintaining high turbine efficiency through the right water releases bounds. This way, the improved solution of the unit commitment problem is manifested by keeping the unit running at high efficiency online, otherwise, the unit is offline. In fact, a similar approach is presented in the work of Lu (Lu et al., 2015); 2- efficient load distribution. Here, it has been proved mathematically that economic load dispatch can be better realized by equal load distribution between working units.

In addition, for a stable mean of power production, effective maintenance management is required. It can play a major role in sustaining the plant's reliability. During the literature review, most studies carried on cascade optimization do not consider preventive maintenance (Guedes et al., 2015). Their work models maintenance scheduling as a continuous variable using a nonlinear auxiliary function. Reservoir volume and the maintenance schedule were optimized, simultaneously, to minimize the thermal generation complement when hydropower is used as a baseload (Guedes et al., 2015). Here, the nonlinearity is overcame in the maintenance procedure by introducing it as a binary sub-problem with a set of linear constraints. Moreover, maintenance costs are incorporated into the objective function.

It is known that multipurpose hydropower reservoirs are designed to provide services beyond power generation, such as water supply for irrigation (Branche, 2017). Villavicencio presented a mixed integer linear programming model for scheduling short-term hydropower operation and takes into account irrigation requirements (Villavicencio et al., 2015). Therefore, in a hydropower-irrigation plant, it is important to consider irrigation agreements in order to have a realistic model.

The research aim is to achieve the optimal revenue, in fact a similar objective is presented in the literature (Pérez-Díaz et al., 2010). Their nonlinear programming model is concerned with maximizing revenues produced from selling the generated energy in a deregulated electricity market with no constraints on power production. Nevertheless, if the cascade hydropower system is operating as load following, it should follow a certain load profile. Any unbalance between supply and demand causes the power system to become unstable with the consequence of failure leading to potential severe economic and technical effects (Moseley and Garche, 2014). Therefore, to optimize the income it is recommended to minimize revenue losses due to power shortage.

Overall, this section seeks an objective to minimize revenue losses due to power deficit, spillage and maintenance tasks taking into account several influencing factors (electricity price, turbine efficiency curve, unit commitment, efficient load distribution, irrigation requirements, repairment tasks,... etc.). The outcome is a novel version of the STHGS model.
In fact, this section consists of the following:

- A new STHGS model is established with an objective to minimize revenue losses due to power deficit, spillage and maintenance tasks.
- Maintenance is integrated into the STHGS model through a set of linear constraints.
- The model is subjected to the improvements suggested by Section 4.5 to maintain high unit efficiency.
- The model is applied to a real case study (Litani Project) to evaluate its effectiveness.
  - Using real and forecasted streamflows, a comparison is carried between the current operation and the simulated ones to validate obtained results.
  - Model performance is tested also during two extreme events: 1- during peak streamflows; 2- generating units are operating near full capacity.

Overall, the outcome is a complete STHGS model which serves as a Decision Support Tool (DST). This multi-functional tool gives the cascade hydropower operator the opportunity to optimize revenue by enhancing the performance of the multi-purpose reservoir system on several levels: appropriate load dispatch, proper on/off unit sequencing, the optimum timing for unit maintenance and scheduled irrigation water allocation.

4.7.1 Mathematical Formulation of the STHGS Model

Objective function

The objective function for the STHGS model is represented in Equation 4.36. It seeks to minimize the revenue losses (€) due to power deficit, spillage and maintenance work throughout the planning horizon while prioritizing critical hydraulic and electrical constraints.

$$E = \min \left[\underbrace{\sum_{t=1}^{T} pr(t) \left(D(t) - \sum_{i=1}^{m} P_i(t) \right)}_{\text{Term 1}} \cdot 1hr + \underbrace{\frac{1}{3600} \cdot \sum_{t=1}^{T} pr(t) \left(\sum_{i=1}^{m} c_i \cdot Sp_i(t) \right)}_{\text{Term 2}} \cdot 1hr \right.$$
$$\left. + \underbrace{\sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{n_i} mc_{ij}(t) z_{ij}(t)}_{\text{Term 3}} \right] \quad (4.36)$$

In Equation 4.36, $pr(t)$ is the power selling price $[€/MWh]$ while $mc_{ij}(t)$ is the maintenance costs $[€/h]$ of unit j of plant i. Term 1: it comprises the financial loss related to power deficit term, $\left(D(t) - \sum_{i=1}^{m} P_i(t) \right) \cdot 1hr$; Term 2: the economic loss associated with spillage. It is known here that spillage is an amount of water that bypasses plant i. Thus, this volume is lost without being utilized for power generation at that plant. Thereby, to determine the financial loss, the reader proceeds as follows: using the coefficient of generation c_i of plant i given in Equation 4.37, the produced quantity $(c_i \cdot Sp_i(t)/3600) \cdot 1hr$ represents the generated energy $[MWh]$ when $Sp_i(t)/3600$ is used as a water release through the turbines of plant i. As a result, the total energy lost due to spillage at time t is $\left(\frac{1}{3600} \sum_{i=1}^{m} c_i \cdot Sp_i(t) \right) \cdot 1hr$. Therefore, the economical loss over the planning period is $\frac{1}{3600} \cdot \sum_{t=1}^{T} pr(t) \left(\sum_{i=1}^{m} c_i \cdot Sp_i(t) \right) \cdot 1hr$. Term 3: the maintenance cost of the generating units. In fact, $\sum_{i=1}^{m} \sum_{j=1}^{n_i} mc_{ij}(t) z_{ij}(t)$ represents inspection cost of all units of the m plants at time t. Then, the total maintenance cost for the entire planning period is given by $\sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{n_i} mc_{ij}(t) z_{ij}(t)$.

Spillage is avoided unless the turbine outflow reaches its maximum (Zhang et al., 2016). In other work, spillage is strongly penalized (penalty function is proportional to the square of the abandoned power through spilling) (Yu et al., 2015). However, in this study, the linear expression "Term 2" is added to the objective function in order to reduce unnecessary spill during the operation of the multi-reservoir system. In the case of no spill, the returned value is zero. On the other hand, if spillage occurs, its aim to reduce the revenue losses related to power deficit, taking into account the losses

induced by the spillage itself. Thus, the optimal operation is defined in a way that ensures reduced spillage. Moreover, one may claim spillage in the upper reservoirs can be useful to improve the efficiency in the production of lower reservoirs. To resolve this issue, turbine inflows bounds are introduced to maintain high efficiency all the time. This way, the occurrence of spillage is concerned with satisfying electricity demands and not to improve efficiency.

Constraints

In multi-purpose reservoirs, several hydraulic and electrical requirements, such as limits on power generation, water discharge rates, spillage, irrigation allocation, minimum uptime/downtime and maintenance duration should all be satisfied simultaneously.
The constraints are the following:

1. According to R.Zhang, the power produced $[MW]$ by unit j of plant i at time t can be written in the form (R. Zhang and et al., 2013):

$$p_{ij}(t) = c_i \cdot q_{ij}(t) \tag{4.37}$$

   In fact, Equation 4.37 is an approximation of the equation $p_{ij} = 0.00981 \cdot \eta_i \cdot h_i \cdot q_{ij}$. If the change in reservoir head is much smaller than the total waterfall elevation, then the head can be assumed fixed. In this case, Equation 4.37 can be considered as a good approximation.

2. The produced power is distributed equally among the different working units within each plant. The reason behind Equation 4.38 is to maintain high efficiency between online units. In fact, this constraint has been explained in Section 4.5:

$$p_{ij}(t) = pw_i \cdot x_{ij}(t) \tag{4.38}$$

$$p_i^{\min} \leq pw_i \leq p_i^{\max} \tag{4.39}$$

3. Total power produced by plant i at time t :

$$P_i(t) = \sum_{j=1}^{n_i} p_{ij}(t) \tag{4.40}$$

4. Power supply must not exceed power demand:

$$\sum_{i=1}^{m} P_i(t) \leq D(t) \tag{4.41}$$

   Moseley claims that excess power generation causes the frequency to rise and voltage to increase (Moseley and Garche, 2014). He continues that action must be taken before leading to potentially

severe economic and technical effects on the power system due to instability. Accordingly, the power system operator may consider restricting output of the generating units in order to ensure stability.

5. Reservoir release from plant i at time t :

$$R_i(t) = \sum_{j=1}^{n_i} 3600 \cdot q_{ij}(t) \tag{4.42}$$

6. Water balance equations:

$$S_1(t+1) = S_1(t) + I_1(t) - R_1(t) - Sp_1(t) - Ir_1(t) - Evp_1(t) - MA_1(t) \tag{4.43}$$

for $t = 1,..,T$.

$$S_i(t+1) = S_i(t) + I_i(t) - R_i(t) - Sp_i(t) - Ir_i(t) - Evp_i(t) - MA_i(t) \tag{4.44}$$

for $t = 1,..,\tau_{i-1}$ and $i = 2,..,m$.

$$\begin{aligned} S_i(t+1) = S_i(t) + I_i(t) + Sp_{i-1}(t-\tau_{i-1}) + R_{i-1}(t-\tau_{i-1}) - R_i(t) \\ - Sp_i(t) - Ir_i(t) - Evp_i(t) - MA_i(t) \end{aligned} \tag{4.45}$$

for $t = \tau_{i-1}+1,...,T$ and $i = 2,..,m$.

7. Reservoir release constraint:

$$R_i^{\min} \leq R_i(t) \leq R_i^{\max} \tag{4.46}$$

Here, the values of $R_i^{\min}$ and $R_i^{\max}$ are given by the MTHGIS model.

8. Reservoir storage limits:

$$S_i^{\min} \leq S_i(t) \leq S_i^{\max} \tag{4.47}$$

Also here, the values of $S_i^{\min}$ and $S_i^{\max}$ are obtained from the MTHGIS model.

9. Spillage constraints:

$$Sp_i^{\min} \leq Sp_i(t) \leq Sp_i^{\max} \tag{4.48}$$

10. Discharge rates constraints when unit j of plant i is on duty:

$$q_i^{\min} \leq q_{ij}(t) \leq q_i^{\max} \tag{4.49}$$

11. Multipurpose hydropower reservoirs are built to provide services beyond electricity generation, such as water supply for irrigation (Branche, 2017). Thus, in order to achieve a realistic model

of such system, it is necessary to consider irrigation demand and bounds:

$$\sum_{t=1}^{T} Ir_i(t) = \mathrm{ID}_i \tag{4.50}$$

$$0 \leq Ir_i(t) \leq Ir_i^{\max} \tag{4.51}$$

In fact, in the hydropower-irrigation reservoir system modeling, ignoring constraints (4.50, 4.51) will produce scenarios beyond compare with the actual operation.

12. Water discharge rate ramp:

$$-rq_i \leq q_{ij}(t) - q_{ij}(t-1) \leq rq_i \tag{4.52}$$

Sometimes, a sudden sharp change in power production may abruptly alter the flow rate of water. The result is a violent change in pressure, leading to rupture of the pipes. This transient problem, known as water hammer effect. In this case, Inequation 4.52 is imposed to reduce the produced shockwave through a smooth flow transition.

13. Any generation unit cannot be turned on or off instantaneously, once it is committed or uncommitted. Thus, it is important to impose minimum uptime/downtime constraints to ensure that a given unit should be on (respectively off) for a certain number of time periods when it is turned on (respectively off). According to Zhu, the constraints are given by (Zhu, 2009):

$$x_{ij}(t) - x_{ij}(t-1) \leq x_{ij}(\gamma), \text{ for } \gamma = t+1, \ldots, \min(t + T_{up} - 1, T) \text{ and } t = 1, \ldots, T \tag{4.53}$$

$$x_{ij}(t-1) - x_{ij}(t) \leq 1 - x_{ij}(\gamma), \text{ for } \gamma = t+1, \ldots, \min(t + T_{down} - 1, T) \text{ and } t = 1, \ldots, T \tag{4.54}$$

14. Preventive Maintenance involves performing maintenance on a power generating unit prior to significant degradation and subsequent failure. This attempt allows problems to be fixed before the unit stops working. Otherwise, further damage to the related components will lead to greater downtime for repairs or replacement.
The Remaining Useful Life (RUL) represents the time during which the generating unit is able to produce power without breakdown (here RUL value is assumed certain).
As a consequence, if $\sum_{s=1}^{t} x_{ij}(s)$ is the number of operational hours for unit j of plant i up till time t, the degradation function $\Delta_{ij}(t)$ for the corresponding unit at time t is then given by the formula (Fitouri et al., 2016):

$$\Delta_{ij}(t) = \frac{\sum_{s=1}^{t} x_{ij}(s)}{RUL_{ij}} \tag{4.55}$$

where RUL_{ij} is the remaining useful life for unit j of plant i during the planning horizon. Moreover, if $\delta_{\max}$ is the maximum threshold of the degradation function, then we certainly have:

$$\sum_{s=1}^{t} x_{ij}(s) \leq \delta_{\max} \cdot RUL_{ij} \tag{4.56}$$

The binary variable z_{ij} is equal to one when the unit is under maintenance and zero otherwise. Then, it is necessary to consider:

$$x_{ij}(t) \leq 1 - z_{ij}(t), \text{ for all } t = 1, ..., T \tag{4.57}$$

If the maintenance window is W_i, then to guarantee that the unit will be under maintenance for the whole window period, the following constraint is imposed:

$$z_{ij}(t) - z_{ij}(t-1) \leq z_{ij}(\gamma), \tag{4.58}$$

for $\gamma = t+1, ..., \min(t+W_i-1, T)$ and $t = 2, ..., T$.
The maintenance is carried once during the planning period:

$$\sum_{t=1}^{T} z_{ij}(t) = W_i \neq 0 \tag{4.59}$$

Finally, the maintenance procedures should start before the critical limit ($\delta_{\max} \cdot RUL_{ij}$) :

$$\sum_{s=1}^{t} x_{ij}(s) \leq \delta_{\max} \cdot RUL_{ij} + \left(t - \delta_{\max} \cdot RUL_{ij}\right) \frac{\sum_{s=1}^{t} z_{ij}(s)}{W_i} \tag{4.60}$$

for $t = \delta_{\max} \cdot RUL_{ij}, ..., T$.
In the following, maintenance constraints are checked for validity and for potential conflicts. During the planning horizon, constraint 4.59 guarantees that at some time $t = t_0$, the quantity $\sum_{s=1}^{t} z_{ij}(s)$ is different from zero while it is equal to zero for $t < t_0$. Here, three cases are distinguished:

- Case 1: for $t = 1, .., t_0 - 1$; $\sum_{s=1}^{t} z_{ij}(s) = 0$ and constraint 4.60 becomes:

$$\sum_{s=1}^{t} x_{ij}(s) \leq \delta_{\max} \cdot RUL_{ij} \tag{4.61}$$

As a results, for $t = 1,..,t_0-1$; $z_{ij}(t) = 0$ and using Inequation 4.57, $x_{ij}(t) = 0$ or 1 as long as $x_{ij}(t)$ satisfies Inequation 4.61. In fact, if 4.61 becomes $\sum_{s=1}^{t} x_{ij}(s) = \delta_{\max} \cdot RUL_{ij}$ at some $t = t_0' < t_0$, then $x_{ij}(t)$ switches exclusively to 0 for all $t = t_0'+1,..,t_0-1$. Otherwise, it violates Inequation 4.61.

- Case 2: for $t = t_0,..,t_0+W_i-1$; $0 < \sum_{s=1}^{t} z_{ij}(s) \leq W_i$ and Inequation 4.60 becomes:

$$\sum_{s=1}^{t} x_{ij}(s) \leq \delta_{\max} \cdot RUL_{ij} + \left(t - \delta_{\max} \cdot RUL_{ij}\right) \cdot \varepsilon \tag{4.62}$$

where ε is any number such that $0 < \varepsilon \leq 1$. Moreover, it is noted that, the quantity $\left(t - \delta_{\max} \cdot RUL_{ij}\right) \geq 0$, since $t = \delta_{\max} \cdot RUL_{ij},...,T$
Therefore, using constraint 4.58, for $t = t_0,..,t_0+W_i-1$; $z_{ij}(t) = 1$ and $x_{ij}(t) = 0$ (by constraint 4.57). It is obvious that $x_{ij}(t)$ satisfies constraint 4.62 since:

$$\sum_{s=1}^{t} x_{ij}(s) = \underbrace{\sum_{s=1}^{t_0-1} x_{ij}(s)}_{\leq \delta_{\max} \cdot RUL_{ij}} + \underbrace{\sum_{s=t_0}^{t} x_{ij}(s)}_{zero}$$

- Case 3: $t = t_0+W_i,...,T$; $\sum_{s=1}^{t} z_{ij}(s) = W_i$ and Inequation 4.60 becomes:

$$\sum_{s=1}^{t} x_{ij}(s) \leq t \tag{4.63}$$

Here, Inequation 4.63 is always valid. Moreover, as maintenance procedures are terminated, $z_{ij}(t)$ is necessarily zero, otherwise it violates constraint 4.59, whereas, $x_{ij}(t)$ is free to be either 0 or 1 according to Inequations 4.57, 4.63.

Therefore, based on the preceding discussion, it is proved that the imposed maintenance constraints in the model are realistic, reliable and induce no conflict among each others.

15. Repair team unavailability for carrying maintenance:

$$z_{ij}(t) = 0, \text{ for } t = t_1,...,t_r \text{ and } r \in \{1,...,T\} \tag{4.64}$$

16. Storage at time $T+1$:

$$S_i(T+1) = \text{final storage of reservoir } i \tag{4.65}$$

4.7.2 Solution Approach of the STHGS Model

Despite the larger number of researches who have previously considered solution approaches, solving STHGS is still a very challenging task. Pereira states that translating these models in a computational language can be a hard task (Pereira et al., 2015). The large number of variables and constraints, important for an accurate representation of any electrical system, makes the code complicated and highly computational resource consuming.

Over time, several methods have been applied to solve the short-term hydropower operation scheduling problem, including: Linear Programming (Piekutowski et al., 1994), Dynamic Programming (DP) (tian Chenga et al., 2009), Mixed-Integer Linear Programming (MILP) (Chang et al., 2001) and Mixed Integer Nonlinear Programming (MINLP) (Catalao and Mendes, 2010). Moreover, Artificial Intelligent (AI) approaches, such as genetic algorithms (Silva et al., 2012), Simulated Annealing (Dudek, 2010) and Particle Swarm Optimization (Hinojosa and Leyton, 2012) have been also adopted to solve short-term hydropower optimal scheduling. All the suggested methods have certain drawbacks. In fact, DP can find the optimal solution of short-term operation scheduling problem theoretically, but it is difficult to solve large-scale problems due to the "curse of the dimensionality". However, Finardia claims that the MILP model has proved to achieve plausible results with fast computational times when compared to nonlinear cases (Finardia et al., 2016). Regarding AI methods, it can be an interesting option, especially when possessing a good knowledge of the problem. This knowledge is manifested through the expertise of the operator to properly calibrate the parameters (Gea et al., 2014). Despite meta-heuristic methods can produce fair solutions within reasonable computation time, the quality of the solutions is difficult to guarantee (Saravanan et al., 2013). Besides that, the limited availability of meta-heuristic tools and the long process of coding, deploying and testing, turned the researcher attention to a less demanding approach. With the advent of very efficient MILP solvers on the market, MILP formulations have become common (Finardia et al., 2016). Thus, a favorable way to solve MINLP model is by converting the nonlinear formulation into a MILP (Gea et al.,2014; Yang et al.,2017; Wang et al.,2017). This goal is achieved through a suitable numerical and algebraic methods, taking into account reducing errors in the final results.

The STHGS problem belongs to the class of NP-hard combinatorial optimization problems (Tseng, 1996). In order to overcome the difficulty regarding nonlinearity of the problem, it is necessary to convert it into a mixed integer problem. This transformation is achieved using simple integer algebra techniques.

Linearization of the produced power represented in Equation 4.38:

$$\begin{cases} p_i^{\min} x_{ij}(t) \leq p_{ij}(t) \leq p_i^{\max} x_{ij}(t) \\ p_i^{\max}\left(x_{ij}(t)-1\right) + pw_i \leq p_{ij}(t) \leq pw_i + p_i^{\min}\left(x_{ij}(t)-1\right) \end{cases} \tag{4.66}$$

Furthermore, the problem encompasses another two nonlinearities: 1- the relationship linking electrical power with water discharge and water head; 2- the relation joining evaporation and water head. In fact, a complete discussion is carried regarding these two issues in the case study.

4.7.3 STHGS Model Implementation - Litani Project

The STHGS formulation is applied in the Litani Project described in Subsection 4.3.2. Here, it is important to note that, certain terms of the general model need to be adjusted in order to satisfy the profile of the Litani Project. The constraints are then as follows:

1. Water balance equation - Qaraoun Lake:

$$S_1(t+1) = S_1(t) + I_1(t) - R_1(t) - Sp_1(t) - Ir_1(t) - Evp_1(t) \tag{4.67}$$

   The water intake for municipal use (MA_i) was removed from the water balance equations due to water inflow (Litani river) into the Dam is polluted.

2. Water balance equation - Marakba Pond (distributor):

$$R_2(t) = I_2(t) + Sp_1(t) + R_1(t) - Ir_2(t) \tag{4.68}$$

3. Hydraulic constraint Azour flow:

$$Sp_2(t) + I_4(t) = I_2(t) + I_3(t),\ t = 1,...,3 \tag{4.69}$$

$$Sp_2(t) + I_4(t) = R_2(t-3) + I_3(t),\ t = 4,...,T \tag{4.70}$$

4. Water balance equation - Anan Lake:

$$S_2(t+1) = S_2(t) + I_4(t) - R_3(t) - Sp_3(t) - Ir_3(t) \tag{4.71}$$

5. Water balance equation - Joun Lake:

$$S_3(t+1) = S_3(t) + I_5(t) + Sp_3(t) + R_3(t) - R_4(t) - Sp_4(t) \tag{4.72}$$

   $t = 1$.

$$S_3(t+1) = S_3(t) + I_5(t) + Sp_3(t) + R_3(t) + Sp_2(t-1) - R_4(t) - Sp_4(t) \tag{4.73}$$

   $t = 2,...,T$.

As it was mentioned in the previous section, the main task of LRA is to meet the provided electricity demand $D(t)$ by "Electricité du Liban (EDL)". In fact, EDL buys the generated hydroelectricity at a fixed rate 24 €$/MWh$ and resells it for a higher price to customers. Here, it should be remarked that the demand profile is set by EDL based on the requirements for transmission regulation and on the demand variability. This issue is resolved by considering the standard deviation of the load demand. In this case, it is given by: $D(t) = \overline{D}(t) + k\sigma(t)$, where $\overline{D}(t)$ is the average power demand, $\sigma(t)$ is the standard deviation and k is a real number obtained from the Equation 4.74;

$$P(D(t) \leq \overline{D}(t) + k\sigma(t)) = (1-\alpha)\% \tag{4.74}$$

where α is the significance level.

According to Robinson, more than 80% of the water in the Qaraoun Lake is used to generate electricity in Markaba, Awali and Charles Helou hydropower plants (Robinson, 2013). The main goal is a controlled release of the stored water in order to minimize revenue losses due to power deficit and unnecessary spillage. Thus, by considering system characteristics, operational requirements and electricity selling price, the objective function encompasses Term 1 and Term 2:

$$E = \min 24 \sum_{t=1}^{T} \left(D(t) - \sum_{i=1}^{3} P_i(t) \right) \cdot 1hr + \frac{24}{3600} \sum_{t=1}^{T} (c_1 \cdot Sp_1(t) + c_2 \cdot (Sp_2(t) + Sp_3(t)) + c_3 \cdot Sp_4(t)) \cdot 1hr \tag{4.75}$$

The assumed values of c_1, c_2 and c_3 are the generation coefficients for Markaba, Awali and Charles Helou respectively. It is known that water spilled at QD cannot be used for power generation at Markaba station. In this case, the lost energy (MWh), at any time t, is equal to the quantity $(c_1 \cdot Sp_1(t)/3600) \cdot 1hr$. By analogy, energy losses (MWh), due to spillage at Anan and Joun, are $(c_2 \cdot (Sp_2(t) + Sp_3(t))/3600) \cdot 1hr$ and $(c_3 \cdot Sp_4(t)/3600) \cdot 1hr$ respectively. Since at LRA, the maintenance stuff are public employees, they are paid fixed monthly salaries to carry usual maintenance tasks. Consequently, Term 3 is dropped from the objective function. In fact, routine maintenance expenses are under fixed costs category.

During the technical investigation executed on Markaba, Awali and Charles Helou hydropower plants, the following observations were made: 1- personal engagement in the on/off sequencing of the electric generating units. Sometimes units are running at a low efficiency level; 2- lack of suitable equipment (data acquisition, distributed control,...etc.); 3- decisions are based on personal expertise; 4- maintenance plan is not optimum. In fact, the operator has become a part of the process where he has to understand the way the plants and each of their parts operate. Moreover, during operational scheduling, calculations are manually performed due to lack of necessary automated instruments. Thus,

the operation depends solely on the skills of the operator which increases the risks regarding safety and reliability of the power system leading to poor hydropower exploitation.

4.7.4 Results and Discussions

The core of the decision system is based on the MILP model. It aims to minimize revenue losses associated with power deficit and spillage while prioritizing critical hydraulic, electrical, irrigation and maintenance constraints. A wise commitment schedule and proper power distribution among hydropower units can decrease significantly power deficit and spillage, and simultaneously increasing the safety and reliability of the power system. In order to test the model rigorously, several operational scenarios are simulated over a period of 3 days. Outcomes are then compared with the current process for validation purposes and to trace operation improvements. In fact, the selected scenarios represent special cases that occurred on different dates:

- Peak flow of Litani river, the main driver of cascade reservoir system. It took place on January 8, 9, 10, 2013.
- Flood of the main reservoir (QD) occurred on March 25, 26, 27, 2013.
- During the irrigation season July 1, 2, 3, 2013; (Irrigation requirements: Bekaa irrigation, 0.168 MCM; Kasmieh irrigation 0.137 MCM; Lebaa irrigation 0.167 MCM).

However, before going further with simulation procedures, certain influencing factors on the STHGS problem need to be considered. In fact, taking into account these factors improves the accuracy and the efficiency of the whole hydro generation operation.

Influencing factors on the STHGS problem

Power generation formula The generated power p $[MW]$ from falling water can be calculated using the formula:

$$p = 0.00981 \cdot \eta_T \cdot q \cdot h \tag{4.76}$$

where η_T is the dimensionless efficiency (eff.) of the hydropower unit, q is the flow $\left[m^3/s\right]$ and h is the height $[m]$ difference between the reservoir surface and the penstock outlet.
Generally, the calculation of η_T does not just account the turbine and generator efficiency, but the head loss due to flow friction in the penstock, i.e.:

$$\eta_T = \text{generator eff.} \times \text{turbine eff.} \times \text{penstock eff.} \tag{4.77}$$

Based on the gathered information from hydropower plants, the generator efficiency in all plants is 0.95%. Regarding the turbine efficiencies, Section 4.5 suggests near optimal average values. The

penstock efficiency for each plant is assumed fixed, their values are obtained using the original observed data. Then the generation coefficient, for any plant, is calculated using the formula $c(h) = 0.00981 \cdot \eta_T \cdot h$. Figure 4.15 represents the generated power as a linear function of the flow at Markaba station for different heights.

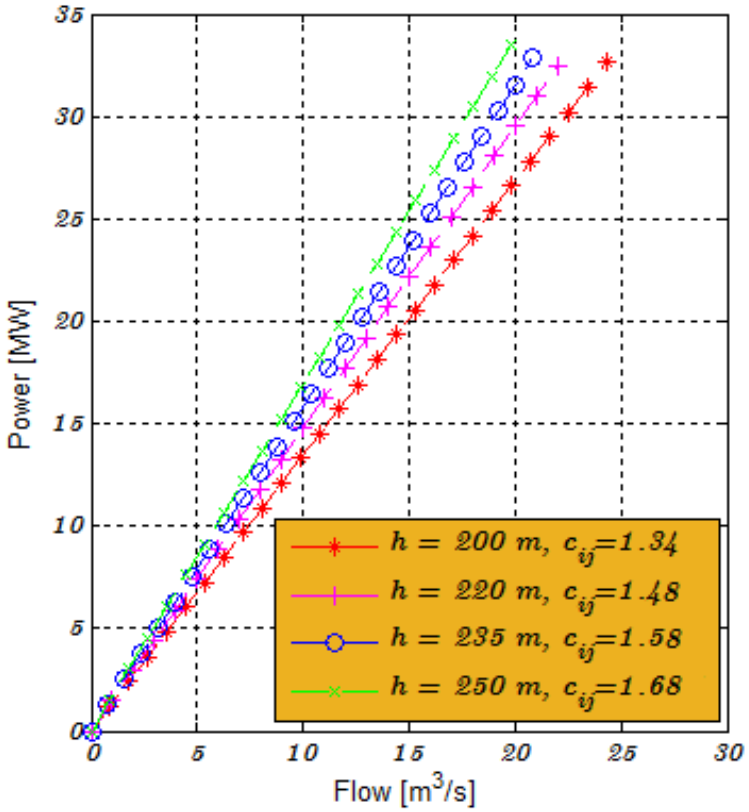


Figure 4.15 Production function of Markaba hydropower plant.

In the following, the effect of head change is investigated on the generation coefficient using two precision indicators. The absolute relative error and absolute error of the generation coefficient within ε m change in height are given in Equations 4.78 and 4.79 respectively:

$$\left| \frac{c(h \pm \varepsilon) - c(h)}{c(h)} \right| = \frac{\varepsilon}{h} \tag{4.78}$$

$$|c(h \pm \varepsilon) - c(h)| = 0.00981 \cdot \eta_T \cdot \varepsilon \tag{4.79}$$

The QD is located 200 m above Markaba plant. Furthermore, due to dead and active storage, the elevation of water in the Dam varies between 35 m and 56 m. Therefore, the total waterfall is at least 235 m and at most 256 m. Using Equations 4.78 and 4.79 with $\varepsilon = 3m$ (maximum change within 3 days), the relative error lies in the interval 0.0117 - 0.0127 and the absolute error is 0.019. The small values of the obtained errors indicate that these linear estimates (Figure 4.15) can be assumed accurate

within 3 m. But although the head is assumed to be constant within the 3 days planning horizon, the head is updated periodically and a new value of the generation coefficient is computed. This way, more realistic power generations results are obtained.

Concerning Awali and Charles Helou stations, their waterfalls are 400 m and 200 m respectively. Whereas, the maximum change in water elevation in Anan lake and in Joun lake, during the planning horizon, may reach 7 m and 5 m respectively. Table 4.10, exhibits the generation coefficients of Awali and Charles Helou stations with their corresponding errors.

Table 4.10 Generation coefficients of Awali and Charles-Helou stations.

Plant	Generation Coefficient	Absolute Relative Error	Absolute Error
Awali	3.20	0.0175	0.049
Charles-Helou	1.62	0.025	0.034

Based on the obtained values of the approximation errors, it may be assumed that the hydraulic heads are fixed in both plants. In this case, each plant is associated with only one coefficient of generation for any planning horizon.

Water losses due to evaporation Estimation of water losses, due to evaporation, is an integral component to ensure optimal short-term hydropower operation. In a way that, taking into account evaporation losses from the exposed water surface of the reservoir can improve the accuracy and precision of the hydropower model. Based on the proposed methodology in Chapter Two, the water volume lost in Qaraoun lake is estimated using the formula:

$$Evp_1 = E_0 \cdot A(h) \tag{4.80}$$

where $A(h) = 0.0001h^2 + 0.2703h - 4.9432$ millions m^2 is the surface area of the reservoir. A 3 m change in water elevation in the reservoir produces a change in water surface area according to the formula:

$$|A(h \pm 3) - A(h)| \simeq 0.0006h + 0.81 \leq 0.8436$$

It can be deduce that, during the warmest days, the difference in areas produces a maximum evaporation volume equal to 6748 m^3. In fact, this amount is negligible relative to the total evaporated volume. Thus, the surface area can be assumed fixed and equal to the lake surface area on day one of the scheduling. Based on that consideration, Table 4.11 displays the water losses due to evaporation over the planning horizon. However, the evaporated water from Anan and Joun lakes is negligible. This is due to the fact that, the exposed surfaces of both lakes are too small to provide evaporation sources.

During the literature review, water losses due to evaporation, in most papers, were neglected. However, in this study, the main concern is to reduce modeling errors by considering various influencing

Table 4.11 Evaporation losses at QD.

	Water losses in thousands of m^3								
Month	Jan.			Mar.			Jul.		
Day	8	9	10	25	26	27	1	2	3
QD	24.2	16.88	8.39	74.1	41.34	48.41	66.15	77.15	88.14

factors, including evaporation in order to present an accurate scheduling tool. To test the impact of evaporation on the scheduling results, two versions of the MILP model were considered: one neglects evaporation while the other does not. Upon solving, the scheduling results came slightly different. Bearing in mind that evaporation term did not increase the complexity of the suggested model, it is decided not to neglect water losses due to evaporation.

Maintenance Concerning maintenance, there is no clear framework provided by the LRA to facilitate the maintenance or the routinely inspection of the three hydropower plants of the Litani Project. However, effective maintenance management can play a major role in sustaining the plant's reliability, which in turn supports stable means of power production. In a sense that, coupling maintenance with hydropower operation can reduce significantly the risk of units failure or performance reduction. As a result, regular inspection procedures are required. But carrying maintenance at an unplanned specific time (repair team unavailability, peak loads,...etc.) may cause a power deficit or spillage. This way, economic profitability is not achieved. However, integration of maintenance constraints within the MILP model enabled the STHGS tool to identify the best time for inspection. This timing has the minimum negative effect on the underlying operation.

Hydrological Forecast The knowledge of the daily river flow is an important step for the resolution of tasks related to optimal short-term hydropower operation. Pérez-Díaz claims that for short time horizon (1 day), the model is deterministic with respect to water inflows (Pérez-Díaz et al., 2010). However, here the hydropower operation is considered up to 3 days ahead. Thus, to maintain optimality, accurate daily reservoir inflow forecasting is extremely important for the operation. In fact, in Chapter Two, a daily forecasting model for Litani river based on Two-Phase Constructive Fuzzy System Modeling (TPC-FSM) was suggested. The obtained outcomes were motivational enough to adopt the predictions as inputs in the STHGS model. As a recall, the predictive accuracy of the used fuzzy model was 84.2% (Nash-Sutcliffe efficiency) for 3 days ahead (Table 4.12). Regarding variability, it was able to preserve 84.6 % of the river variation. Therefore, accomplished results can provide an essential support to enhance the operational simulation.

Nevertheless, to fully investigate the performance of the STHGS tool, all water resources should be modeled too. Unfortunately, data concerning the flow of Ain Zarqa, Jezzine and Awali are scarce. To resolve this issue, Yevjevich claims that lognormal Equation 4.81 is the probability distribution that

Table 4.12 Litani river flow forecast - 3 days ahead for different dates.

	Litani River flow $[m^3/s]$								
Month	Jan.			Mar.			Jul.		
Day	8	9	10	25	26	27	1	2	3
Actual	134	124	100.6	14.68	13.75	13.63	1.09	0.76	0.82
Predicted	134	123	102.5	14.68	17.85	13.11	1.09	0.69	0.7

represents most reasonably the daily stream flow (Yevjevich, 1984).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \tag{4.81}$$

where the parameters μ and σ are respectively the mean and standard deviation of the variable's natural logarithm.

Using the available flow data, standard deviations were determined for Ain Zarqa, Jezzine and Awali streams. Fortunately, the obtained values showed that the streamflows possess small variability. Thus, the flow can be assumed to be constant for the entire 3 days planning horizon (Tables 4.13 and 4.14).

The probability density function given in Equation 4.81 is used to determine the Confidence Interval (CI) for different flows. In fact, the CI can play an important role to avoid unnecessary spills.

Table 4.13 Ain Zarqa and Jezzine flow forecast - 3 days ahead for different dates.

	Ain zarqa and Jezzine springs $[m^3/s]$								
Month	Jan.			Mar.			Jul.		
Day	8	9	10	25	26	27	1	2	3
Actual	6.67	4.21	4.72	3.27	2.51	3.74	2.73	2.41	2.1
Predicted	6.67	6.67	6.67	3.27	3.27	3.27	2.73	2.73	2.73
CI - 90%	[0.75, 6.81]			[0.44, 4.03]			[0.76, 2.35]		

Numerical Results and Analysis

To assess the validity and reliability of the proposed novel MILP formulation of the STHGS problem, it was compared with the actual operation implemented in the Litani Project. The computer, on which all of the computations were performed, was an Intel Core i7-5500U CPU @ 2.40 GHz, 12.0 GB RAM, running MS-Windows 10 (64-bits). The Human Machine Interface (HMI) display is based on a tight collaboration between Mathematical Programming, LINGO 16.0 and EXCEL. The linear version of the STHGS model has 6328 variables of which 1289 are integer variables and 9273 are

Table 4.14 Awali flow forecast - 3 days ahead for different dates.

	Awali river $[m^3/s]$								
Month	Jan.			Mar.			Jul.		
Day	8	9	10	25	26	27	1	2	3
Actual	7.87	7.51	7.02	3.59	3.92	3.45	0.55	0.63	0.60
Predicted	7.87	7.87	7.87	3.59	3.59	3.59	0.55	0.55	0.55
CI - 90%	[0.4, 11.8]			[1.63, 12.9]			[0.01, 2.44]		

equality and inequality constraints. Regarding execution time, all the simulated scenarios in this paper had a runtime between 15 and 102 seconds. The outcome is a simplified operating environment for hydropower scheduling. It allows the operators to quickly determine the optimal rules of water release, on/off sequencing, load distribution and maintenance time of generating units at the three plants.

The daily hydropower generation is expected to follow a given pattern of load demand at every hour over the planning horizon (Figure 4.16). The available water for the entire scheduling period is allocated for power generation in a way to fit, as much as possible, energy demanded issued by EDL.

In order to compare the simulated with the actual operation, the following settings are considered: the minimum downtime/uptime for the different generating units is assumed to be 2 hours; the maintenance window is set to 4 hours for routine maintenance. Generally, routine maintenance such as inspection, cleaning, tightening of nuts and bolts at suitable intervals is essential to minimize the risk of breakdown (Paish, 2002). The time delay is 3 hours between Qaraoun dam and Anan lake, while it is about 1 hour between Azour overflow and Joun lake.

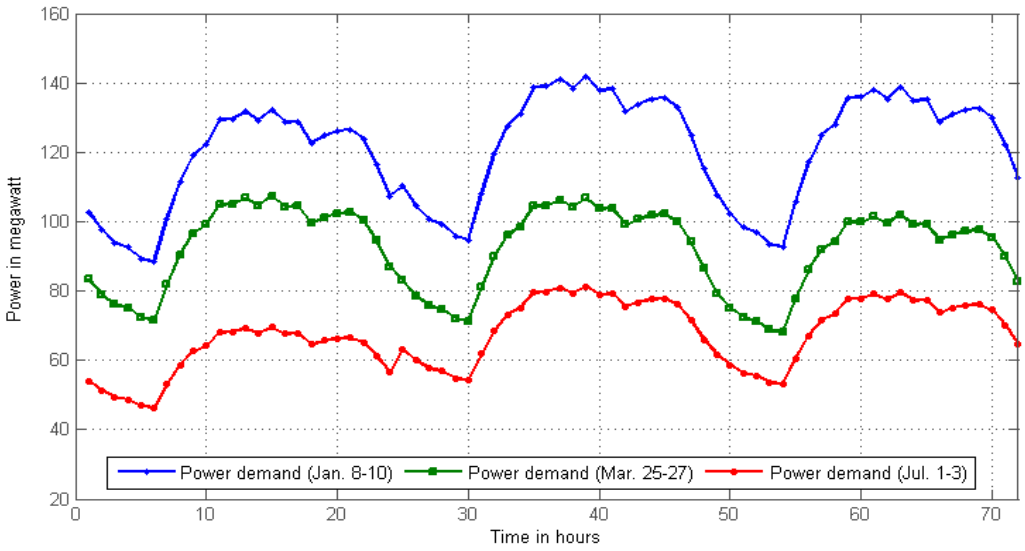


Figure 4.16 Power demand profile over different periods: Jan., 8-10, Mar., 26-27 and Jul, 1-3, 2013.

Table 4.15 Discharge and spillage of the actual\simulated operation.

Date	8-10 Jan.			25-27 Mar.			1-3 Jul.		
Release (*MCM*)	Actual	Simulation-1	Simulation-2	Actual	Simulation-1	Simulation-2	Actual	Simulation-1	Simulation-2
Markaba	3.439	3.179	3.185	2.691	2.825	2.872	2.262	2.590	2.487
Awali	4.786	4.763	4.761	3.760	3.643	3.632	2.820	2.712	2.716
Charles Helou	6.831	6.802	6.801	4.811	4.590	4.562	3.066	2.865	2.958
Spill (*MCM*)	Actual	Simulation-1	Simulation-2	Actual	Simulation-1	Simulation-2	Actual	Simulation-1	Simulation-2
Qaroun	0	0	0	0.246	0	0	0	0	0
Azour	0.109	0	0	0.104	0	0	0	0	0
Joun	0	0	0	0	0	0	0	0	0

Comparison of Results For the sake of comparison, two different simulations are considered (Table 4.15): Simulation-1 uses real flow data while Simulation-2 takes into account hydrological forecasts given in Tables 4.12, 4.13 and 4.14. In fact, the importance of Simulation-2 is to check the model's response to the deviation in stream flows. After solving the described MILP optimization problem, the total release, total spillage and the actual values of the current operation are exhibited in Table 4.15 for different dates of the year 2013. In fact, the presented simulations' outcomes were obtained as a result of the coordinated optimal plan between the three plants with zero energy deficit and zero spillage (revenue losses due to energy deficit and spillage is zero). In addition, the irrigation requirements were fully satisfied. It is noted that irrigation demand constraint was added to the model since part of the released water is diverted to irrigation on 1-3 July. In fact, removing the constraint, the difference between the actual releases and the simulated ones would be bigger.

Based on more than 50 years of experience, the operators of the cascade plants reached , using trial and error approach, a near optimal operation. This explains the close values between optimal releases of both simulations and the actual ones in the specified dates. In fact, release deviations (*MCM*) at Markaba, Awali and Charles Helou range between 0.077 - 0.13, 0.011 - 0.058 and 0.014 - 0.111 respectively (Table 4.15).

Furthermore, despite the deviation between the actual and predicted flows, the STHGS model managed to reproduce close releases between Simulation-1 and Simulation-2 (Table 4.16) in the three specified dates. Therefore, the small deviation in releases indicates that the suggested model is valid and suitable for hydropower operation optimization.

Table 4.16 Absolute difference between releases of Simulation-1 and Simulation-2 (over three days span).

	Absolute difference [*MCM*]		
Date	8-10 Jan.	25-27 Mar.	1-3 Jul.
Markaba	0.006	0.047	0.103
Awali	0.002	0.011	0.004
Charles Helou	0.001	0.028	0.093

Clearly, the uncoordinated strategy didn't utilize water properly. Spillage occurred at Azour overflow during January 8-10 and another significant amount of spillage occurred at both Qaraoun and Azour spillways during March 25-27. This is mainly due to the unplanned discharge rates. In practice, unnecessary spillage should be avoided since water is lost without being utilized for power generation at a certain level, unless it can reduce the power deficit . Alternatively, the release simulations presented in this study show that the spillage problem was avoided by employing a common pool management strategy. Table 4.17 shows the percentage of revenue losses during the actual and the simulated operation. Indeed, the induced spillage term in equation 4.75 has played an important role in reducing revenue losses and stopping unnecessary spills.

Table 4.17 Percentage of revenue loss due to spillage with respect to the total revenue achieved at the specified dates.

	Revenue loss		
Date	8-10 Jan.	25-27 Mar.	1-3 Jul.
Actual operation	1.12%	3.14%	0%
Simulation-1	0%	0%	0%
Simulation-2	0%	0%	0%

The last two reservoirs Anan and Joun have a very small storage capacity when compared to QD. Consequently, any sharp deviation above the presumed stream flow can easily trigger unnecessary spillage either at Azour overflow or Joun spillway by its upstream reservoir if water is released without precaution. This emerging problem is due to the stochastic nature of streamflows. However, to overcome this issue, the possible resolution of unnecessary spillage can be:

- at Azour, by decreasing the maximum discharges rate of the Qaraoun reservoir using suitable bounds, while at
- Joun, by reducing the maximum releases from Anan lake.

Actually, in this study and due to hydrological variation, the maximum discharge rate is reduced by an amount equal to the confidence interval upper bound given in Tables 4.13 and 4.14. This way, spillage can be prevented up to 90%. Despite the reduction in maximum releases, the STHGS tool managed to supply enough energy to satisfy the EDL demand without spillage occurrence. Therefore, by using the predicted flow of Litani river provided by TPC-FSM model and by fixing the other streamflows, STHGS model proved to be a plausible tool for achieving the following: 1- zero power deficit; 2- zero spillage; 3- zero irrigation shortage; 4- almost similar releases between Simulation-1 and Simulation-2. However, it is known that Litani river follows a pluvial regime. Thus, it is likely to have a sharp change in the river's flow (Chapter Two - Figure 2.21). In this case, it is recommended that the operator considers the new inputs and re-schedule the hydropower system. In fact, this task does not impose a

problem, especially that, the running time of the model is short. As a result, the rescheduling process can be triggered every time a sharp change in streamflow occurs assuming that the flow may disrupt the existing plan through reservoir flood or unnecessary spillage.

Table 4.18 illustrates, as an example, the power generation and the maintenance scheduling at the three power plants during the period January 8-10. The bold zeros represent the best timing for inspection during that period. The reader may notice how the model dealt with suggested settings: i- equal load distribution among online units; ii- minimum downtime/uptime limits are satisfied; iii- generating units working near optimal turbine efficiency; iv- maintenance tasks are carried out successfully.

Due to scarcity of records in the Litani Project, there is no exact information about the performance of the installed units. However, the operator claims that sometimes generating units do not work at high efficiency. To test his claim, it was determined, using the derived generation coefficients, the estimated power associated with the actual releases. Bearing in mind the estimation error presented in Subsubsection 4.7.4, the estimated power during March 25-27 is given in Table 4.19

If the units in the hydropower plants are working near optimal efficiency, it is expected that the actual generated power is near the estimated power. However, based on Table 4.19, it is not the case. In fact, taking into account estimation error the revenue losses range between 0.45%-3.79%. Nevertheless, in the suggested model, the revenue losses due to low efficiency is avoided since the online units are kept running at high efficiency all the time. Thus, by maintaining high generation efficiency water will not be wasted.

Based on the obtained outcomes, the MILP model manifested in the STHGS tool proved its effectiveness through the ability to avoid power deficit and unnecessary spillage. These results was a motivation to investigate further potentials of the STHGS tool by replacing certain data of the period January 8-10 by their extreme values. Based on recommendations of the director of the Litani project, the performance of the model is tested on two interesting extreme situations:

Case 1. The streamflows of both Bisri-Jezzine and Awali are pushed to 6.81 m^3/s and 11.8 m^3/s respectively (CI upper bounds).

Case 2. The load profile is increased by 25%. In this case, the hydropower plants are working close to full capacity.

Before going forward with the simulation of Case 1 and Case 2, the operator of Litani Project claims the following: **1-** peak loads can be met only for few hours before the power production drops to lower levels due to storage and hydraulics constraints. **2-** spillage takes place due to different reasons such as **2-a** flood of the reservoir; **2-b** overflow in the pipeline system; **2-c** certain generating unit or station is out of duty. Here, the main motive of the spill is to replace the missed water releases. For example, if Makaba station is out of duty, the operator is obliged to consider spillage to take advantage of the full operational capability of both Awali and Charles Helou stations, otherwise their operation is very limited.

Table 4.18 The detailed optimal power (MW) scheduling for Markaba, Awali and Charles Helou stations during the period January 8-10.

Period	Markaba		Awali			Charles Helou		Period	Markaba		Awali			Charles Helou	
	Unit 1	Unit 2	Unit 1	Unit 2	unit 3	Unit 1	Unit 2		Unit 1	Unit 2	Unit 1	Unit 2	unit 3	Unit 1	Unit 2
1	17.00	0	18.91	18.91	0	23.98	23.98	**37**	12.27	12.27	34.33	0	34.33	23.98	23.98
2	17.00	0	19.37	19.37	0	20.88	20.88	**38**	12.27	12.27	32.99	0	32.99	23.98	23.98
3	17.00	0	28.90	0	0	23.98	23.98	**39**	11.74	11.74	0.00	35.20	35.20	23.98	23.98
4	17.00	0	27.63	0	**0**	23.98	23.98	**40**	12.27	12.27	0.00	32.71	32.71	23.98	23.98
5	0	0	32.66	32.66	**0**	23.98	**0**	**41**	12.27	12.27	21.91	21.91	21.91	23.98	23.98
6	0	17.00	27.49	27.49	**0**	16.32	**0**	**42**	12.27	12.27	29.58	0	29.58	23.98	23.98
7	0	17.00	29.94	29.94	**0**	23.98	**0**	**43**	0	17.00	35.20	0	35.20	23.22	23.22
8	0	17.00	23.56	23.56	23.56	23.98	**0**	**44**	0	17.00	0	35.20	35.20	23.94	23.94
9	17.00	0	18.05	18.05	18.05	23.98	23.98	**45**	12.27	12.27	0	31.63	31.63	23.98	23.98
10	17.00	0	**0**	28.67	28.67	23.93	23.93	**46**	11.56	11.56	20.58	20.58	20.58	23.98	23.98
11	12.27	12.27	**0**	35.20	35.20	17.26	17.26	**47**	0	17.00	20.04	20.04	20.04	23.98	23.98
12	12.27	12.27	**0**	30.64	30.64	21.89	21.89	**48**	0	0	0	33.59	33.59	23.98	23.98
13	12.27	12.27	**0**	31.32	31.32	22.23	22.23	**49**	11.56	0	0	24.12	24.12	23.98	23.98
14	12.27	12.27	0	30.49	30.49	21.81	21.81	**50**	11.56	0	0	21.35	21.35	23.98	23.98
15	17.00	0	0	34.03	34.03	23.60	23.60	**51**	11.56	0	0	19.44	19.44	23.98	23.98
16	17.00	0	21.88	21.88	21.88	22.99	22.99	**52**	12.27	12.27	13.23	13.23	13.23	16.42	16.42
17	12.27	12.27	30.41	30.41	0	21.77	21.77	**53**	12.27	12.27	15.04	15.04	15.04	**0**	23.98
18	12.27	12.27	28.38	28.38	0	20.74	20.74	**54**	12.27	12.27	14.69	14.69	14.69	**0**	23.98
19	12.27	12.27	29.06	29.06	0	21.09	21.09	**55**	11.56	11.56	22.10	22.10	22.10	**0**	16.32
20	12.27	12.27	29.50	0	29.50	21.31	21.31	**56**	12.27	12.27	35.20	0	35.20	**0**	22.12
21	0	17.00	32.15	0	32.15	22.65	22.65	**57**	11.56	11.56	34.55	0	34.55	16.32	16.32
22	0	17.00	0	31.24	31.24	22.19	22.19	**58**	12.27	12.27	35.20	0	35.20	16.59	16.59
23	0	17.00	0	28.84	28.84	20.98	20.98	**59**	11.82	11.82	34.05	34.05	0	21.99	21.99
24	0	17.00	25.77	**0**	25.77	19.42	19.42	**60**	12.27	12.27	32.73	32.73	0	22.94	22.94
25	0	**0**	31.12	**0**	31.12	23.98	23.98	**61**	12.27	12.27	0	32.76	32.76	23.98	23.98
26	0	**0**	28.29	**0**	28.29	23.98	23.98	**62**	12.27	12.27	0	33.24	33.24	22.19	22.19
27	17.00	**0**	17.84	**0**	17.84	23.98	23.98	**63**	12.27	12.27	33.49	0	33.49	23.59	23.59
28	16.24	**0**	17.54	0	17.54	23.98	23.98	**64**	**0**	17.00	34.95	0	34.95	23.98	23.98
29	0	17.00	15.40	0	15.40	23.98	23.98	**65**	**0**	17.00	35.10	0	35.10	23.98	23.98
30	0	11.56	12.60	12.60	12.60	22.66	22.66	**66**	**0**	17.00	32.86	32.86	0	23.01	23.01
31	0	11.56	0	24.31	24.31	23.98	23.98	**67**	**0**	17.00	33.57	33.57	0	23.37	23.37
32	0	17.00	0	28.22	28.22	23.14	23.14	**68**	12.27	12.27	21.02	21.02	21.02	22.34	22.34
33	0	17.00	0	31.37	31.37	23.98	23.98	**69**	12.27	12.27	30.12	0	30.12	23.98	23.98
34	17.00	0	0	35.20	35.20	21.81	21.81	**70**	11.56	11.56	32.77	0	32.77	20.61	20.61
35	17.00	0	29.72	29.72	29.72	16.32	16.32	**71**	11.56	11.56	0	25.61	25.61	23.98	23.98
36	12.27	12.27	26.02	26.02	26.02	18.18	18.18	**72**	11.56	11.56	0	21.07	21.07	23.66	23.66

Table 4.19 Efficiency test: actual power versus estimated power.

	Generated power $[MW]$		Estimation Error $[MW]$	Revenue loss
	Actual	Estimated	Range	Range
Actual Operation	6630	6770	6660-6881	0.45%-3.79%

Under extreme circumstances, exceptional measures have to be employed. Thus, it is wise to reduce the inspection time in order to avoid, as much as possible, power deficit and unnecessary spillage.

The maintenance window is re-adjusted to 3 hours. Other parameters are kept fixed. After simulating both situations, Figure 4.17 exhibits power demand versus power supply during the assigned planning horizon. Apparently, in Case 1, the MILP model managed to present a plan with 0% power deficit. However, in Case 2, this quantity rises to 5.6%. Unsurprisingly, the power shortage is expected to happen in Case 2. This is due to inconsistency in flow capacity of the pipelines joining the different stations (Figure 4.4). In addition, the small storage capacity of Anan (170,000 m^3) and Joun (300,000 m^3) lakes leads to the fact that the reservoirs would reach dead storage in short time. As a result, the entire generation operation is negatively affected. Moreover, in Case 2, peak loads are met, for few hours (matches operator's claim **1**) after which the power supply deviated below the demand (Figure 4.17). In parallel, the storage of the small lakes had reached the dead storage at different periods (Figure 4.18) before they were filled back at the end of the planning horizon (due to the final storage constraint). The re-filling procedure brings down the water releases for power generation, which in turn explains the significant power shortage in the last stage of the scheduling. Meanwhile, removing the final storage constraint (Equation 4.65), the power deficit could be reduced for a price. It is more likely to have a state where either Anan or Joun lake are near dead storage. Such situation has a severe consequence on the cascade operation, environment, fisheries and recreation. Therefore, it is recommended to avoid removing the final storage constraints.

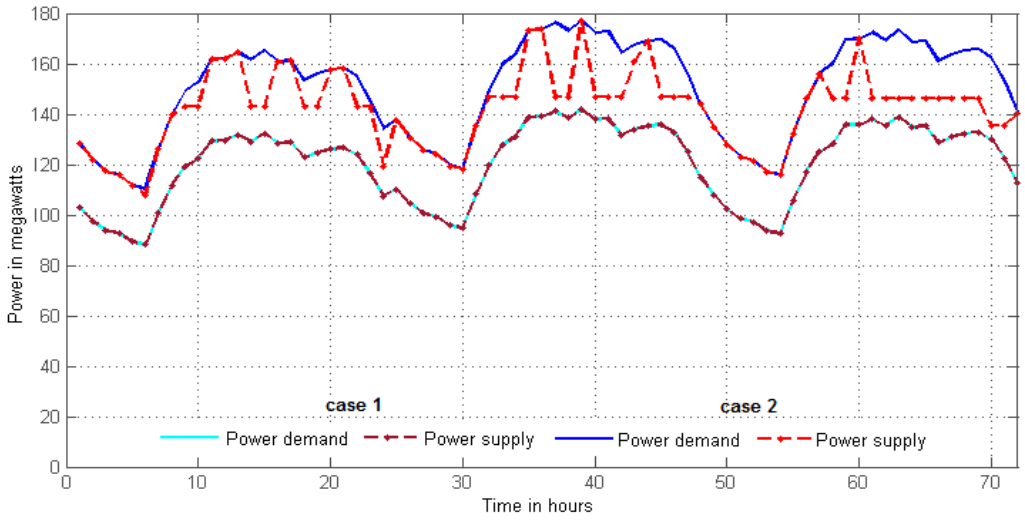


Figure 4.17 Power supply profile provided by the MILP model in response to the power demand described in Cases 1 and 2.

Another note, the discharge rate at Awali station can reach 33 m^3/s, while the maximum release at the Charles Helou station is 30 m^3/s. Thus, taking into consideration the Awali river inflow into Joun lake, Awali station cannot operate near full capacity all the time, unless spillage occurs at Joun at some

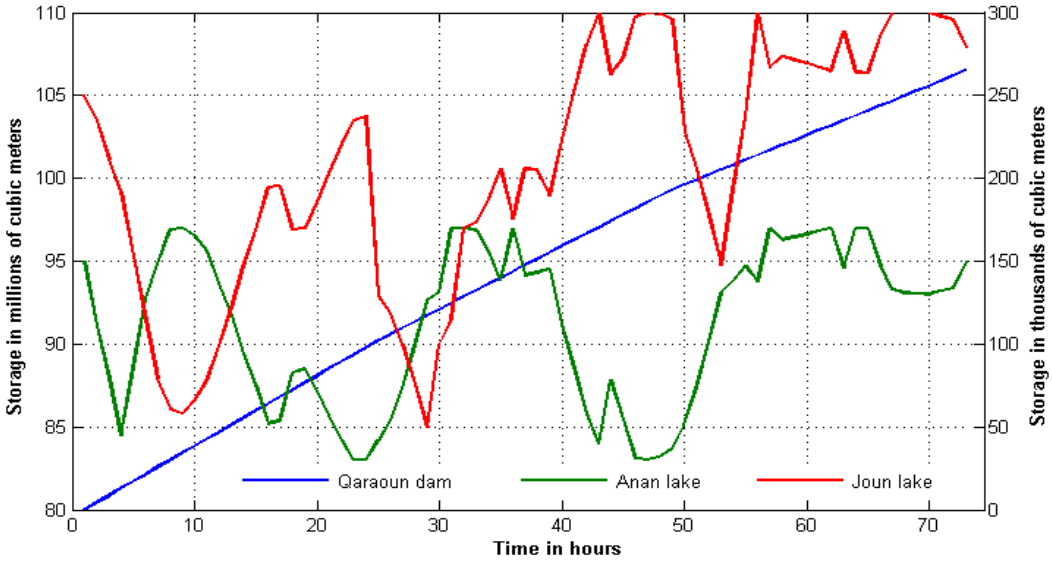


Figure 4.18 Case 2: Qaraoun Dam storage in millions of m^3 and the storage of Anan and Joun lakes in thousands of m^3 during the planning horizon.

time. In fact, Table 4.20 supports this assertion (matches operator's claim **2-a)**. Furthermore, one may notice also in Case 2 spillage from QD (Table 4.20). Upon checking the simulation, the spill occurred during the maintenance time of one of the generating units at Markaba (matches operator's claim **2-c**). The reason was to compensate the missed water source associated with the unit under maintenance.

Furthermore, Figure 4.17 provides an insight into the shortage of energy every hour. Therefore, based on Figure 4.17, EDL can look for other external power sources in order to compensate deficit in the generated power. This way power shortage issue is resolved.

Someone may claim that removing spillage penalty term will give more flexibility to the hydropower operation. To test this claim, extreme Case 2 is simulated such that the spillage term (Term 2) is removed from the objective function. The obtained outcomes are compared with the results presented in Table 4.20 (Case 2). Although the power shortage is reduced to 4.6%, the spills from Qaraoun dam and Joun lake are increased by 115% and 110% respectively. It is obvious that the reduction in power shortage is insignificant when compared with the significant increase in spillage. Therefore, the objective function given in equation 4.75 is recommended in order to preserve water.

Table 4.20 Further results on the optimal scheduling (Case 1 and Case 2).

Simulation	Spillage (*MCM*)			Power deficit	Financial	Financial
	Qaraoun	Anan	Joun	percentage	losses (€)	revenue (€)
Case 1.	0	0	0.23	0%	2486	207912
Case 2.	0.052	0	0.208	5.80%	17360	244848

Although Case 1 and Case 2 occasionally happen, the linear version of the STHGS model manages to provide an optimal operation with relatively small financial losses.

Maintenance Results Maintenance ensures reduced O & M costs; increased durability; reliable and uninterrupted power supply source. In parallel, if maintenance is not properly scheduled, it can promote power deficit and unnecessary spillage. The original intention of the STHGS tool is to integrate units' maintenance related activities within the hydropower operation. In fact, the goal is achieved by tightly connecting operation variable x_{ij} with the maintenance variable z_{ij} through a series of Inequations (4.57, 4.59 and 4.60). Thereby, for an optimal hydropower generation operation, a proper on/off unit sequencing is required, which in turn cannot be realized without the right maintenance timing. According to the simulations carried during the special events period, the MILP model manages to present a inspection plan with zero financial losses (as an example, Table 4.18). During extreme events: in Case 1, the units of Charles Helou were running all the time at full capacity except during maintenance. The reason was to soak up the extreme flow of Awali river, in addition to the releases of Awali plant. However, during the compulsory maintenance, the excess water volume still needs to be discharged, as a consequence 0.16 *MCM* was spilled. Furthermore, in the current operation, the operator claims that maintenance is avoided during peak power demand in order not to impede electricity production. Figure 4.19 shows the peak demand (red strip) considered in Case 2. As expected, the MILP model succeeded to schedule inspection away from the red strips. As a final attempt to check for flaws, the MILP model is also tested using a randomly generated maintenance intervals. The outcomes show significant amount of spilled water and power shortage, whereas all the produced revenue losses were greater than € 17360. In fact, sometimes the revenue losses had reached € 20000. Thus, it is clear now that the schedule provided by the STHGS tool represents the optimum plan for maintenance.

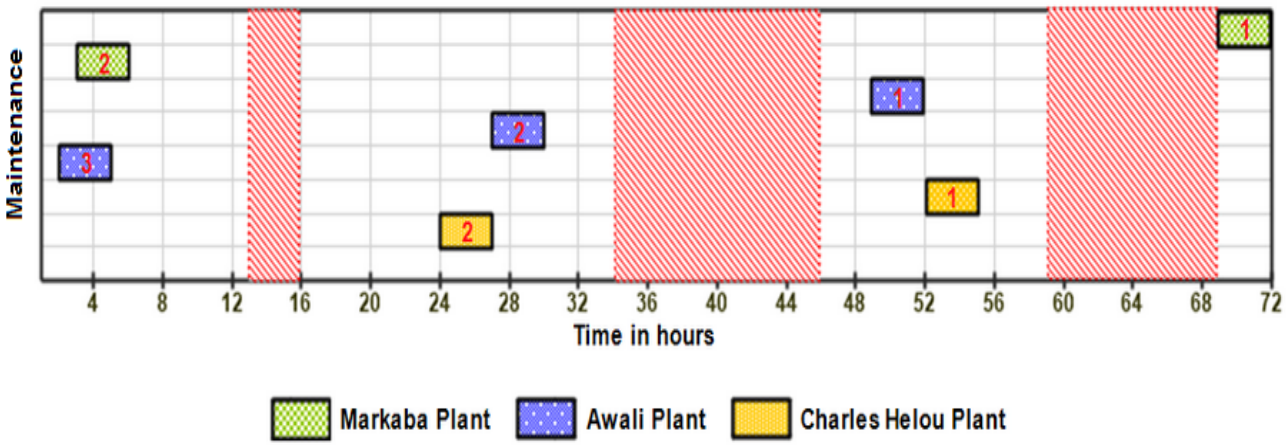


Figure 4.19 Case 2: Units maintenance schedule at Markaba, Awali and Charles Helou plants.

Models' Running Time For the proposed MILP optimization problem, the optimality gap tolerance is set to zero (stopping criterion) for all scenarios. As for the execution times, they range between 15 and 102 seconds. Here, the reader may notice a wide variation ($\sigma = 30.4$ sec.) in running times. In fact, according to Finardia, MILP performance is significantly influenced by data input, which interpret the excessively high computational times for some simulations (Finardia et al., 2016). Nevertheless, the fast convergence presents an attractive choice to hydro-operators. The STHGS tool allows them to carry different scenario analysis in a short time period before plan implementation. Furthermore, in a case of sudden failure of a generating unit, the operator has the ability to rapidly re-schedule and set a new plan to avoid insufficient power generation. This can be achieved by re-adjusting releases, where the offline unit can be compensated by putting another one into service. Overall, the STHGS tool has been successfully applied for energy management to get optimized energy deficit. Within a short time, it gives a complete insight to DM concerning water discharges, on/off sequencing and the reservoir storage during the whole planning period.

In general, the presented results show that the simulated operation intersects the current operation at different points (water releases, power fitting, dealing with system drawbacks, maintenance,...etc.). This due to the fact that, the decision-making process is based on humans expertise who gained experience over the years. In a word, the results fully demonstrate validity and effectiveness of the suggested STHGS tool in solving the cascade hydropower operation under special, extreme conditions and using hydrological predictions.

4.8 Conclusion

In this chapter, four major interconnected topics are considered: 1- estimating financial losses due to evaporation in the Litani Project; 2- improving the solution of the unit commitment and load dispatch problems using the efficiency curve scheme; 3- Medium-Term Hydro Generation - Irrigation Scheduling (MTHGIS) model ; 4- Short Term Hydro Generation Scheduling (STHGS) model.

In Section 4.4, an approach is described in order to study the annual financial impact of evaporation on Litani Project. First, based on the methodology presented in Chapter Two, the evaporation volume is estimated during the year 2013. The achieved result showed that the estimated water loss is around 17.88 *MCM*. Afterward, the economical effect of the lost water due to evaporation is evaluated. Unfortunately, the evaporated volume produced a tangible financial loss on hydropower and irrigation sectors. It is estimated that the hydropower loss is around 0.76 million euros, while the irrigation loss is drastic reaching a value of 39.22 million euros. Based on these results, the evaporation term cannot be simply neglected. Thereby, in this research, it is included in the medium as well as in the short term planning.

Section 4.5 presents a method to improve the solution of the unit commitment and economical load dispatch problems. In fact, using the efficiency curve scheme, two improvements are made

for the described problems: 1- maintaining near optimal turbine efficiency through the right water releases bounds. This way, the solution of the unit commitment problem is enhanced by keeping the unit running at high efficiency online, otherwise, it is offline; 2- distributing load efficiently among generating. Here, it is proved mathematically that economic load dispatch can be better realized by equal load distribution between working units. Therefore, by considering these measures, any generating unit will be working at its high efficiency during its scheduled plan. As a consequence, less water is released and more energy is produced.

Section 4.6 introduces a Medium-Term Hydro Generation-Irrigation Scheduling (MTHGIS) model. Its objective is to minimize power, irrigation and municipal water deficits. To test the model's plausibility, it was validated using a real case study (Litani Project). In fact, a comparison is carried between the real operation implemented during the year 2011 and the simulated one. The outcomes of the simulation came close to the data of the real operation. As a result, the model successfully passed the test for reliability. Afterward, several scenarios of various load profiles are simulated. The carried analysis, on the Lebanese case, relied on two conditions: 1- power and irrigation requirements should be satisfied; 2- the reservoir storage should not be near the dead storage at any time. The results indicated that the power production should not exceed 0.15 standard deviation above average during a normal year, otherwise the two conditions are violated. Overall, the introduced MTHGIS model managed, to a certain extent, to give a general overview of the hydropower-irrigation operation in the medium run. It also initialized boundary conditions with certain flexibility for the more detailed short term models.

Section 4.7 suggests a detailed Short-Term Hydro Generation Scheduling (STHGS) model. The significant innovations of the proposed model mainly include three points: 1- objective function represents revenue losses due to power deficit, spillage and maintenance tasks in the hydropower operation; 2- based on recommendations of the Section 4.5, the solutions of unit commitment and load dispatch sub-problems are enhanced within the STHGS model; 3- maintenance procedures are introduced to the model as a set of linear constraints as well as repairment costs are integrated into the objective function. During the modeling process, the achieved STHGS model came up with a rather complicated form with many nonlinear terms. However, to reduce its complexity, the STHGS problem was reformulated into a MILP model using numerical and algebraic techniques. Afterward, the performance of the suggested MILP is investigated in the Litani Project during special events (such as peak flow of the driving river, flood of the main reservoir and during irrigation season) and during extreme cases (such as peak flow of all water resources and hydropower stations are operating near full capacity). The achieved outcomes are evaluated by comparing simulations with details about the current method. For example, in the actual operation, a significant amount of spillage had occurred in the system on different dates. However, the use of a spillage term in the objective function, spillage is avoided except in extreme situations. At the same time, the intelligent integration of maintenance into the model guaranteed the economic profitability of hydropower plant. It is manifested through the determination of the best timing for machine inspection with minimum economical losses. Overall, the

establishment of the Intelligent Control-Maintenance-Management hydropower system has improved the short term operation through a series of optimal decisions starting with water discharges, on/off sequencing, load distribution and the best maintenance time.

In general, for optimal operation of the cascade plants, a Human Machine Interface (HMI) coupled with a decision core based on MTHGIS and STHGS paradigm, will be linked to a SCADA system for remote monitoring and control. It will effectively reduce the faults committed by human operators through centralizing the whole process with a fully automated system.

Chapter 5

Conclusions and Perspectives

> True optimization is the revolutionary contribution of modern research to decision processes
>
> *George Dantzig*

Development of a reliable and efficient hydropower-irrigation operation scheme requires the involvement of several disciplines such as hydrology, data mining, modeling, mathematical optimization, hydraulics and economics. In order to achieve this goal, a systematic literature review has been carried out to study various influencing factors on hydropower-irrigation operation. These factor mainly are: 1- streamflow forecasting models; 2- evaporation estimations; 3- optimization of water allocation and cropping pattern under deficit/non-deficit irrigation; 4- medium and short term optimal hydropower operation plans.

This chapter consists of five sections. Section 5.1 shows how different modules in the dissertation are interconnected for an efficient water resources management plan. Section 5.2 summarizes the methodology employed in hydrological modeling. The section also gives a closer insight about the results that emerged when the suggested methods are applied in the Litani Project. Section 5.3 recalls the two-stage mathematical programming model for optimal multi-crop planning and profit distribution, and offers some recommendations to DM on the basis of the analyzed results (Litani Project). Section 5.4 exhibits the optimal operation of cascade hydropower-irrigation plants followed in the research. It also explains why the proposed scheduling framework could be adopted in water resources planning projects similar to the Litani Project. Section 5.5 presents the future work.

5.1 Overview

In this dissertation, a modular modeling approach was adopted in which detailed modules were developed separately. The main strength of the modular approach is its ability to go into more details in each sub-field, and the ability to be independently improved and updated. These modules cover hydrological forecast, agricultural planning and hydropower modeling. During the research progress, different modules are linked through the backward and forward exchange of output data (Figure 5.1).

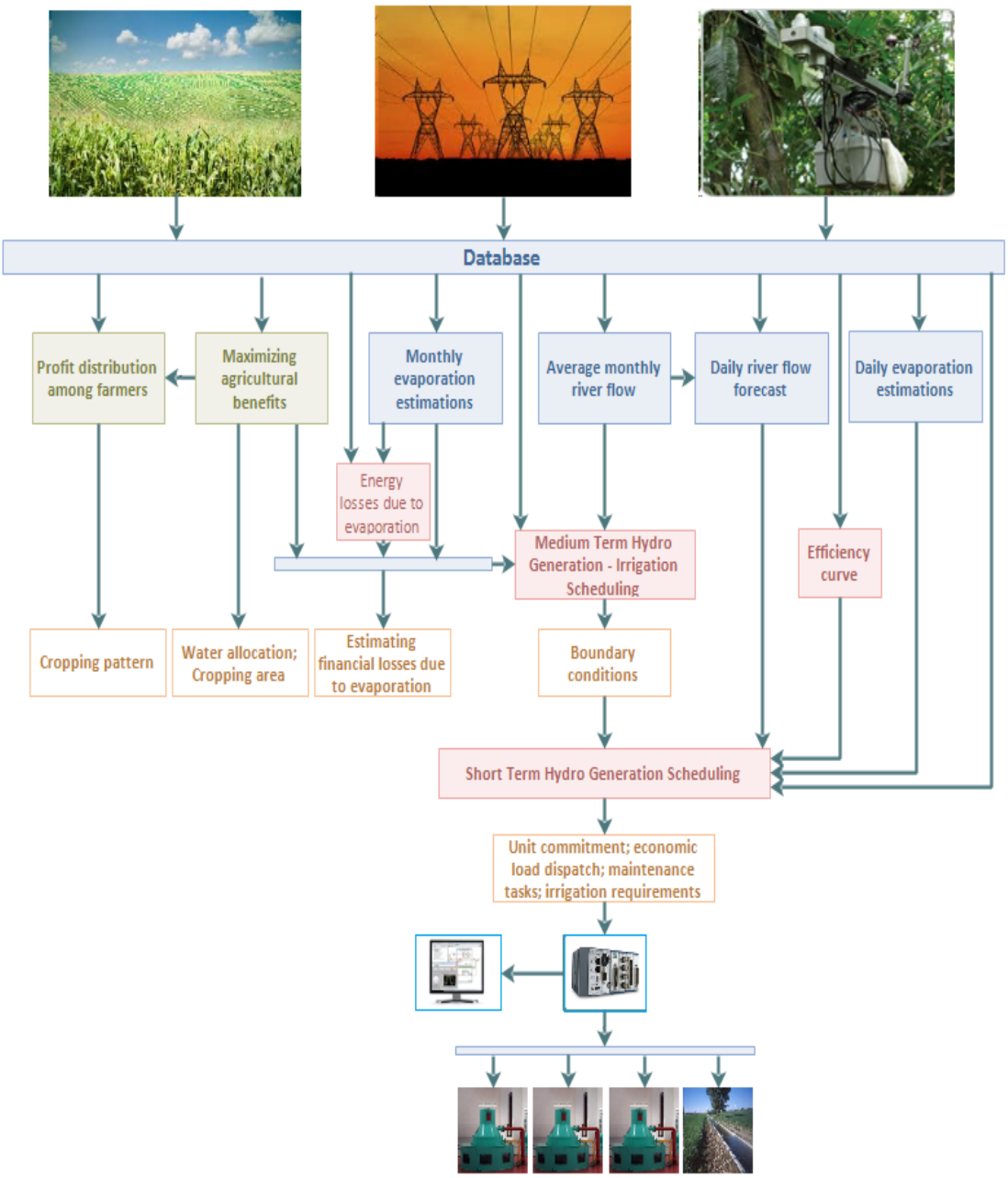


Figure 5.1 Different system modules connected to each other.

5.2 Hydrological Modeling

In this research, a preliminary statistical data analysis was conducted in order to determine the quality and the quantity of the collected hydro-meteorological data from the Litani basin. Unfortunately, the data suffer from insufficiency, inaccuracy, asymmetry[1] and sometimes unreliability in the information provided by the gauging stations. However, during the literature review, Fuzzy inference appears to be quite competent in handling these drawbacks. Thereby, a variant of Fuzzy inference, known as Constructive Fuzzy System Modeling (C-FSM), was used to model daily flow of the Litani river. In fact, the successful implementation of the C-FSM model was a motive to move forward to deal with the asymmetry issue. As a result, a hybrid model is presented based on Two Phase Constructive Fuzzy System Modeling (TPC-FSM). Here, the suggested TPC-FSM model out performed the classical C-FSM model. The results were promising with an overall good concordance (93%) between observed and predicted values for the longest lead day. Furthermore, data scarcity problem persists with the reservoir's surface evaporation estimation. However, the research managed to deliver a model, based on the Nonlinear Least Square (NLS) method and Simplified Penman formula, to estimate evaporation in poorly monitored lakes. In fact, this research offers several advantages in the hydrology field:

1. It incorporates climate-hydrology linkages using theoretical and empirical relationships, such as functional expressions relating: (i)- temperature-rainfall to river flow; (ii)- temperature-relative humidity- dew point to evaporation.

2. Although the data suffer from different drawbacks (scarcity, asymmetry, non-homogeneity, noisy,...etc.), the research suggests different Fuzzy inference models that are capable of reproducing daily river flow values accurately.

3. It presents a less-dimensional method to estimate lake evaporation as an alternative to commonly used approaches.

4. It provides a reliable hydrological prediction to help set up optimal hydropower-irrigation scheduling plans.

In developing countries, hydro-meteorological data required to quantify water availability are usually scarce, asymmetric and heterogeneous. However, with available low-resolution historical records, this research managed to reproduce water resource scenarios in the Litani basin. Therefore, the adopted methods here can be applied to other basin sharing the same problem. In fact, the methodologies are generic and are applicable to any other basin in the world.

[1] The term hydro-meteorological measurements asymmetry means that the climatic measurements exist for a shorter period in comparison with the hydrological ones.

5.3 Multi-Crop Planning and Profit Distribution

One of the main aims of cascade hydropower systems is irrigation. In fact, the research comprises two interlaced models: 1- Multi-Crop Planning; 2- Profit Distribution.

First, a Relaxed Nonlinear Programming (RNLP) model is established that describes Multi-Crop Planning (MCP) problem. One recalls that, during the MCP problem modeling, it is assumed the water availability for irrigation is fixed at every stage. Afterward, the model is evaluated in a likewise situation using real data from Bekaa valley. Here, the monthly diverted amount of water from Qaraoun dam to Bekaa valley relies on the experience of the Decision Maker. Based on the provided data, the achieved net profit is 13.073 million euros. Unfortunately, the diverted water is not fully utilized. Here, the research went a step forward by considering an adjusted version (AMCP) of the MCP model. It has the ability to re-arrange water availability in a way to maximize profit. Thus, using the same settings, the adjusted version managed to raise the profit to 17.112 million euros, that is a 31% increase in profit. As a result, some recommendations to DM are made:

1. At the case study site, there exists a huge information gap, especially when it relates cropping pattern to water availability. Thereby, by having pre-knowledge of the types of crops to be cultivated in the assigned area, the adjusted version of the MCP model will give the exact irrigation water requirements at each stage. At this level, it is guaranteed that no water will be wasted during the irrigation period. Nevertheless, for a successful AMCP model's application, data collection (crop types, cropping area, total water quota,...etc.) is necessary and recommended to the management authorities.

2. Since irrigation is a key component in any cascade hydropower-irrigation system, knowing the exact irrigation profile is vital for a reliable medium-term plan.

Furthermore, this research has suggested Medium-Term Hydro Generation-Irrigation Scheduling (MTHGIS) model for cascade systems. The objective is to minimize irrigation and power deficits. The allocation model has dealt with the average irrigation demand based on the historical data set. However, irrigation demand may be subjected to change due to modifications in cropping types or area. Here, the AMCP model can play an important role by providing the MTHGIS model with the irrigation demands each month. This way, the risk that water releases for irrigation is not fully utilized is reduced.

Another important issue that was discussed in the research is profit distribution among farmers or stakeholders. During the literature review, all studies optimize revenue by putting a general cropping plan. But there was no clear framework on how to implement the plan at the local level. As a result, several legitimate questions had emerged: 1- What are the types of crops are to be planted in each farmland within the planned area?; 2- What is the percentage of each crop type to be cultivated in each parcel?; 3- Can any suggested cropping plan on the local level guarantee equity among farmers or stakeholders? As a matter of fact, without setting any rules at the farmland level, the obvious scenario

is most of the farmers will select the most profitable crop suggested in the general plan. As a result, the proposed plan will be highly selective without a clear framework and it will not be fully exploited. The Profit Distribution (PD) paradigm came to resolve this issue based on a predefined cooperative policy. This method is effective in dealing with the complexity of managing profit among several farmland owners involved in the same agricultural project. However, one of the potential drawbacks is manifested in countries with un-free or oriented economy. This approach grants the highest profit to the farmer with the largest activity area (it is a little bit greedy). In fact, the resolution of the greedy issue can be achieved by replacing the activity area to total area ratio with a weight factor. This weight factor is proportional to the activity area, the farmer's financial situation and other criteria. This way, beginner farmers can improve their holdings in order to sustain their small business.

5.4 Operation of Cascade Hydropower-Irrigation Plants

Integrated reservoir operation is a must in any cascade hydropower-irrigation system. It should take into consideration agriculture activities, human settlements, industrial needs, recreation and environmental aspects. In fact, this dissertation has presented two phase reservoir operation plan: 1- Medium-Term Hydro Generation-Irrigation Scheduling (MTHGIS) model; 2- Short-Term Hydro Generation Scheduling (STHGS) model. During the modeling process, a significant attention has been paid to reduce errors in the obtained models. This was achieved by a careful treatment of various influencing factors (evaporation, unit's generation performance, waterfall head,...etc.) in the system. As a result, the two models were successfully validated by using data from the Litani Project. Thereby and based accomplished results, the MTHGIS and STHGS models can form the core of a Decision Support Tool (DST) for reservoir water management. This multi-functional tool gives the cascade system operator the opportunity to enhance the hydropower-irrigation operation by:

1. **Setting medium-term plan-** The main role of the plan is drawing a road map for the short-term scheduling problem and it is manifested in the following points:

    - The medium term module is based on multi-scenario deterministic optimization. It is able to simulate various operation using different input profiles (power, irrigation, streamflow). This way, the project operator can investigate the flexibility of the hydropower-irrigation operation concerning: reservoir storage; water discharges; power production; irrigation allocation.
    - The MTHGIS model can generate an individual endpoint water value description for the use in the STHGS model.

2. **Setting short-term plan-** The aim of the short-term hydropower production plan is to optimize revenue through the following:

- Optimal water releases for hydropower generation and irrigation
- Efficient unit performance
- Proper on/off sequencing of generating units within the hydropower plant
- Appropriate load dispatch
- Best timing for unit's maintenance
- Scheduled irrigation water allocation

Up to date hydropower-irrigation system planning remains an active research field and many studies are issued every year. However, the interesting aspect of this work is that it manages to divide the master hydropower-irrigation problem into several important sub-problems. The practical solution of each sub-problem acts as a supportive tool in enhancing the hydropower-irrigation operation. It is known that, in Lebanon and in many developing countries, the hydropower-irrigation operation is determined using a trial and error approach. The main reasons behind are: 1- poor quality and quantity of the available data; 2- absence of reliable data acquisition systems; 3- lack of expert involvement; 4- absence of efficient scheduling plans; 5- financial barriers for hydropower development. In fact, this research succeeded to deliver cheap and non-expensive approaches to overcome all mentioned obstacles and they are summarized in the following:

1. Different hydrological models are suggested that have the ability to deal with data drawbacks such as scarcity, inconsistency, heterogeneity and asymmetry.
2. A mathematical model is presented capable of providing reliable irrigation profiles.
3. A complete non-expensive DST is proposed to help cascade hydropower-irrigation system operators to make expert scheduling plans on both medium and short run.

In conclusion, the exhibited DST represents an attractive solution to the developing countries in hydropower-irrigation sector. Despite of the existing challenges, the DST managed to generate efficient plans useful for multi-goal water resources management.

5.5 Future Work

The outcomes of this research set a foundation for potential future work in this subject area. The following specific recommendations are made regarding future research areas to expand this work:

1. The PD model will be extended so that the profit is distributed between farmers according to a weight factor. This weight factor depends on several criteria such as activity area, the farmer's financial situation, the practices that may sustain environmental sound agriculture,..., etc. In order to determine these weights, a Fuzzy logic based model will be applied to generate their values.

2. The Pareto Front set of the MTHGIS problem will be constructed using an Adaptive Weighted Sum method. In addition, to deal with the stochastic nature of river flow, an effective approach will be suggested based on Model Predictive Control (MPC). Here, the operational decisions are acquired for the most probable inflow scenario provided by the forecasting model adjusted to the planning horizon.

3. The STHGS model will be tested on larger scale projects outside Lebanon and under dynamic electricity prices.

4. Based on the financial assessment of water evaporation carried at Qaraoun Lake, further study will be applied. It aims to determine the impact of installing floating PV panels over Qaraoun lake on evaporation, electric power and agriculture.

5. Taking into consideration floating PV panels installations, the STHGS model will be extended to handle such hybrid systems.

Bibliography

[1] A. Afshar and M. Mariño. Optimization models for wastewater reuse in irrigation. *Journal of Irrigation and Drainage Engineering*, (115):185–202, 1989.

[2] A. Afshara, O. Haddad, M.A. Marino, and B.J.Adams. Honey-bee mating optimization (hbmo) algorithm for optimal reservoir operation. *Journal of the Franklin Institute*, 2007.

[3] Silvia Maria Alessio. Digital signal processing and spectral analysis for scientists: Concepts and application. *Springer*, 2016.

[4] M. Aqil, I. Kita, A. Yano, and S. Nishiyama. A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *Journal of Hydrology*, (337):22–34, 2007.

[5] S. Asadia, J. Shahrabia, P. Abbaszadehb, and S. Tabanmehra. A new hybrid artificial neural networks for rainfall runoff process modeling. *International Work Conference on Artificial Neural Networks*, (121):470–480, 2013.

[6] Tao Bai, Jian xia Chang, Fi-John Chang, Qiang Huang, Yi min Wang, and Guang sheng Chen. Synergistic gains from the multi-objective optimal operation of cascade reservoirs in the upper yellow river basin. *Journal of Hydrology*, (532):758–767, 2015.

[7] Mohammad Ebrahim Banihabiba, Arezoo Ahmadianb, and Farimah Sadat Jamalib. Hybrid darima-narx model for forecasting long-term daily inflow to dez reservoir using the north atlantic oscillation (nao) and rainfall data. *GeoResJ*, (13):9–16, 2017.

[8] Sasan Barak and Saeedeh Sadegh. Forecasting energy consumption using ensemble arima–anfis hybrid algorithm. *Electrical Power and Energy Systems*, (82):92–104, 2016.

[9] D. Bolton. The computation of equivalent potential temperature. *Monthly Weather Review*, 1980.

[10] B. Bouchon-Meunier, M. Rifqi, and M.J. Lesot. Similarities in fuzzy data mining: From a cognitive view to real-world applications. *Computational Intelligence: Research Frontiers - Springer*, (5050):349–367, 2008.

[11] G. Box and G. Jenkins. Time series analysis: forecasting and control. *San Francisco: Holden-Day*, 1970.

[12] Emmanuel Branche. The multipurpose water uses of hydropower reservoir: The share concept. *Comptes Rendus Physique*, (18):469–478, 2017.

[13] W.A. Brock, W.D. Dechert, J.A., Scheinkman, and B. Le Baron. A test for independence based on the correlation dimension. *Econ. Rev.*, (15):197–235, 1996.

[14] A. Buck. New equations for computing vapor pressure and enhancement factor. *Journal of Applied Meteorology and Climatology*, 1981.

[15] California WaterBlog. Water managers drop the ball on hetch hetchy. *https://californiawaterblog.com*, 2016.

[16] J. P. S. Catalao and H. M. I. Pousinho ; V. M. F. Mendes. Mixed-integer nonlinear programming approach for short-term hydro scheduling. *IEEE Latin America Transactions*, (8):658–663, 2010.

[17] Fi John Chang and Meng Jung Tsai. A nonlinear spatio-temporal lumping of radar rainfall for modelling multi-step-ahead inflow forecasts by data-driven techniques. *Journal of Hydrology*, (535):256–269, 2016.

[18] G.W. Chang, M. Aganagic, J.G. Waight, J. Medina, T. Burton, and S. Reeves ; M. Christoforidis. Experiences with mixed integer linear programming based approaches on short-term hydro scheduling. *IEEE Transactions on Power Systems*, (16):743–749, 2001.

[19] Y.C. Cheng and S. Li. Fuzzy time series forecasting with a probabilistic smoothing hidden markov model. *IEEE Transactions on Fuzzy Systems*, (20):291–304, 2012.

[20] S.L. Chiu. A cluster estimation method with extension to fuzzy model identification. *World Congress on Computational Intelligence, Proceedings of the Third IEEE Conference*, 1994.

[21] V.T. Chow, D.R. Maidment, and L.W. Mays. Applied hydrology. *McGraw-Hill, New York*, 1988.

[22] P. Coulibaly and C.K. Baldwin. Nonstationary hydrological time series forecasting using nonlinear dynamic methods. *Journal of Hydrology*, (307):164–174, 2005.

[23] Q. Cuia, X. Wanga, C. Lib, Y. Caia, and P. Liangd. Improved thomas–fiering and wavelet neural network models for cumulative errors reduction in reservoir inflow forecast. *Journal of Hydro-environment Research*, (In Press), 2015.

[24] C.W. Dawson, R.J. Abrahart, and L.M. See c. Hydrotest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling and Software*, (22):1034–1052, 2007.

[25] S. Deodhar. Elementary engineering hydrology. *Pearson Education*, 2008.

[26] Y.B. Dibike and D.P. Solomatine. River flow forecasting using artificial neural networks. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, (26):1–7, 2001.

[27] Jason Donev. Energy education. *http://energyeducation.ca/*, 2017.

[28] G. L. Doorman. Hydropower scheduling. compendium in course elk15. *NTNU, Trondheim*, 2012.

[29] G. Dudek. Adaptive simulated annealing schedule to the unit commitment problem. *Electric Power Systems Research*, (80):465–472, 2010.

[30] ECODIT-led Consortium. *Strategic environmental assessment for the new water sector strategy for Lebanon*. Regional Governance and knowledge generation project, 2015.

[31] Brian Everitt. *The Cambridge Dictionary of Statistics*. Cambridge, UK New York: Cambridge University Press, 1998.

[32] FAO. *Food and Agriculture Organization of the United Nations*. 2017.

[33] FAO - Water Report. Deficit irrigation practices. *ftp://ftp.fao.org/agl/aglw/docs/wr22e.pdf*, 2002.

[34] FCH. First-hydro-company: one of the uk's most dynamic electricity generators. *http://www.fhc.co.uk/*, 2017.

[35] Thomas S. Ferguson. *LINEAR PROGRAMMING: A Concise Introduction*. University of California at Los Angeles, 2010.

[36] E.C. Finardia, F.Y.K. Takigawab, and B.H. Britoc. Assessing solution quality and computational performance in the hydro unit commitment problem considering different mathematical programming approaches. *Electric Power Systems Research*, (136):212–222, 2016.

[37] M. Firat. Comparison of artificial intelligence techniques for river flow forecasting. *Hydrol. Earth Syst. Sci*, (12):123–139, 2008.

[38] C. Fitouri, N. Fnaiech, C. Varnier, F. Fnaiech, and N. Zerhouni. A decison-making approach for job shop scheduling with job depending degradation and predictive maintenance. *IFAC-PapersOnLine*, (49):1490–1495, 2016.

[39] H. Galavi and L.Teang Shui. Neuro-fuzzy modelling and forecasting in water resources. *Scientific Research and Essays*, (7):2112–2121, 2012.

[40] Zhen Gao, Di Long, Guoqiang Tang, Chao Zeng, Jiesheng Huang, and Yang Hong. Assessing the potential of satellite-based precipitation estimates for flood frequency analysis in ungauged or poorly gauged tributaries of china's yangtze river basin. *Journal of Hydrology*, (550):478–496, 2017.

[41] N.K. Garg and A. Ali. "two-level optimization model for lower indus basin". *Agricultural Water Management*, (36):1–21, 1998.

[42] N. K. Garga and M. S. Dadhich. Integrated non-linear model for optimal cropping pattern and irrigation scheduling under deficit irrigation. *Agricultural Water Management*, (140):1–13, 2014.

[43] Xiaolin Gea, Lizi Zhang, Jun Shu, and Naifan Xu. Short-term hydropower optimal scheduling considering the optimization of water time delay. *Electric Power Systems Research*, (110): 188–197, 2014.

[44] D. George and M. Mallery. Spss for windows step by step: A simple guide and reference. *Boston: Pearson 10th edition*, 2010.

[45] P. E. Georgiou, D. M. Papamichail, and S. G. Vougioukas. Optimal irrigation reservoir operation and simultaneous multi-crop cultivation area selection using simulated annealing. *Irrigation and Drainage*, (55):129–144, 2006.

[46] P. E. Georgiou, D. M. Papamichail, and S. G. Vougioukas. Optimal irrigation reservoir operation and simultaneous multi-crop cultivation area selection using simulated annealing. *Irrigation and Drainage*, (55):129–144, 2006.

[47] L.S.M. Guedes, D.A.G. Vieiraa, A.C. Lisboaa, and R.R. Saldanhab. A continuous compact model for cascaded hydro-power generation and preventive maintenance scheduling. *International Journal of Electrical Power and Energy Systems*, (73):702–710, 2015.

[48] Qinkai Han, Fanman Meng, Tao Hu, and Fulei Chu. Non-parametric hybrid models for wind speed forecasting. *Energy Conversion and Management*, (148):554–568, 2017.

[49] M. Henkel. *21st Century Homestead: Sustainable Agriculture III: Agricultural Practices*. Lulu.com, 2015.

[50] V.H. Hinojosa and C. Leyton. Short-term hydrothermal generation scheduling solved with a mixed-binary evolutionary particle swarm optimizer. *Electric Power Systems Research*, (92): 162–170, 2012.

[51] L. Huamani, R. Ballini, I.G. Hidalgo, P.S Franco Barbosa, and A.L. Francato. Daily reservoir inflow forecasting using fuzzy inference systems. *IEEE International Conference on Fuzzy systems*, (4):201–213, 2011.

[52] D.A. Hughes, K.V. Heal, and C. Leduc. Improving the visibility of hydrological sciences from developing countries. *Hydrological Sciences Journal*, (59):1627–1635, 2014.

[53] International Resources Group (IRG). Litani river basin management plan. *USAID*, 2012.

[54] A. Jain and A.M. Kumar. Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, (7):585–592, 2007.

[55] A.W. Jayawardena and A.B. Gurung. Noise reduction and prediction of hydrometeorological time series: dynamical systems approach vs. stochastic approach. *Journal of Hydrology*, (228): 242–264, 2000.

[56] ME Jensen. Water consumption by agricultural plants. *NWISRL Publications, Academic Press, New York*, (2):1–22, 1968.

[57] J.K. Kaldellis, E-mail, D.S. Vlachou, and G. Korbakis. Techno-economic evaluation of small hydro power plants in greece: a complete sensitivity analysis. *Energy Policy*, (33):1969–1985, 2005.

[58] F. Karam, J. Breidy, C. Stephan, and J. Rouphael. Evapotranspiration, yield and water use efficiency of drip irrigated corn in the bekaa valley of lebanon. *Agricultural Water Management*, (127):125–137, 2003.

[59] I. Y. Kim and O. L. de Weck. Adaptive weighted sum method for multiobjective optimization: a new method for pareto front generation. *Structural and Multidisciplinary Optimization*, (31): 105–116, 2006.

[60] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science, New Series*, (220):671–680, 1983.

[61] O. Kisi, J. Shiri, and B. Nikufa. Forecasting daily lake levels using artificial intelligence approaches. *Computers and Geosciences*, (4):169–180, 2012.

[62] D. Nagesh Kumar and M. Janga Reddy. Multipurpose reservoir operation using particle swarm optimization. *Journal of Water Resources Planningand Management*, (133):192–201, 2007.

[63] D. Nagesh Kumar, K. Srinivasa Raju, and B. Ashok. Optimal reservoir operation for irrigation of multiple crops using genetic algorithms. *Journal of Irrigation and Drainage Engineering*, (132):123–129, 2006.

[64] L.B. Leopold, M.G. Wolman, and J.P. Miller. Fluvial processes in geomorphology. *New York: Dover Publications*, 1995.

[65] E.T Linacre. A simple formula for estimating evaporation rates in various climates, using temperature data alone. *Agricultural Meteorology*, 1977.

[66] DP Loucks, JR Stedinger, and DH Haith. Water resource systems planning and analysis. *Prentice Hall, Eaglewood Cliffs, NJ*, 1981.

[67] LRA. The characteristics of the litani river. *http://www.litani.gov.lb/*, 2016.

[68] P. Lu, J. Zhou, C. Wang, Q. Qiao, and L. Mo. Short-term hydro generation scheduling of xiluodu and xiangjiaba cascade hydropower stations using improved binary-real coded bee colony optimization algorithm. *Energy Conversion and Management*, (91):19–31, 2015.

[69] I. Luna, S. Soares, and R. Ballini. A constructive-fuzzy system modeling for time series forecasting. *International Joint Conference on Neural Networks*, 2007.

[70] José Fernando Ortega Álvarez, José Arturo de Juan, ValeroJosé María Tarjuelo, and Martín-BenitoEulogio López Mata. an economic optimization model for irrigation water management. *Irrigation Science*, (23):61–75, 2004.

[71] G. Magnus. Versuche über die spannkräfte des wasserdampfs. *Ann. Phys. Chem.*, 1844.

[72] R.T. Marler and J.S. Arora. The weighted sum method for multi-objective optimization: New insights. *Struct. Multidiscip. Optim.*, (41):853–862, 2010.

[73] Mathworks. *Least Squares Model Fitting Algorithms*. 2013.

[74] D. Mcjannet, F. Cook, and S. Burn. Evaporation reduction by manipulation of surface area to volume ratios: overview, analysis and effectiveness. *Technical Report for the Urban Water Security Research Alliance, Brisbane*, 2008.

[75] D.L. McJannet, M.P. Stenson I.T. Webster, and B.S. Sherman. Estimating open water evaporation for the murray-darling basin. *CSIRO Murray-Darling Basin Sustainable Yields Project*, 2008.

[76] O. Mehmmet. Comparison of fuzzy inference systems for streamflow prediction. *Hydrological Sciences*, (54):261–273, 2009.

[77] MEW. *Policy paper for the electricity sector*. Report 1. Beirut: Ministry of Energy and Water (MEW), 2010.

[78] S.H. Moore and M.C Wolcott. Using yield maps to create management zones in field crops. *Journal of Irrigation and Drainage Engineering*, (43):12–13, 2000.

[79] Patrick Moseley and Jurgen Garche. Electrochemical energy storage for renewable sources and grid balancing. *Newnes*, 2014.

[80] P.C. Nayak and K.P. Sudheer. Fuzzy model identification based on cluster estimation for reservoir inflow forecasting. *Hydrological Processes*, (22):827–841, 2008.

[81] C. M. Neale, K. Yilmaz, I. Yucel, and H. Gupta. New approaches to hydrological prediction in data-sparse regions. *IAHS Press*, 2009.

[82] P. Newbold, W.L. Carlson, and B.M. Thorne. Statistics for business and economics. *Fifth Version, Prentice Hall, Upper Saddle River, NJ*, 2003.

[83] R. Nisbet, G. Miner, and J. Elder. Handbook of statistical analysis and data mining applications. *Elsevier*, 2009.

[84] Hamideh Noory, Abdol Majid Liaghat, Masoud Parsinejad, and Omid Bozorg Haddad. Optimizing irrigation water allocation and multicrop planning using discrete pso algorithm. *Journal of Irrigation and Drainage Engineering*, (138):437–444, 2012.

[85] Esin Onbaşoglu and Linet Ozdamar. Parallel simulated annealing algorithms in global optimization. *Journal of Global Optimization*, (19):27–50, 2001.

[86] P.R. Onta, R. Loof, and M. Banskota. Performance based irrigation planning under water shortage. *Irrigation and Drainage Systems*, (9):143–162, 1995.

[87] O. Paish. Small hydro power: technology and current status. *Renew Sustain Energy Rev*, (6): 537–56, 2002.

[88] Z. Pap. Crop rotation constraints in agriculturalproduction planning. *Intelligent SISY. 6th International Symposium on Systems and Informatics*, 2008.

[89] H.L. Penman. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London, Series A Mathematical and Physical Sciences*, 1948.

[90] H.L. Penman. Vegetation and hydrology. technical committee no. 53. *Commonwealth Agricultural Bureaux, Harpenden, UK.124*, 1963.

[91] Sergio Pereira, Paula Ferreira, and A.I.F. Vaz. A simplified optimization model to short-term electricity planning. *International Journal of Electrical Power and Energy Systems*, (93): 2126–2135, 2015.

[92] M. Piekutowski, T. Litwinowicz, and R.J. Frowd. Optimal short-term scheduling for a large-scale cascaded hydro system. *IEEE Transactions on Power Systems*, (9):805–811, 1994.

[93] A. Porporato and L. Ridolf. Multivariate nonlinear prediction of river flows. *Journal of Hydrology*, (248):109–122, 2001.

[94] J. I. Pérez-Díaz, J. R. Wilhelmi, and J. Sánchez-Fernández. Short-term operation scheduling of a hydropower plant in the day-ahead electricity market. *Electric Power Systems Research*, (80): 1535–1542, 2010.

[95] C.H.B. Priestley, , and R.J. Taylor. On the assessment of the surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 1972.

[96] I. Pulido-Calvo and M.M. Portela. Application of neural approaches to one-step daily flow forecasting in portuguese watersheds. *Journal of Hydrology*, (332):1–15, 2007.

[97] J. Zhou R. Zhang and, S. Ouyang, X. Wang, and H. Zhang. Optimal operation of multi-reservoir system by multi-elite guide particle swarm optimization. *Electrical Power and Energy Systems*, (48):1535–1542, 2013.

[98] M. Janga Reddy and D. Nagesh Kumar. Optimal reservoir operation using multi-objective evolutionary algorithm. *Springer*, 2006.

[99] REN21. Renewable energy policy network for the 21st century. *Renewables 2015 Global Status Report, GSR series*, 2015.

[100] J. Robinson. Sustainable land management in the qaroun watershed. *UNDP, Lebanese Ministry of the Environment*, 2013.

[101] A.M. Rushworth, A.W. Bowman, M.J. Brewer, and S.J. Langan. Distributed lag models for hydrological data. *Biometrics*, (69):537–544, 2013.

[102] B. Saravanan, S. Das, S. Sikri, and Dp. Kothari. A solution to the unit commitment problem – a review. *Electric Power Systems Research*, (2):223–236, 2013.

[103] Science Alert. The world's largest floating solar power plant is being built in japan. *https://www.sciencealert.com*, 2016.

[104] SENSIRON company. *Application Note Dew-point Calculation*. 2006.

[105] M. Silva, H.Morais, and Z. Vale. An integrated approach for distributed energy resource short-term scheduling in smart grids considering realistic power system simulation. *Energy Conversion and Management*, (64):273–288, 2012.

[106] John M. Staatz. Farmers' incentives to take collective action via cooperatives: A transaction-cost approach. *Cooperative Theory: New Approaches, USDA ACS Service Report*, pages 87–107, 1987.

[107] Stat Trek. *Coefficient of Determination definition*. 2017.

[108] K. Sudheer, A. Gosain, and K. Ramasastri. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrol. Process*, (16):1325–1330, 2002.

[109] Chun tian Chenga, Sheng li Liaoa, Zi-Tian Tanga, and Ming yan Zhaob. Comparison of particle swarm optimization and dynamic programming for large scale hydro unit load dispatch. *Energy Conversion and Management*, (50):3007–3014, 2009.

[110] CL. Tseng. On power system generation unit commitment problems. *University of California at Berkeley*, 1996.

[111] J. Valiantzas. Simplified versions for penman evaporation equation using routine weather data. *Journal of Hydrology*, (331):690–702, 2006.

[112] J. Valiantzas. Simplified forms for the standardized fao-56 penman monteith reference evapotranspiration using limited weather data. *Journal of Hydrology*, (505):12–23, 2013.

[113] J.D. Velasquez and V. Palade. Advanced techniques in web intelligence: Web user browsing behaviour and preference analysis. *Springer*, 2013.

[114] D. Verner, D. Lee, and M. Ashwill. Increasing resilience to climate change in the agricultural sector ofthe middle east: The cases of jordan and lebanon. 2013.

[115] Jorge Villavicencio, Rodolfo Ramirez, and S. Eduardo Pereira Bonvallet. Short term self scheduling of hydro power plants with intra-day regulation capacity. case study of los molles and sauzal in chile. *IEEE Power and Energy Society General Meeting*, pages 1–5, 2015.

[116] N. De Vos and T. Rientjes. Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation. *Hydrology and Earth System Sciences*, (9):111–126, 2005.

[117] Jiayang Wang, Shengli Liao, Chuntian Cheng, and Benxi Liu. Milp model for short-term hydro scheduling with head-sensitive prohibited operating zones. *World Environmental and Water Resources Congress*, 2017.

[118] Ping Wang, Hong Zhang, Zuodong Qin, and Guisheng Zhang. A novel hybrid-garch model based on arima and svm for pm2.5 concentrations forecasting. *Atmospheric Pollution Research*, (8):850–860, 2017.

[119] W. Wang, P.H.A.J.M. Van Gelder, J.K. Vrijling, and J. Ma. Forecasting daily streamflow using hybrid ann models. *Journal of Hydrology*, (324):383–399, 2006.

[120] Wen Wang. Stochasticity, nonlinearity and forecasting of streamflow processes. *IOS Press*, 2006.

[121] R. Wardlaw and J. Barnes. Optimal allocation of irrigation watersupplies in real time. *Journal of Irrigation and Drainage Engineering*, (125):345–354, 1999.

[122] William W.S. Wei. Time series analysis : Univariate and multivariate methods. *Pearson Addison Wesley*, 2006.

[123] World Energy Council. World energy resources. *Hydropower*, 2016.

[124] World Weather Online. *http://www.worldweatheronline.com/*. 2016.

[125] C.L. Wu, K.W. Chau, and Li Y.S. Methods to improve neural network performance in daily flows prediction. *Journal of Hydrology*, (69):537–544, 2009.

[126] C.L. Wu, K.W. Chau, and C. Fan. Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. *Journal of Hydrology*, (389): 146–167, 2012.

[127] Li Xue-Zhen, Xu Li-Zhong, and Chen Yan-Guo. Implicit stochastic optimization with data mining for reservoir system operation. *International Conference on Machine Learning and Cybernetics*, 2010.

[128] Linfeng Yang, Chen Zhang, Jinbao Jian, Ke Meng, Yan Xu, and Zhaoyang Dong. A novel projected two-binary-variable formulation for unit commitment in power systems. *Applied Energy*, (187):732–745, 2017.

[129] Tiantian Yang, Xiaogang Gao, Scott Lee Sellars, and Soroosh Sorooshian. Improving the multi-objective evolutionary optimization algorithm for hydropower reservoir operations in the california oroville–thermalito complex. *Environmental Modelling and Software*, (69):262–279, 2015.

[130] Vujica Yevjevich. Structure of daily hydrologic series. *Water Resources Pubns*, 1984.

[131] U. Yildiran, İ. Kayahan, M. Tunç, and S. Şisbot. Milp based short-term centralized and decentralized scheduling of a hydro-chain on kelkit river. *International Journal of Electrical Power and Energy*, (69):1–8, 2015.

[132] Wu Xin Yu, Cheng Chun Tian, Shen Jian Jian, Luo Bin, Liao Sheng Li, and Li Gang. A multi-objective short term hydropower scheduling model for peak shaving. *Electrical Power and Energy Systems*, (68):278–293, 2015.

[133] L. Zadeh. Optimality and non-scalar-valued performance criteria. *IEEE Transactions on Automatic Control*, (8):59–60, 1963.

[134] L.A. Zadeh. Fuzzy sets. *Information and Control*, (8):338–353, 1965.

[135] B. Zemadim, M. McCartney, and B. Sharma S. Langan. A participatory approach for hydrometeorological monitoring in the blue nile river basin of ethiopia. *IWMI*, 2014.

[136] G. Zhang. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, (50):159–175, 2003.

[137] Xueqing Zhang, Xiang Yu, and Hui Qin. Optimal operation of multi-reservoir hydropower systems using enhanced comprehensive learning particle swarm optimization. *Journal of Hydro-environment Research*, (10):50–63, 2016.

[138] Jizhong Zhu. *Optimization of Power System Operation*. Wiley, 2009.

Appendix A

Poster Presentations

The following two posters are the result of two papers that are addressed below:

1. **An Optimal Multi-Crop Planning Approach Implemented Under Deficit Irrigation**
Abstract: Multi-Crop Planning optimization model for cropping pattern and water allocation is introduced as a nonlinear programming problem. Its solution promotes an efficient use of water with a flexibility to keep the chosen crops at either full or deficit irrigation throughout different stages so that the net financial return is maximized within certain production bounds and resources constraints. The problem-solution approach is as follows: at first a preliminary mathematical tools are presented involving existence, benchmark linear models and a relaxation formulation, second two meta-heuristic algorithms Simulated Annealing (SA) and Particle Swarm Optimization (PSO) are implemented as a numerical technique for solving the MCP problem. The particularity of our approach consists of using the solution of a linear problem as an initial guess for the SA, while for PSO the particle swarm is initiated in the neighborhood of that solution.
2. **Daily River Flow Prediction Coupled with Data Processing Techniques: A Comparative Study between Constructive Fuzzy Systems and Autoregressive Models**
Abstract: Daily river flow forecast is an essential step for real-time hydro-power reservoir operation. Its purpose is to asset the decision making process of determining water storage in the reservoir in order to ensure optimal and reliable operational policy. The paper aim is seeking, in a region where meteorological and hydrological data are insufficient, inaccessible and sometimes unreliable, a data driven model based on Constructive Fuzzy Systems and capable of extracting the foremost from the available data with high prediction efficiency relative to an Autoregressive method. A case study was applied to Litani River in the Bekaa Valley - Lebanon using 4 years of rainfall, temperature and river flow daily measurements. A reference Autoregressive (AR) model, a classical Constructive Fuzzy System Modeling (C-FSM) and the Constructive Fuzzy System Modeling coupled with Moving Average (C-FSM_MA) are trained. Upon testing, the last two models have shown primarily competitive performance and accuracy with the ability of preserving the day-to-day variability up to 12 days ahead prediction task.
The first poster was presented at Melecon, 2016, whereas the second one was exhibited at JFL3 Troisièmes journées franco-libanaises, 2015.

Optimal Multi-Crop Planning Implemented Under Deficit Irrigation

Bassam BOU-FAKHREDDINE, Sara ABOU-CHAKRA, Imad MOUGHARBEL, Yann POLLET and Alain FAYE

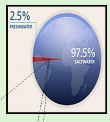
Joint Work Between Doctoral School Of Science and Technology (EDST) - Lebanese University and Ecole Doctorale d'Informatique, Télécommunications et Electronique (EDITE) - Conservatoire National des Arts et Métiers

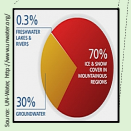
Abstract

Multi-Crop planning (MCP) optimization model for cropping pattern and water allocation is introduced as a nonlinear programming problem. Its solution promotes an efficient use of water with a flexibility to keep the chosen crops at either full or deficit irrigation throughout different stages.
The net financial return is maximized within certain resources constraints as well as **production bounds** .

Introduction

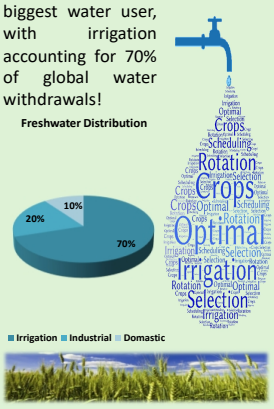
- Total volume of water on Earth is about 1.4 billion Km³





- Volume of freshwater resources is around 35 million Km³.
- About 105000 Km³ of freshwater can be accessed by humans directly from the surface.

Agriculture is the biggest water user, with irrigation accounting for 70% of global water withdrawals!



Problematic

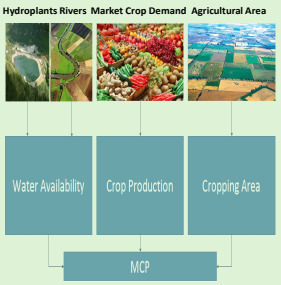
Multi-crop planning model

The aim of the MCP is to obtain the optimal farm plan for allocating irrigation areas in a multi-cropping system described in 1 and 2:

1- Objective functions

Maximize net profit of the produced yield over the planning horizon.

2- Constraints

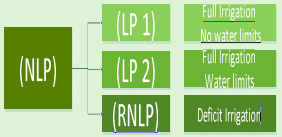


Nonlinear Programming Problem (NLP)

Methodology

Theoretical based approach

Two linear formulations (LP 1) , (LP 2) and a relaxed (RNLP) version were extracted from the (NLP) model:



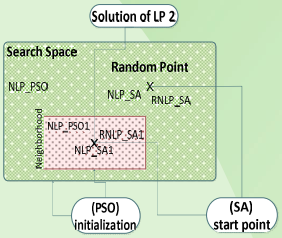
Numerical approach

Two meta-heuristic algorithms:

- Simulated Annealing (SA) and
- Particle Swarm Optimization (PSO)

Algorithms Initialization Scenarios

- o Random
- o Using LP 2 solution



Results

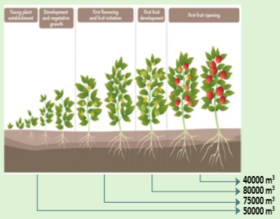
Numerical Example

An experimental evaluation for each algorithm aims:

- Optimal cropping pattern and
- Irrigation scheduling

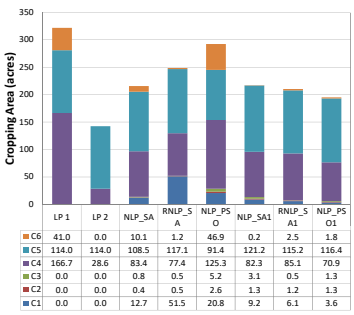
Given :

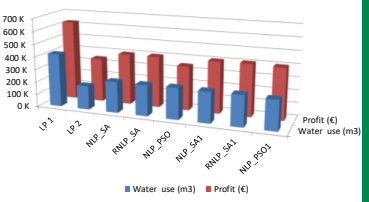
- o A lot of six crops (C1, C2,…)
- o Max spread area 322 acres
- o Crop Demand
- o Water availability at every stage
- o Available water 245000 m³



Simulations

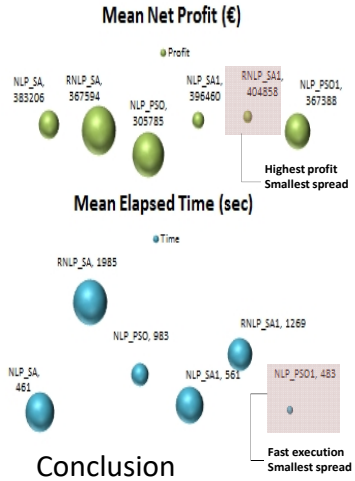
The programs were coded in MATLAB language and ran on Intel Core i7-5500U CPU @ 2.40 GHZ, 12.0 GB RAM.
The following figures show the recommended optimum crop pattern, consumed water and the net profit for each crop:





Performance and Reliability of each Scenario

Algorithms are run for several times with different initialization methods. Tests carried with the aid of Coefficient of Variation (ball size):



Conclusion

The computational results lead us to consider, in the future work, the real capabilities of the suggested scenarios.
They will be implemented with real data obtained from the Bekaa-Valley region near Qaraoun reservoir- Lebanon.

Daily River Flow Prediction Coupled with Data Processing Techniques: A Comparative Study between Constructive Fuzzy Systems and Autoregressive Models

Bassam BOU-FAKHREDDINE, Sara ABOU-CHAKRA, Imad MOUGHARBEL, Yann POLLET and Alain FAYE

Joint Work Between Doctoral School Of Science and Technology (EDST) - Lebanese University and Ecole Doctorale d'Informatique, Télécommunications et Electronique (EDITE) - Conservatoire National des Arts et Métiers

ABSTRACT

Daily river flow forecast is an essential step for real-time hydro-power reservoir operation. Its purpose is to asset the decision-making process of determining water storage in the reservoir in the course of ensuring optimal and reliable operational policy. Our aim is seeking, in case of meteorological and hydrological data limitation, a data driven model based on Constructive Fuzzy Systems that is capable of extracting the foremost from the accessible data with high prediction efficiency relative to an Autoregressive method.

A case study was applied to Litani River in the Bekaa valley - Lebanon using 4 years of rainfall, temperature and river flow daily measurements.

A reference Autoregressive model, a classical Constructive Fuzzy System Modeling and Constructive Fuzzy System coupled with moving average are trained. Upon validation, the last two models have shown primarily a competitive performance and accuracy for a multi-step ahead prediction task.

Overview

Framework for River flow Forecasting

Gathering data	• Rainfall • Temperature • River flow	• Machgara weather station • Joub Jannin Gauging station
Data Organization	• Distribute data over 3 classes	• Training, Calibration, Validation
Pre-Modeling data processing	• Input selection • Unsupervised data classification • Trends, periodicity • Noise filtering	• Correlation analysis • Subtractive clustering • Standardization • Moving average
Model Formulation	• Linear • Non-Linear	• Autoregressive Modeling • Constructive Fuzzy Modeling
Training	• Autoregressive model • Constructive Fuzzy model	• Yule-Walker equation • Expectation Maximization algorithm
Calibration	• Cluster radius	• Best efficiency
Validation	• Testing model	•
Predication	• Decision making	• Flood control • Irrigation • Hydropower

BASIC MODELS

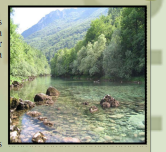
1. Autoregressive Model

An autoregressive model defines the next random variable in a sequence as an explicit linear function of previous ones within a time frame.

Input
- Previous river flow

Data processing
- Standardization
- Correlation analysis

Output
- River flow up to 12 days ahead



2. Constructive Fuzzy Modeling

Typically, Multiple Input Single Out-put (MISO) model structure that is based on first order Takagi Sugeno fuzzy system is composed by a set of M fuzzy rules. Its representative power is manifested through its capability of describing a highly complex nonlinear system using a small number of simple rules.

Input
- Rainfall
- Temperature
- Previous river flow

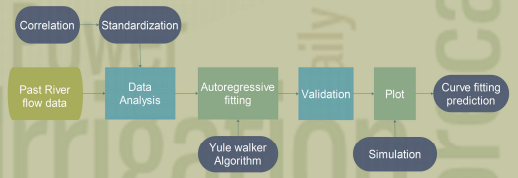
Data processing
- Standardization
- Normalization
- Data transformation
- Correlation analysis
- Moving average

Output
- River flow up to 12 days ahead

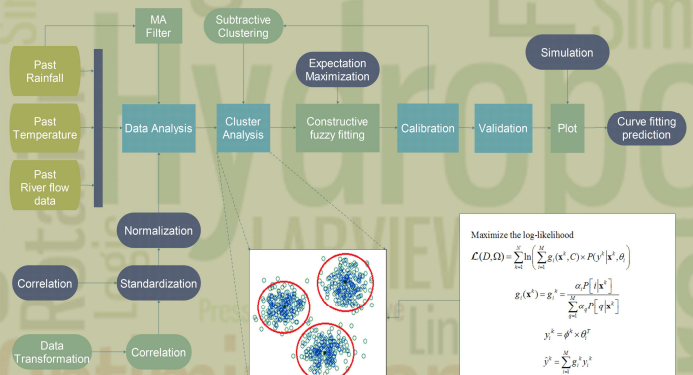
MODELING

Daily River Flow Forecasting Models

Autoregressive Modeling (AR)



Constructive Fuzzy Systems Modeling (C-FSM)



Maximize the log-likelihood

$$\mathcal{L}(D,\Omega) = \sum_{k=1}^{N} \ln\left(\sum_{i=1}^{M} g_i(\mathbf{x}^k, C) \times P(y^k|\mathbf{x}^k, \theta_i)\right)$$

$$g_i(\mathbf{x}^k) = g_i^k = \frac{\alpha_i P[\,|\mathbf{x}^k]}{\sum_{j=1}^{M} \alpha_q P[\,q|\mathbf{x}^k]}$$

$$y_i^k = \phi^k \times \theta_i^T$$

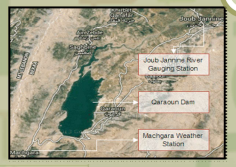
$$\hat{y}^k = \sum_{i=1}^{M} g_i^k y_i^k$$

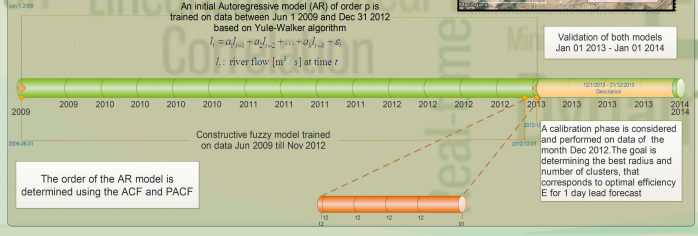
SIMULATIONS

Litani River Flow Estimate

Case study: Litani River

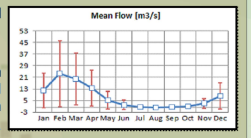
The Litani is the largest river in Lebanon in length reaching 170 Km. Its watershed covers an area of 2160 km^2 that is fed with rainfall average around 700 mm/year, or about 764 Mm^3.



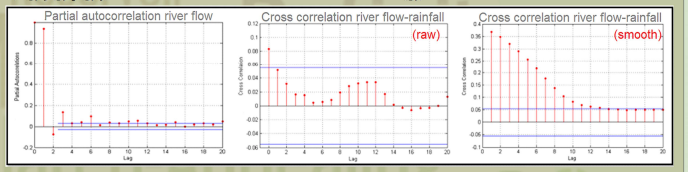


Data Processing

Investigation the relationships between hydrological and meteorological data collected from Litani Basin.



Input Selection: The input – output pattern is based on correlation analysis. Inputs are past and present observations for rainfall $(...,P_{t-2},P_{t-1},P_t)$, temperature $(...,T_{t-2},T_{t-1},T_t)$ and past river flow data $(...,Q_{t-3},Q_{t-2},Q_{t-1})$. The output is the present river flow (Q_t).

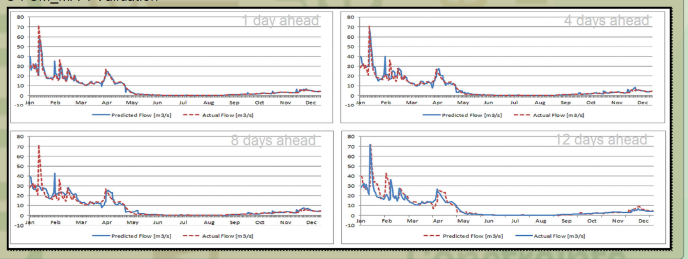


Results

Models performance and validation was conducted considering the Root Mean Square Error (RMSE-[m^3/s]), Mean Absolute Error (MAE-[m^3/s]) and the mass curve coefficient (E).

Model	Performance index	Horizon [h] 1	2	3	4	5	6	7	8	9	10	11	12
AR(7)	RMSE [m^3/s]	3.454	3.700	4.075	4.192	5.003	5.240	4.811	4.633	4.862	5.094	5.034	6.668
C-FSM 1		3.377	4.276	4.623	5.206	4.687	4.950	4.523	5.099	4.564	4.912	4.640	4.924
C-FSM_MA 1		3.349	3.534	3.868	3.763	4.987	4.491	4.637	4.373	4.640	4.754	4.827	5.774
C-FSM 2		3.346	4.283	4.615	5.184	4.510	4.985	4.777	5.040	4.708	5.060	4.822	5.125
C-FSM_MA 2		3.301	3.486	3.816	3.698	4.899	4.625	4.213	4.181	4.644	4.913	4.948	4.960
C-FSM 3		3.324	3.413	3.717	3.613	4.941	4.183	4.455	4.471	4.995	5.063	5.119	4.647
C-FSM_MA 3		3.344	3.496	3.823	3.722	4.891	4.351	4.448	4.281	4.807	4.885	4.979	5.127
C-FSM 4		3.352	4.178	3.691	3.527	5.020	4.371	4.831	4.585	5.360	5.297	5.354	4.715
C-FSM_MA 4		3.348	3.433	3.632	3.474	4.885	3.757	4.203	4.203	4.807	4.798	5.054	4.350
AR(7)	MAE [m^3/s]	0.951	1.203	1.458	1.524	1.824	1.957	1.939	1.961	1.916	2.090	2.046	2.695
C-FSM 1		0.924	1.260	1.438	1.743	1.759	1.786	1.743	2.160	1.873	2.048	1.745	1.801
C-FSM_MA 1		0.917	1.156	1.369	1.403	1.751	1.760	1.804	1.826	1.821	1.880	1.887	2.412
C-FSM 2		0.928	1.291	1.409	1.762	1.727	1.833	1.923	2.213	1.970	2.240	1.950	2.136
C-FSM_MA 2		0.907	1.138	1.347	1.379	1.767	1.764	1.658	1.722	1.804	2.027	1.954	2.227
C-FSM 3		0.842	1.031	1.241	1.285	1.715	1.540	1.709	1.715	1.818	1.965	1.887	1.943
C-FSM_MA 3		0.884	1.110	1.326	1.368	1.734	1.673	1.753	1.746	1.812	1.900	1.867	2.168
C-FSM 4		0.841	1.229	1.289	1.301	1.783	1.620	1.834	1.799	2.053	2.209	2.198	1.973
C-FSM_MA 4		0.898	1.122	1.264	1.311	1.788	1.482	1.725	1.705	1.800	1.831	1.922	1.870
AR(7)	E	0.878	0.860	0.830	0.820	0.743	0.719	0.763	0.780	0.758	0.734	0.740	0.544
C-FSM 1		0.888	0.820	0.791	0.734	0.784	0.762	0.799	0.747	0.797	0.765	0.789	0.764
C-FSM_MA 1		0.885	0.872	0.847	0.855	0.745	0.796	0.780	0.806	0.782	0.771	0.762	0.663
C-FSM 2		0.890	0.820	0.791	0.736	0.800	0.758	0.776	0.753	0.784	0.751	0.772	0.744
C-FSM_MA 2		0.888	0.876	0.852	0.860	0.754	0.783	0.819	0.823	0.781	0.755	0.750	0.751
C-FSM 3		0.887	0.881	0.858	0.866	0.750	0.821	0.797	0.795	0.744	0.737	0.731	0.779
C-FSM_MA 3		0.885	0.875	0.850	0.858	0.755	0.806	0.797	0.812	0.763	0.755	0.746	0.731
C-FSM 4		0.885	0.821	0.860	0.873	0.742	0.804	0.761	0.785	0.706	0.712	0.706	0.772
C-FSM_MA 4		0.885	0.879	0.865	0.876	0.755	0.855	0.819	0.819	0.763	0.764	0.738	0.806

C-FSM_MA 4 Validation



CONCLUSION

A comparative study was carried between the AR, C-FSM and C-FSM coupled with MA methods for multi-step ahead daily river flow time series forecasting, where models capabilities were tested on Litani river.

Despite the reported illegal activities along the river such as domestic and industrial discharges, water diversion and garbage dumping , the results of suggested models came very satisfactory. In fact, their performances were too promising in the context of river flow prediction especially when compared with the latest studies carried on Litani river.

Bassam Bou-Fakhreddine

Modeling, Control and Optimization of Cascade Hydroelectric-Irrigation Plants Operation and Planning

Résumé

Ce travail de recherche vise à optimiser la procédure opérationnelle des centrales hydroélectriques en cascade afin de les utiliser efficacement pour la production d'électricité et l'irrigation. Le défi consistait à trouver le modèle le plus réaliste basé sur la caractéristique stochastique des ressources en eau, sur la demande en énergie et sur le profil d'irrigation. Tous ces aspects sont affectés à court et à long terme par un large éventail de conditions différentes (hydrologique, météorologique et hydraulique). Au cours de ce projet, une étude bibliographique a été réalisée afin d'identifier les problèmes techniques qui empêchent l'utilisation efficace des centrales hydroélectriques dans les pays en développement. Le système est modélisé numériquement en tenant compte de toutes les variables et paramètres impliqués dans le fonctionnement optimal. L'approche la plus appropriée est choisie afin de maximiser l'utilisation efficace de l'eau et de minimiser les pertes économiques, où différents scénarios sont simulés afin de valider les solutions adoptées.

Mots-clés : Hydroélectricité, Irrigation, Optimisation, Modélisation, Opération, Planification, Data Mining.

Résumé en anglais

This research work aims to optimize the operational procedure of cascade hydro plants in order to be efficiently used for power generation and irrigation. The challenge was to find the most realistic model based on the stochastic feature of water resources, on the power demand and on the irrigation profile. All these aspects are affected on the short and on the long run by a wide range of different conditions (hydrological, meteorological and hydraulic). During this project a bibliographic study was done in order to identify the technical issues that prevent the efficient use of hydro plants in developing countries. The system is numerically modelled taking into consideration all the variables and parameters involved in the optimal operation. The most appropriate approach is chosen in order to maximize the efficient use of water and to minimize economical losses, where different scenarios are simulated in order to validate the adopted suggestions.

Keywords: Hydroelectricity, Irrigation, Optimization, Modeling, Operation, Planning, Data Mining.