Fuzzy Interval Numbers (*FIN*s): Lattice Theoretic Tools for Improving Prediction of Sugar Production from Populations of Measurements[1]

Vassilis G. Kaburlasos
Dept. of Industrial Informatics, Div. of Computing Systems
Technological Educational Institution of Kavala
GR-65404 Kavala, Greece

***Abstract* – This work presents novel mathematical tools developed during a study of an industrial-yield prediction problem. The set F of *Fuzzy Interval Numbers*, or *FIN*s for short, is studied in the framework of lattice theory. A *FIN* is defined as a mapping to a metric lattice of *generalized intervals*, moreover it is shown analytically that the set F of *FIN*s is a metric lattice. A *FIN* can be interpreted as a convex fuzzy set, moreover a statistical interpretation is proposed here. Algorithm CALFIN is presented for constructing a *FIN* from a population of samples. An *underlying* positive valuation function implies both a metric distance and an inclusion measure function in the set F of *FIN*s. Substantial advantages, both theoretical and practical, are shown. Several examples illustrate geometrically on the plane both the utility and the effectiveness of novel tools. It is outlined comparatively how some of the proposed tools have been employed for improving prediction of sugar production from populations of measurements for Hellenic Sugar Industry, Greece.**

***Index Terms* – Fuzzy Interval Numbers (FINs), Lattice Theory, Prediction Models, Sugar Production.**

I. INTRODUCTION

The amount of sugar required for the needs of the Greek market is supplied, at large, by the production of Hellenic Sugar Industry (HSI). An early season accurate prediction of sugar production is critical for planning effectively the annual sugar-beet campaign. A detailed problem description in the HSI domain and extensive experimental results using various prediction methods have been reported in the literature. More specifically, [27] reports prediction results using both intelligent-clustering and first-principles modelling techniques; furthermore the work in [42] shows results by a probabilistically motivated predictor, namely Bayesian Combined Predictor (BCP), whose local predictors include linear regressors, artificial neural networks, polynomial predictors, etc. Lately, the application of a nearest neighbour classifier has been presented for prediction of sugar production based on *Fuzzy Interval Numbers*, or *FIN*s for short [41]. While the emphasis in [41] is primarily on an application of *FIN*s in the HSI domain, this work details an analytic study of *FIN*s. In addition, a few experimental results are briefly presented here comparatively.

[1] This work has been supported in part by Hellenic Sugar Industry, Greece from July 2000 to June 2001.

A *FIN* is presented here as a "mathematical object", which can be interpreted as a convex fuzzy set; additional interpretations are shown including a statistical interpretation. In the context of this work a *FIN* is computed from a population of samples using algorithm CALFIN. A *FIN* is represented as a set of *generalized intervals*. Rigorous analysis is carried out based on lattice theory [3], [10], [19]. It is shown that the set F of *FIN*s is a metric (mathematical) lattice. An *underlying* positive valuation function implies both a metric distance and an inclusion measure function in F. A practical advantage of the techniques introduced here is that a much larger number of metric distances as well as inclusion measure functions between *FIN*s can be defined, both theoretically and practically, for tuning performance in an application.

A potentially important outcome of this work is that the mathematical tools presented here could be effective for general fuzzy system design. It is remarkable that even though the *lattice ordering* relation of fuzzy sets was acknowledged since the introduction of fuzzy set theory [58], as well as during its later development [59], the ordering relation in question has not been taken advantage of in practice. For instance, in [32] the ordering of *fuzzy numbers* is addressed using a user-defined satisfaction function, whereas in [57] a method is proposed for ranking *fuzzy numbers* based on signed distances. This work suggests a useful alternative for dealing with fuzzy numbers.

The layout of this paper is as follows. Section II details metric lattices of generalized intervals. Section III shows both metric distances and inclusion measure functions in the set of *FIN*s. Section IV presents algorithm CALFIN for constructing a *FIN* from a population of measurements, moreover a statistical interpretation of a *FIN* is presented. Section V shows comparatively experimental results for prediction of sugar production based on *FIN*s. Section VI concludes by summarizing the contribution of this work and discussing future work. Mathematical proofs of useful theoretical results are included in the Appendix.

II. LATTICES M^h OF GENERALIZED INTERVALS

The notion "interval" has been used for both analysis and design in mathematics as well as in engineering [1], [4], [29], [33], [35], [36], [49]. The notion *generalized interval* has been introduced in regression and modelling problems [22], [24], [41]. A *generalized interval* is defined in the following.

Definition 1: A *generalized interval of height h* is a mapping given by

1) If $x_1<x_2$ then $\quad \mu_{[x_1,x_2]^h}(x)=\begin{cases} h, & x_1 \le x \le x_2 \\ 0, & otherwise \end{cases}$, else

2) if $x_1>x_2$ then $\quad \mu_{[x_1,x_2]^h}(x)=\begin{cases} -h, & x_1 \ge x \ge x_2 \\ 0, & otherwise \end{cases}$, else

3) $\quad \mu_{[x_1,x_1]^h}(x)=\begin{cases} \{-h,h\}, & x=x_1 \\ 0, & otherwise \end{cases}$,

where $h\in(0,1]$.

An interpretation of a generalized interval depends on an application; for instance a positive generalized interval could indicate the *presence*, whereas a negative generalized interval could indicate the *absence* of certain features. No specific interpretation is necessary for the theoretical analysis in this section.

The above definition implies that $\mu_{[x_1,x_2]^h}(x)$, for $x_1 \neq x_2$, is a function of x; nevertheless, $\mu_{[x_1,x_1]^h}(x)$ is not a function because $x=x_1$ is mapped to both values $-h$ and h. A generalized interval will be denoted, more compactly, as $[x_1,x_2]^h$. More specifically, if $x_1<x_2$ ($x_1>x_2$) then $[x_1,x_2]^h$ is called *positive* (*negative*) *generalized interval*; furthermore $[x_1,x_1]^h$ is called *trivial generalized interval*. The set of positive (negative) generalized intervals of height h will be denoted by M_+^h (M_-^h); furthermore the corresponding set of trivial generalized intervals will be denoted by M_0^h. The set of generalized intervals of height h will be denoted by M^h; it follows $\mathsf{M}^h = \mathsf{M}_-^h \cup \mathsf{M}_0^h \cup \mathsf{M}_+^h$, where M_-^h, M_0^h and M_+^h are pairwise disjoint. The set-union of the M^hs, $h\in(0,1]$ is the set M of *generalized intervals*, symbolically $\mathsf{M}=\bigcup_{h\in(0,1]}\mathsf{M}^h$. Likewise the symbols $\mathsf{M}_-=\bigcup_{h\in(0,1]}\mathsf{M}_-^h$, $\mathsf{M}_0=\bigcup_{h\in(0,1]}\mathsf{M}_0^h$ and $\mathsf{M}_+=\bigcup_{h\in(0,1]}\mathsf{M}_+^h$ denote the sets of negative, trivial and positive generalized intervals, respectively. A couple of useful functions regarding generalized intervals are defined in the following.

The *support* (of a generalized interval) is a function which maps a generalized interval to its conventional interval support set; in particular $support([x_1,x_2]^h)=[x_1,x_2]$ for positive generalized intervals, $support([x_1,x_2]^h)=[x_2,x_1]$ for negative generalized intervals, whereas $support([x_1,x_1]^h)=\{x_1\}$ for trivial generalized intervals. Furthermore, function *sign* maps a positive generalized interval to number +1, a negative generalized interval to number –1, moreover it maps a trivial generalized interval to number 0.

A. Metric Distances and Inclusion Measure Functions in Lattice M^h

An *ordering* relation in the set M^h, $h\in(0,1]$ of generalized intervals is introduced in the following.

(OR1) $[a,b]^h \leq_{\mathsf{M}^h} [c,d]^h \Leftrightarrow support([a,b]^h) \subseteq support([c,d]^h)$, for both $[a,b]^h$ and $[c,d]^h$ in M_+^h,

(OR2) $[a,b]^h \leq_{\mathsf{M}^h} [c,d]^h \Leftrightarrow support([c,d]^h) \subseteq support([a,b]^h)$, for both $[a,b]^h$ and $[c,d]^h$ in M_-^h, and

(OR3) $[a,b]^h \leq_{\mathsf{M}^h} [c,d]^h \Leftrightarrow support([a,b]^h) \cap support([c,d]^h) \neq \varnothing$, for $[a,b]^h$ in M_-^h and $[c,d]^h$ in M_+^h.

We remark that the ordering relation $\leq_{\mathsf{M}^h}$ also accommodates *trivial generalized intervals*. The subscript in symbol "$\leq_{\mathsf{M}^h}$" was meant to explicitly identify the underlying set M^h. Note that when symbol $\leq$ is employed here without a subscript it refers to the total ordering relation in the lattice R of real numbers.

It can be easily shown that the ordering relation $\leq_{M^h}$ is a *partial ordering* relation[2]. Note that partial ordering relation $\leq_{M^h}$ does not hold for all pairs of generalized intervals. If neither $r_1 \leq_{M^h} r_2$ nor $r_2 \leq_{M^h} r_1$ hold for $r_1, r_2 \in \mathsf{M}^h$, then r_1 and r_2 are called *incomparable generalized intervals*, symbolically $r_1 \|_{M^h} r_2$ (Fig.1). It is understood that if $r_1 \in \mathsf{M}^{h_1}$ and $r_2 \in \mathsf{M}^{h_2}$ (with $h_1 \neq h_2$) then generalized intervals r_1 and r_2 are *incomparable* because they are elements of different partially ordered sets. The following result fully characterizes the partial ordering relation $\leq_{M^h}$ in M^h [22], [41].

Proposition 2: The partially ordered set M^h of generalized intervals of height h is a (mathematical) lattice[3].

For any two generalized intervals in M^h there exist both a greatest lower bound (g.l.b.) and a least upper bound (l.u.b.) in M^h [22], [24], [41]. Fig.2 enumerates exhaustively all cases which could be encountered in the computation of the *lattice join* ($r_1 \vee_{M^h} r_2$) and the *lattice meet* ($r_1 \wedge_{M^h} r_2$) for two generalized intervals r_1 and r_2. We remark that M^h is not a complete lattice because the lattice R of real numbers is not a complete lattice [10]. Moreover for $x,y \in \mathsf{R}$ it holds $x \wedge_{\mathsf{R}} y = x \wedge y = min\{x,y\}$, and $x \vee_{\mathsf{R}} y = x \vee y = max\{x,y\}$; note also that the elements of lattice R are *totally ordered*, that is for a pair or real numbers x,y it is either $x \leq y$ or $y \leq x$. A useful function in a general lattice L is shown next.

Definition 3: A *valuation* in a lattice L is a real function v: $\mathsf{L} \rightarrow \mathsf{R}$ which satisfies $v(x)+v(y)=v(x \vee_{\mathsf{L}} y)+v(x \wedge_{\mathsf{L}} y)$, $x,y \in \mathsf{L}$. A valuation is called *monotone* if and only if $x \leq_{\mathsf{L}} y$ implies $v(x) \leq v(y)$, and *positive* if and only if $x <_{\mathsf{L}} y$ implies $v(x) < v(y)$ for $x,y \in \mathsf{L}$.

A positive valuation function $v(.)$ in a lattice L implies a *metric distance*[4] d: $\mathsf{L} \times \mathsf{L} \rightarrow \mathsf{R}$ given by $d(x,y)=v(x \vee_{\mathsf{L}} y)-v(x \wedge_{\mathsf{L}} y)$ for $x,y \in \mathsf{L}$ [3]. Furthermore, it is well-known that a positive valuation function $v(.)$ in a complete lattice L implies an *inclusion measure function* k: $\mathsf{L} \times \mathsf{L} \rightarrow [0,1]$ given by $k(x,u)= \dfrac{v(u)}{v(x \vee_{\mathsf{L}} u)}$, where the latter function quantifies a degree of inclusion of a lattice element x into another one u [24], [25], [38], [39]. Since a lattice M^h of generalized intervals is not a complete lattice the following definition is proposed here for an inclusion measure.

---

[2] A *partial ordering* relation, symbolically $\leq_{\mathcal{S}}$, in a set $\mathcal{S}$ has to be: (PO1) *reflexive* ($x \leq_{\mathcal{S}} x$), (PO2) *antisymmetric* ($x \leq_{\mathcal{S}} y$ and $y \leq_{\mathcal{S}} x$ imply $x=y$), and (PO3) *transitive* ($x \leq_{\mathcal{S}} y$ and $y \leq_{\mathcal{S}} z$ imply $x \leq_{\mathcal{S}} z$), where $x,y,z \in \mathcal{S}$.

[3] A *lattice* is a partially ordered set L any two of whose elements have a greatest lower bound (g.l.b.) or "meet" denoted by $x \wedge_{\mathsf{L}} y$, and a least upper bound (l.u.b.) or "join" denoted by $x \vee_{\mathsf{L}} y$. A lattice L is called *complete* when each of its subsets X has a l.u.b. and a g.l.b in L.

[4] A *metric distance* in a set $\mathcal{S}$ is a real function d: $\mathcal{S} \times \mathcal{S} \rightarrow \mathsf{R}$ which satisfies: (MD1) $d(x,y) \geq 0$, $x,y \in \mathcal{S}$, (MD2) $d(x,y)=0 \Leftrightarrow x=y$, $x \in \mathcal{S}$, (MD3) $d(x,y) = d(y,x)$, $x,y \in \mathcal{S}$, - *Symmetry*, (MD4) $d(x,y) \leq d(x,z) + d(z,y)$, $x,y,z \in \mathcal{S}$ - *Triangle Inequality*.

Definition 4: Let L be a non-complete lattice. An *inclusion measure* σ in L is a map σ: $\mathsf{L}\times\mathsf{L}\rightarrow[0,1]$ such that for $u,w,x\in\mathsf{L}$ the following conditions are satisfied

(IM1) $\sigma(x,x) = 1$,

(IM2) $u <_{\mathsf{L}} w \Rightarrow \sigma(w,u) < 1$, and

(IM3) $u \leq_{\mathsf{L}} w \Rightarrow \sigma(x,u) \leq \sigma(x,w)$ - Consistency Property.

In can be argued that $\sigma(x,u)$ indicates a degree of inclusion of x in u [25], [39]. Therefore, from henceforward the notations $\sigma(x,u)$ and $\sigma(x\leq_{\mathsf{L}}u)$ will be used interchangeably.

For the totally-ordered lattice R of real numbers, in particular, any strictly increasing function $f(.)$ is a positive valuation function, which can be used for introducing a positive valuation function in a lattice M^h of generalized intervals as detailed in the following.

Proposition 5: Let f: $\mathsf{R}\rightarrow\mathsf{R}$ be a strictly increasing real function in R, namely *underlying positive valuation function*. Then the real function v: $\mathsf{M}^h\rightarrow\mathsf{R}$ given by $v([a,b]^h)=sign([a,b]^h)c(h)\int_a^b [f(x)-f(a)]dx$ is a *positive valuation* function in M^h, where function *sign*: $\mathsf{M}^h\rightarrow\{-1, 0, +1\}$ has been defined above, c: $(0,1]\rightarrow\mathsf{R}^+$ is a positive real function, and $\int(.)$ is the conventional integral operator.

We remark that the positive multiplicative coefficient $c(h)$ has been inserted above for normalization; for instance it could be $c(h)=h$ which indicates that, all other things being equal, a larger value of $v(.)$ is assigned to a generalized interval of a larger height. Proposition 5 can be proved easily by considering exhaustively all the cases shown in Fig.2 so as to confirm that function v satisfies the conditions for a positive valuation function of definition 3.

Given a positive valuation function in M^h there follows a metric distance in M^h given by $d_{\mathrm{h}}(x,y)=v(x\vee_{\mathsf{M}^h} y)-v(x\wedge_{\mathsf{M}^h} y)$ as it has been shown above. As a special application consider $f(x)=x$ and $c(h)=h$ in proposition 5; it follows that the distance between two generalized intervals $[a,b]^h$ and $[c,d]^h$ equals $d_{\mathrm{h}}([a,b]^h,[c,d]^h)= h(|a-c|+|b-d|)$. Furthermore, a metric can be defined in the space $\mathsf{M}=\bigcup_{h\in(0,1]}\mathsf{M}^h$ as follows

$$d_{\mathsf{M}}([a,b]^{h_1},[c,d]^{h_2})= \frac{1}{2}[d_{\mathrm{h}}([a,b]^{h_1},[c,d]^{h_1}) + d_{\mathrm{h}}([a,b]^{h_2},[c,d]^{h_2})] \qquad (1)$$

It is straightforward to show that the aforementioned function d_{M} satisfies conditions (MD1)–(MD4) of footnote 4; note that for $h_1=h_2$ it follows $d_{\mathsf{M}}\equiv d_{\mathrm{h}}$. An inclusion measure is introduced in the following in the

lattice $\mathsf{M}^h\backslash\mathsf{M}_-^h$ of non-negative generalized intervals of height h. More specifically, it is straightforward to show that the following function

$$k(x,u)= \frac{v(u)}{v(x \vee_{\mathsf{M}^h} u)} \tag{2}$$

satisfies conditions (IM1)-(IM3) of definition 4, therefore $k(x,u)$ is an inclusion measure function in the lattice $\mathsf{M}^h\backslash\mathsf{M}_-^h$ (in fact, the aforementioned proof follows along the proof of Theorem 5 in [39], the latter proof is given in a complete lattice). Note that lattice $\mathsf{M}^h\backslash\mathsf{M}_-^h$ is a non-complete one because the set R of real numbers is a non-complete (totally-ordered) lattice [10]. It follows that the lattice $\mathsf{M}^h\backslash\mathsf{M}_-^h$ of non-negative generalized intervals is a *fuzzy lattice* whose definition is shown in the following.

Definition 6: A *fuzzy lattice* is a pair $\langle \mathsf{L},\mu \rangle$ where L is a crisp lattice and $(\mathsf{L}\times\mathsf{L},\mu)$ is a fuzzy set such that $\mu(x,y)=1$ if and only if $x\leq_{\mathsf{L}} y$.

We remark that a fuzzy set is denoted in this work by a pair $(\mathtt{X},\mu)$, where $\mathtt{X}$ is the universe of discourse and μ is a fuzzy membership function μ: $\mathtt{X}\rightarrow[0,1]$. The collection of fuzzy lattices is referred to as *framework of fuzzy lattices* or *FL-framework* for short [25], [38], [39]. It follows that $\langle \mathsf{M}^h\backslash\mathsf{M}_-^h, k\rangle$ is a fuzzy lattice. The function $k(x\leq_{\mathsf{M}^h} u)$ quantifies a degree of inclusion of non-negative generalized interval x to the non-negative generalized interval u in a "fuzzy lattice sense". That is, given a positive valuation function $v(.)$ in M^h, ratio $k(x\leq_{\mathsf{M}^h} u)= \frac{v(u)}{v(x \vee_{\mathsf{M}^h} u)}$ expresses the size of interval u in terms of the size of the smallest interval which includes both intervals x and u, the latter size is considered as unit.

B. Remarks, Perspectives and Examples

In the definition of positive valuation function v: $\mathsf{M}^h\rightarrow\mathsf{R}$, in proposition 5, the calculation of the definite integral $\int_a^b [f(x)-f(a)]dx$ is required. Without loss of generality it can be assumed $f(0)=0$ because for any strictly monotone increasing, integrable real function $f_1(.)$ with $f_1(.)\neq 0$ another one $f_0(.)$ with $f_0(0)=0$ can always be defined from $f_1(.)$ by subtracting $f_1(0)$ from $f_1(x)$, $\forall x\in\mathsf{R}$. Functions $f_1(.)$ and $f_0(.)$ give identical results when either of them is used for computing the definite integral $\int_a^b [f_1(x)-f_1(a)]dx$ due to the linearity

property of the integral operator: $\int_a^b [f_1(x) - f_1(a)]dx = \int_a^b [\{f_1(x) - f_1(0)\} - \{f_1(a) - f_1(0)\}]dx = \int_a^b [f_0(x) - f_0(a)]dx$. In the sequel it is assumed $f(0)=0$ for all underlying positive valuations functions. The cardinality of all metric distances in a set M^h of generalized intervals is shown in the following paragraph.

Proposition 5 implies that there is a one-one correspondence between metric distances in M^h and strictly increasing functions in the set R of real numbers. Since there are $\aleph_1$ different strictly increasing functions in R, where $\aleph_1$ is the cardinality of the set R of real numbers [22]; there follow $\aleph_1$ different metric distances in M^h.

The essential role of a positive valuation function v: $\mathsf{L}\rightarrow\mathsf{R}$ is known to be a mapping from a lattice L of *semantics* to the mathematical field R of real numbers for carrying out computations [23]. Regarding, in particular, the domain R of an underlying positive valuation function $f(.)$ note that the aforementioned domain is treated here as a *totally ordered lattice*; that is, the operations *addition* and *multiplication* in the domain of $f(x)$ were of no concern. Nevertheless, the range of function $f(.)$ was treated as the conventional *mathematical field* [20] of real numbers where both the *addition* and the *multiplication* operations involving $f(x_1)$ and $f(x_2)$ have been employed, e.g. for computing the integral $\int_a^b [f(x) - f(a)]dx$ in proposition 5. The practical implication of the latter remark is that the tools presented here are applicable in any *totally-ordered* universe of discourse L provided an underlying positive valuation function in L.

From a mathematical viewpoint the set M^h of generalized intervals is still a metric lattice for $h>1$. Nevertheless, interest is focused in this work on metric lattices M^h with $h\in(0,1]$ because generalized intervals in the latter lattices can be interpreted as a-cuts of fuzzy sets as explained below. The following examples illustrate on the plane the calculation of distance $d(.,.)$ between generalized intervals in M^h.

Example 7: Consider two real numbers a and b ($\neq a$). Let $[a,a]^1$ and $[b,b]^1$ be the corresponding trivial generalized intervals in M_0^1 shown in Fig.3(a). Consider generalized intervals $[a,a]^1 \vee_{\mathsf{M}^1} [b,b]^1$ and $[a,a]^1 \wedge_{\mathsf{M}^1} [b,b]^1$ in Fig.3 (b) and (c), respectively. Using the positive valuation function $v(.)$ defined in proposition 5 with underlying positive valuation function $f(x)=x$, it follows $v([a,a]^1 \vee_{\mathsf{M}^1} [b,b]^1)=c(1)|a-b|$ furthermore $v([a,a]^1 \wedge_{\mathsf{M}^1} [b,b]^1)=-c(1)|a-b|$, where $|.|$ is the absolute value of its (real number) operand. In conclusion, $d_1([a,a]^1,[b,b]^1)=2c(1)|a-b|$. Choosing $c(1)=0.5$, equality $d([a,a]^1,[b,b]^1) = |a-b|$ is forced, the latter is the conventional distance between numbers a and b.

Using a different underlying positive valuation function than $f(x)=x$, the distance between two generalized intervals can change as demonstrated in the following example.

Example 8: Consider generalized intervals $[-1,1]^1$ and $[2,4]^1$. Fig.4(a) and Fig.4(b) show, respectively, two different underlying positive valuation functions in the set $\mathbb{R}$ of real numbers. On the one hand, a steeply increasing underlying positive valuation function f_1 is shown in Fig.4(a), that is in particular $f_1(x)=x^3$. The computation of positive valuation function $v(.)$ with $f=f_1$ and $c(1)=0.5$, results in $d([-1,1]^1,[2,4]^1)=f([-1,1]^1 \vee_{M^1} [2,4]^1)-f([-1,1]^1 \wedge_{M^1} [2,4]^1)=32.5+3.5=36$. On the other hand, a slowly increasing (saturated) underlying positive valuation function is shown in Fig.4(b), that is in particular the logistic function $f_2(x) = \frac{2}{1+e^{-x}} - 1$. The computation of positive valuation function of $v(.)$ with $f=f_2$ and $c(1)=0.5$, results in $d([-1,1]^1,[2,4]^1) = f([-1,1]^1 \vee_{M^1} [2,4]^1)-f([-1,1]^1 \wedge_{M^1} [2,4]^1) \approx 0.713072 + 0.149738 \approx 0.862810$.

The latter example has demonstrated that different *underlying positive valuation functions* result in different distances between "fixed" intervals. Apparently, an underlying positive valuation function is an instrument for introducing non-linearities. The author of this work expects that the definition of a "good" underlying positive valuation function depends on an application. Since the goal of this work is to introduce novel tools rather than to optimise the utility of the (novel) tools introduced here, the positive valuation $f(x)=x$ has been employed for simplicity henceforward.

III. THE METRIC FUZZY LATTICE $\mathbb{F}$ OF *FIN*S

The previous section has studied lattices $\mathbb{M}^h$, $h\in(0,1]$ of generalized intervals. It was explained how $\aleph_1$ different positive valuation functions can be introduced in a lattice $\mathbb{M}^h$, where $\aleph_1$ is the cardinality of the set $\mathbb{R}$ of real numbers, moreover a positive valuation function in $\mathbb{M}^h$ implied both a metric distance d_h and an inclusion measure function k; furthermore, a metric d_M was introduced in the space $\mathbb{M}= \bigcup_{h\in(0,1]} \mathbb{M}^h$. Based on the previous analysis useful tools are introduced in this section. Consider the following definition.

Definition 9: A *Fuzzy Interval Number* (*FIN*) is a continuous function F: $(0,1]\to\mathbb{M}^h_+$ (*positive FIN*), or F: $(0,1]\to\mathbb{M}^h_0$ (*trivial FIN*), or F: $(0,1]\to\mathbb{M}^h_-$ (*negative FIN*) such that $h_1\le h_2 \Rightarrow support(F(h_1)) \supseteq support(F(h_2))$, where $0 < h_1 \le h_2 \le 1$.

We remark that without loss of generality the generalized interval $F(1)$ can be a trivial generalized interval. Also, in the above definition a *FIN F* is required to be a *continuous* function in the sense that for $h_0 \in (0,1)$, given $\varepsilon>0$, there is a $\delta(\varepsilon)>0$ such that for all $h \in (0,1)$ with $|h-h_0|<\delta$ it follows $d_M(F(h),F(h_0))<\varepsilon$. The set of *FIN*s is denoted by F; more specifically the sets of *positive*, *trivial* and *negative FIN*s will be denoted, respectively, by F_+, F_0 and F_-. Fig.5 shows a positive *FIN* (F_p), a trivial *FIN* (F_t) and a negative *FIN* (F_n) along the horizontal axis.

We point out explicitly that a *FIN* is a "number" not adhered to a specific interpretation; in other words, a *FIN* is regarded as an abstract "mathematical object". *FIN*s could have various interpretations and uses. For instance [26] employs a normed linear space of *FIN*s, including both positive and negative *FIN*s, for function approximation. Note that positive *FIN*s can be interpreted as conventional convex fuzzy sets. Furthermore, a statistical interpretation for a positive *FIN F* is presented here below. An ordering relation has been introduced in the set F of *FIN*s as follows.

Definition 10: Let $F_1, F_2 \in \mathsf{F}$, then $F_1 \leq_{\mathsf{F}} F_2 \Leftrightarrow F_1(h) \leq_{M^h} F_2(h)$, for all h in $(0,1]$.

That is, a *FIN* F_1 is smaller-than or equal-to another *FIN* F_2 if and only if every generalized interval $F_1(h)$ of F_1 is smaller-than or equal-to the corresponding generalized interval $F_2(h)$ of F_2 for $h \in (0,1]$. The following proposition establishes that the ordering relation $\leq_{\mathsf{F}}$ is a *partial ordering relation*.

Proposition 11: The set F of *FIN*s is a *partially ordered set* because the ordering relation $\leq_{\mathsf{F}}$ is

(a) *reflexive*: $F_1 \leq_{\mathsf{F}} F_1$,

(b) *antisymmetric*: $F_1 \leq_{\mathsf{F}} F_2$ and $F_2 \leq_{\mathsf{F}} F_1 \Rightarrow F_1 = F_2$, and

(c) *transitive*: $F_1 \leq_{\mathsf{F}} F_2$ and $F_2 \leq_{\mathsf{F}} F_3 \Rightarrow F_1 \leq_{\mathsf{F}} F_3$, for F_1, F_2, F_3 in F.

The proof of proposition 11 is shown in the Appendix.

It turns out that the partially ordered set F is a lattice. The following proposition introduces a metric distance in lattice F.

Proposition 12: Let F_1 and F_2 be *FIN*s in lattice F. A *metric distance* function d_K: $\mathsf{F}\times\mathsf{F}\rightarrow\mathsf{R}$ is given by $d_K(F_1,F_2)= c\int_0^1 d_h(F_1(h),F_2(h))dh$, where c is a positive normalizing coefficient, $d_h(F_1(h),F_2(h))$ is a metric distance between generalized intervals $F_1(h)$ and $F_2(h)$, and $\int(.)$ is the conventional integral operator.

The proof of proposition 12 is shown in the Appendix.

A specific value for the *positive normalizing coefficient c* in proposition 12 depends on the application. Furthermore, we point out that *negative* generalized intervals are used implicitly in the computation of metric d_K. More specifically the computation of d_K is based on the metric distance $d_h(F_1(h),F_2(h))$ whose computation involves explicitly negative generalized intervals.

It has been pointed out above that positive *FIN*s can be interpreted as conventional convex fuzzy sets. Recall from [58] that a fuzzy set is *convex* if and only if the crisp set Γ_a, namely *a-level set* or equivalently *a-cut* defined by $\Gamma_a=\{x|\mu(x) \geq a\}$, is convex for a in (0,1]. It follows that when a FIN A is interpreted as a convex fuzzy set then the a-cut Γ_a of A is a closed interval such that $a \leq b$ implies $\Gamma_a \supseteq \Gamma_b$, more specifically it is $\Gamma_h = support(F(h))$. It turns out, when positive *FIN*s are interpreted as convex fuzzy sets, that the lattice ordering relation $\leq_F$ between (positive) *FIN*s is identical with the standard fuzzy subset relation [58].

The metric d_K presented above reduces to the metric distance $d_1(X,Y)= \frac{1}{2}\int_{a=0}^{1}(|x_l^a - y_l^a| + |x_r^a - y_r^a|)da$ between two fuzzy numbers X and Y reported in the literature [5], [6], where $[x_l^a, x_r^a]$ and $[y_l^a, y_r^a]$ stand for the corresponding a-cuts of the fuzzy numbers X and Y, respectively. Note that lattice theory is not mentioned in [5], [6]. Using the terminology and notation presented in this work the latter metric distance $d_1(X,Y)$ can be regarded as a specific form of the metric distance d_K for underlying positive valuation function $f(x)=x$, $c(h)=1$ and $c=\frac{1}{2}$.

It is also interesting to compare metric distance d_K with another well-known metric distance between convex fuzzy sets given by the following equation

$$d_p(u,v)= (\int_0^1 d_H([u]^a,[v]^a)^p da)^{1/p} \quad (3)$$

whose calculation is based on the Hausdorf metric distance d_H between the a-cuts $[u]^a$ and $[v]^a$ of two fuzzy sets u and v, respectively, as detailed in [12]; note also that the aforementioned metric $d_p(u,v)$ for both p=1 and p=∞ also appears in [45], [60]. The Hausdorf metric $d_H(.,.)$ is a generalization of the distance between two points in a metric space to two compact nonempty subsets of the space [60]. For practical and theoretical purposes the membership functions of the fuzzy sets involved in the computation of $d_p(u,v)$ have been restricted to be *upper semicontinuous* [12]. Advantages and disadvantages of the metric distances d_p and d_K are discussed comparatively in the following.

The metric distance d_p can be computed in space $\mathbf{R}^N$ for any number N of dimensions based on the Hausdorf metric d_H, whereas the metric d_K is computed in $\mathbf{R}^N$ only via its computation in $\mathbf{R}$ based on the metric d_h. It follows that the metrics d_H and d_h produce different results in space $\mathbf{R}^1$; for instance consider two conventional intervals $[a,b]$ and $[c,d]$. Then based on the *Hausdorf separation* function [12] the

Hausdorf metric is given by $d_H([a,b],[c,d])= \max\{|a-c|,|b-d|\}$, whereas metric d_h with underlying positive valuation function $f(x)=x$ and $h=1$ is given by $d_h([a,b],[c,d])= |a-c|+|b-d|$. Moreover the employment of d_h produces "intuitively expected" results, whereas the Hausdorf metric d_H may produce "counter-intuitive" results as explained in the following. For instance $d_H([1,2],[3,9])= \max\{2,7\}= 7$, moreover $d_H([1,2],[8,9])= \max\{7,7\}= 7$; whereas, $d_h([1,2],[3,9])= 2+7= 9$, moreover $d_h([1,2],[8,9])= 7+7= 14$. In words, the metric d_h concludes that intervals [1,2] and [3,9] are closer to each other than intervals [1,2] and [8,9] are, the latter is considered "intuitive". Whereas, the Hausdorf metric d_H concludes that intervals [1,2] and [3,9] are as far from each other as intervals [1,2] and [8,9] are, the latter is "counter-intuitive".

Furthermore there is an overwhelming advantage, both theoretically and practically, for the employment of metric d_K versus metric d_p as explained in the following. From a theoretical point of view, there are $\aleph_0$ different metrics d_p for all different integer values p=1,2,3,… whereas there exist $\aleph_1=2^{\aleph_0} > \aleph_0$ different metrics d_K, where $\aleph_0$ and $\aleph_1$ are the cardinalities of the sets of integers and real numbers, respectively. From a practical point of view note that only a finite number of metric distances between fuzzy numbers can be computed due to the finite word-length of a digital computer. More specifically, on the one hand, only a small number of distinct metrics d_p, p=1,2,… can be used in practice because number $d_{p_0}(u,v)$ is practically no different than number $d_{p_0+1}(u,v)$ in a digital computer for large p_0. On the other hand, the capacity to calculate metric d_K based on any strictly increasing function f can potentially produce a much larger (finite) number of metrics d_K thus taking full advantage of the existing digital computer memory resources.

The aforementioned advantages extend to inclusion measure functions σ_K in the non-complete lattice $\mathsf{F}_{nn}=\mathsf{F}\backslash\mathsf{F}_-$ as explained in the following.

Proposition 13: Let F_1 and F_2 be non-negative *FIN*s in $\mathsf{F}_{nn}=\mathsf{F}\backslash\mathsf{F}_-$. An *inclusion measure* function σ_K: $\mathsf{F}_{nn}\times\mathsf{F}_{nn}\to[0,1]$ is defined by $\sigma_K(F_1,F_2)= \int_0^1 k(F_1(h),F_2(h))dh$, where $k(F_1(h),F_2(h))$ is an inclusion measure function between non-negative generalized intervals $F_1(h)$ and $F_2(h)$, and $\int(.)$ is the conventional integral operator.

The proof of proposition 13 is shown in the Appendix.

Since $\sigma_K(F_1,F_2)$ quantifies a degree of inclusion of *FIN* F_1 into *FIN* F_2 symbol $\sigma_K(F_1\leq_F F_2)$ might be used instead of symbol $\sigma_K(F_1,F_2)$ in the sequel. Both computations of the metric distance d_K and the inclusion measure σ_K between *FIN*s involve the computation of a definite integral. Furthermore note that both d_K and σ_K constitute mathematically sound alternatives for quantifying the "relative position" on a *FIN* with respect to another one, which (alternatives) could be useful in applications.

Alternative definitions of a fuzzy inclusion index have been proposed in the literature for quantifying a degree of inclusion of a fuzzy set into another one [7], [30], [44], [47]. The aforementioned alternative definitions have been proposed in a fuzzy set-theoretic context furthermore the fuzzy sets involved need to overlap otherwise the corresponding inclusion index equals zero. On the other hand, the fuzzy inclusion measure σ_K proposed here can quantify a degree of inclusion of one fuzzy set into another one in a "fuzzy lattice sense". Therefore σ_K is typically non-zero for non-overlapping fuzzy sets; the latter has been valuable in various pattern recognition applications as pointed out in [25] in a discussion on the extended choice (Weber) function. It should also be pointed out that there exists at least one more inclusion measure function σ given as follows

$$s(x \leq_L u) = \frac{v(x \wedge_L u)}{v(u)} \tag{4}$$

where x, u are elements of a general lattice L with positive valuation function v:$\mathsf{L}\rightarrow\mathsf{R}$; it follows $s(x\leq_L u)= 0$ for $x\|_L u$ [40]. If positive *FIN*s are interpreted as convex fuzzy sets then it follows that the fuzzy degree of inclusion $s(x\leq_L u)$ of a fuzzy set A to a non-overlapping fuzzy set B equals zero.

A number of examples are shown in the following to illustrate the utility of the aforementioned tools σ_K and d_K in the lattice F_+ of positive *FIN*s.

Example 14: Consider *FIN*s E and F_1 in Fig.6(a) with triangular membership functions. *FIN*s F_{1L} and F_{1R} shown, respectively, in Fig.6 (b) and (c) have been produced from *FIN* F_1 by moving the top of F_1 by 0.2 to the left and to the right, respectively. It is expected intuitively that *FIN* F_1 should be included (in a fuzzy lattice sense) in *FIN* E less than it does *FIN* F_{1L}; nevertheless *FIN* F_1 should be included in *FIN* E more than it does *FIN* F_{1R}. Computation of numbers $\sigma_K(F_1\leq_F E)$, $\sigma_K(F_{1L}\leq_F E)$ and $\sigma_K(F_{1R}\leq_F E)$ as shown in proposition 13 has resulted in $\sigma_K(F_1\leq_F E)\approx 0.2512$, $\sigma_K(F_{1L}\leq_F E)\approx 0.2537$ and $\sigma_K(F_{1R}\leq_F E)\approx 0.2488$, as expected intuitively.

Example 15: In this example the degree of inclusion σ_K is calculated of *FIN* F_1 into both *FIN*s E_1 and E_2 shown in Fig.7(a). By inspecting Fig.7(a) it might be intuitively expected that the degree of inclusion $\sigma_K(F_1\leq_F E_2)$ of *FIN* F_1 into *FIN* E_2 should be larger than the degree of inclusion $\sigma_K(F_1\leq_F E_1)$ of *FIN* F_1 into *FIN* E_1 (in a fuzzy lattice sense) because *FIN* E_2 "leans more towards" *FIN* F_1 than *FIN* E_1 does. Using formula $\sigma_K(F_1\leq_F F_2)= \int_0^1 k(F_1(h), F_2(h))dh$ of proposition 13, it follows $\sigma_K(F_1\leq_F E_1)\approx 0.2465$ and $\sigma_K(F_1\leq_F E_2)\approx 0.2602$ as expected intuitively. Fig.7(b) plots functions $k(F_1(h),E_i(h))$ i=1,2 versus h for $h\in(0,1]$. The integral of a function $k(F_1(h),E_i(h))$ i=1,2, that is the area under a curve plotted in Fig.7(b), equals the corresponding degree of inclusion of *FIN* F_1 into *FIN* E_i, i=1,2.

Equipped with an inclusion measure $\sigma_K(.,.)$ the lattice $\mathsf{F}_{nn}=\mathsf{F}\setminus\mathsf{F}_{-}$ of non-negative *FIN*s becomes a *fuzzy lattice* $<\mathsf{F}_{nn},\sigma_K>$. Recall that metric d_K is also available in the lattice F_{nn}. Both the inclusion measure σ_K and the metric distance function d_K can be used for quantifying the "relative position" of two *FIN*s. In particular, inclusion measure $\sigma_K(F_1 \leq_F F_2)$ can be used for quantifying a *fuzzy lattice degree of inclusion* of *FIN* F_1 into *FIN* F_2, whereas the metric distance $d_K(F_1,F_2)$ can be used for quantifying the *proximity* of *FIN*s F_1 and F_2. It follows that lattice F_{nn} could be used for unifying rigorously *heterogeneous data* including fuzzy sets, numbers and intervals [24], [37]. However, the employment of σ_K could be misleading as demonstrated in the following example.

Example 16: Fig.8 shows two pairs of *FIN*s, namely pair E and F_1 (above) and pair E and F_2 (below). Using formula $\sigma_K(F_1 \leq_F F_2)= \int_0^1 k(F_1(h), F_2(h))dh$ of proposition 13, it was calculated $\sigma_K(E \leq_F F_1)=0.1256$ and $\sigma_K(E \leq_F F_2)=0.3768$, that is *FIN* E is included more in *FIN* F_2 than it is in *FIN* F_1, as it might be expected intuitively by inspecting Fig.8. Nevertheless it is $\sigma_K(F_1 \leq_F E) = 0.3140 = \sigma_K(F_2 \leq_F E)$, that is *FIN* F_1 is included in *FIN* E as much as *FIN* F_2 is included in *FIN* E. The latter counter-intuitive equality is due to the fact that the computation of $k(F_i(h),E(h))$, i=1,2 in proposition 13 is based on the computation of $v(F_i(h) \vee_{M^h} E(h))$ i=1,2 where the left-end of generalized interval $F_i(h)$ i=1,2 has no effect in the computation of $F_i(h) \vee_{M^h} E(h)$; it follows $\sigma_K(F_1 \leq_F E) = \sigma_K(F_2 \leq_F E)$.

The aforementioned counter-intuitive result can be amended using the metric distance d_K instead. More specifically it has been computed $d_K(E,F_2) \approx 5.1667 \leq 5.8334 \approx d_K(E,F_1)$; the latter inequality is expected intuitively from Fig.8. It is worthwhile pointing out that in extensive experiments using real world data in the Hellenic Sugar Industry (HSI) domain regarding prediction of sugar production, the employment of d_K has resulted in more accurate predictions than the employment of σ_K.

The following section IV illustrates how a positive *FIN* can be constructed from a population of samples (measurements), moreover a statistical interpretation for a (positive) *FIN* F is proposed.

IV. CONSTRUCTION AND A STATISTICAL INTERPRETATION OF A *FIN*

This section presents algorithm CALFIN for constructing a positive *FIN* from a population of numbers. A statistical interpretation of a constructed *FIN* is also shown. Finally, it is explained how the techniques introduced in this work can be used for calculating a (metric) distance between either two sets of real numbers or two probability density functions.

A. Calculating a FIN from a Population of Measurements

Consider a vector of real numbers (samples/measurements) $x=[x_1,x_2,\ldots,x_N]$ such that the numbers $x_1,x_2,\ldots,x_N$ are ordered incrementally, that is $x_1 \le x_2 \le \ldots \le x_N$. A *FIN* can be constructed according to the following procedure. Calculate the *median*[5] number *median(x)* of the N numbers in vector $x=[x_1,x_2,\ldots,x_N]$. Insert number *median*(x) in a vector *pts*. Split vector x into two "half vectors", these are *x_left* which contains the numbers of vector x smaller than *median(x)*, and vector *x_right* which contains the numbers of vector x larger than *median(x)*. The aforementioned procedure is repeated again for vectors *x_left* and *x_right* in order to calculate the numbers *median(x_left)*, *median(x_right)*, which are both inserted in vector *pts*. The previous procedure is repeated recursively $log_2 N$ times, until "half vectors" are computed including a single number x_i, $i=1,\ldots,N$; the latter numbers are *median* numbers by definition. The computed median values are stored (sorted) in vector *pts* whose entries constitute the abscissae values of a positive *FIN*'s membership function. The corresponding ordinate values are computed in vector *val* as explained in the following.

CALFIN: An Algorithm for Calculating a *FIN*

1. Let x be a vector of real numbers.
2. Order incrementally the numbers in vector x.
3. Initially vector *pts* is empty.
4. function *calfin*(x)
5. { while ($|x| \neq 1$)
6. *med*:= *median*(x)
7. insert *med* in vector *pts*
8. *x_left*:= elements in vector x less-than number *median*(x)
9. *x_right*:= elements in vector x larger-than number *median*(x)
10. *calfin*(*x_left*)
11. *calfin*(*x_right*)
12. endwhile
13. } //function *calfin*(x)
14. Sort vector *pts* incrementally.
15. Let $|pts|$ denote the cardinality of vector *pts*. Store in vector *val*, $|pts|/2$ numbers from 0 up to 1 in steps of $2/|pts|$ followed by another $|pts|/2$ numbers from 1 down to 0 in steps of $2/|pts|$.

[5] The *median*(x), where $x=[x_1,x_2,\ldots,x_N]$ is a vector, is defined to be a number such that half of the N numbers $x_1,x_2,\ldots,x_N$ are smaller than *median*(x) and the other half are larger than *median*(x); for instance, the *median*($[x_1,x_2,x_3]$) with $x_1 < x_2 < x_3$ equals x_2, whereas the *median*($[x_1,x_2]$) with $x_1 < x_2$ was computed here as *median*($[x_1,x_2]$)= $(x_1 + x_2)/2$.

Algorithm CALFIN computes two vectors, namely *pts* and *val*, where vector *val* includes the degrees of fuzzy membership of the corresponding real numbers in vector *pts*. Note that during a preprocessing step, when identical numbers appear in the original vector x, then only the first one is retained whereas each of the subsequent identical numbers is replaced by the sum of its previous number and a very small user-defined positive number ε; in this way identical numbers are replaced by distinct numbers in an arbitrarily small neighbourhood. Algorithm CALFIN produces a positive *FIN* with membership function $\mu(x)$ such that $\mu(x)=1$ for exactly one number x.

Fig.9 displays *FIN G* constructed from 222 random numbers generated according to the normal probability density function N(0,1) with *mean* 0 and *standard deviation* 1. Note that the maximum value (1) of *FIN G* is attained close to the *mean* 0 of probability density function N(0,1).

B. A Statistical Interpretation of FINs, etc.

Due to the aforementioned construction algorithm of a *FIN F* it follows that when algorithm CALFIN is applied on a vector $x=[x_1,x_2,\ldots,x_N]$ of numbers then approximately 100(1-h) % of the n numbers in vector x are within interval *support*($F(h)$). If a large number of samples is drawn independently according to a probability distribution function $f_0(x)$ and a *FIN F* is constructed using algorithm CALFIN, then interval *support*($F(h)$) constitutes, by construction, an *interval of confidence at level-h*. That is, a random number drawn according to the probability distribution function $f_0(x)$ is expected to be included in interval *support*($F(h)$) with confidence 100(1-h)%.

The previous analysis implies a one-one correspondence between *FIN*s and *probability density functions* (*pdf*s) as explained in the following. Based on the one-one correspondence between *pdf*s and *cumulative distribution functions* (*CDF*s), a one-one correspondence will be shown between *CDF*s and *FIN*s. In the one direction, it is shown how a *CDF* $G(x)$ is mapped to a *FIN F*: Let x_0 be such that $G(x_0)=0.5$. Let the membership function $\mu_F(.)$ of a *FIN F* be defined such that $\mu_F(x)=2G(x)$ for $x \le x_0$ furthermore $\mu_F(x)=2[1-G(x)]$ for $x \ge x_0$. In the other direction, it is shown how a *FIN F* is mapped to a *CDF*: Let x_0 be such that $\mu_F(x_0)=1$. A *CDF* $G(x)$ is defined such that $G(x)=0.5\mu_F(x)$ for $x \le x_0$ furthermore $G(x)=1-0.5\mu_F(x)$ for $x \ge x_0$. From the previous it follows that a *FIN* which corresponds to a uniform probability distribution over a range $[a,b]$ is a *FIN* with isosceles triangular membership function of unit height with base $[a,b]$. The previous analysis also implies that metric $d_K(.,.)$ can be employed for calculating a distance between either two probability density functions or two cumulative distribution functions.

It might be useful to point out that another function for calculating proximity of two probability distributions is the so-called *Kullback-Leibler* (*KL*) *distance* [17]. On the one hand, a disadvantage of KL-distance is that the *CDF*s involved in the calculations need to be defined on the same elements of a set otherwise spurious results are produced such as 0 or ∞. On the other hand, distance d_K can be used to

calculate the proximity of two *CDF*s defined even on mutually disjoint intervals of numbers. Finally note that KL-distance is not a metric because it does not satisfy laws D3 (Symmetry) and D4 (Triangle Inequality).

An additional potential for algorithm CALFIN stems from the construction of a *FIN* from a population of numbers. Therefore, function $d_K(F_1,F_2)$ can also be used for computing a (metric) distance between two populations of numbers as demonstrated in the following example.

Example 17: Three populations of random numbers were generated randomly according to the uniform probability distribution in the ranges [0,2], [0.9,1.1], and [0.2, 2.2], respectively. Each population included 111 random numbers. The averages of the three populations of random numbers have been, respectively, 1.0467, 1.0030 and 1.1625. Fig.10 shows the three *FIN*s constructed using algorithm CALFIN. The following distances were calculated: $d_K(F_1,F_2)$=0.2843, $d_K(F_2,F_3)$=0.3480 and $d_K(F_1,F_3)$=0.1002. The aforementioned distances d_K confirm what might have been expected by observing Fig.10, that is the proximity of the two populations of random numbers in the ranges [0,2] and [0.2, 2.2] is larger than the proximity of the two populations of numbers in the ranges [0,2] and [0.9,1.1], even though both [0.9,1.1] $\subseteq$ [0,2] and the two populations of numbers drawn randomly in the ranges [0,2] and [0.9,1.1] have nearly equal average values. It is worthwhile pointing out that should the distances between the population averages were used then the first two populations of numbers would be the nearest to each other.

This example also demonstrates experimentally the fact that a *FIN F*, constructed from *N* numbers drawn randomly according to the uniform probability distribution, has nearly an isosceles triangular membership function. More specifically in the limit when *N* becomes very large then *FIN F* becomes an isosceles triangle as it has been confirmed experimentally using progressively larger values of *N* than *N*=111.

V. PREDICTION OF SUGAR PRODUCTION BASED ON *FIN*S

The sugar-production prediction problem is outlined briefly in this section, in the HSI (Hellenic Sugar Industry) domain, followed by experimental results and discussion.

The majority of sugar produced in Europe is extracted from the storage root of *Beta Vulgaris L* (common name: *sugar-beet*), which is planted in early spring and harvested during late-summer/early-fall. In particular in Greece, sugar production is organized in five agricultural districts with factories located at Larisa, Platy, Serres, Xanthi, and Orestiada. In the beginning of a year a set of *pilot fields* is selected in every agricultural district for sampling. An early season accurate prediction of sugar production is critical for safeguarding the continuous operation of a sugar production factory since interruptions are very costly.

Sample measurements of ten *production variables* and eight *meteorological variables* were available in this work for eleven years from 1989 to 1999 from three agricultural districts, namely *Larisa*, *Platy*, and

Serres [41]. The *production variables* included: 1) average root weight, in *g*, 2) POL: percentage of sugar in fresh root weight, 3) *a*-amino-Nitrogen (*a-N*), in *meq/100g root*, 4) potassium (K), in *meq/100g root*, 5) sodium (Na), in *meq/100g root*, 6) Leaf Area Index (LAI): leaf area per field area ratio, 7) TOP: plant top weight, in *kg/hectare*, 8) Roots Weight (RW), in *kg/hectare*, 9) Nitrogen-test (N-test): NO_3-N content in petioles, in $mg.kg^{-1}$, and 10) the planting date. Sugar production was calculated as POL*RW. The *meteorological variables* included: 1) average daily temperature, in ^oC, 2) maximum daily temperature, in ^oC, 3) minimum daily temperature, in ^oC, 4) relative humidity, 5) wind speed, daily average in *miles/hour*, 6) daily precipitation, in *mm*, 7) daily evaporation, in *mm*, and 8) sunlight, in *hours/day*. The production variables were sampled every 20 days, whereas the meteorological variables were sampled daily. The term *population of measurements* is used here to denote either, first, a collection of *a production variable* samples obtained during 20 days from all pilot fields or, second, a collection of *meteorological variable* samples obtained daily in an agricultural district during the aforementioned 20 days.

In the context of this work prediction was effected "by classification" using a nearest neighbour classifier as explained in the following. Each training year was labelled by an expert as "poor", "medium", or "good" according to the corresponding industrial sugar yield; finally a testing year was assigned the category label of its nearest neighbour training year. Two types of distances were employed here comparatively. First, the conventional Euclidean distance between the average values of populations of measurements; second, the metric distance d_K between *FIN*s computed, respectively, from the corresponding populations of measurements using algorithm CALFIN. Details regarding the application of a nearest neighbour classifier *FINkNN* are shown in [41]. Some aspects of the application of classifier *FINkNN* are outlined in the following using public meteorological data for the months of July and August for the port of Thessaloniki area, that is where sugar-beets are grown in Greece.

Fig.11(a) displays four *FIN*s, namely *T89*, *T92*, *T96* and *T99* computed by algorithm CALFIN from daily average temperatures for July and August of the years 1989, 1992, 1996 and 1999, respectively. Furthermore, Fig.11(b) displays four *FIN*s, namely *H89*, *H92*, *H96* and *H99* computed likewise from daily relative humidities for July and August, respectively. Table I displays the distances d_K between all pairs of *FIN*s among *T89*, *T92*, *T96* and *T99*, whereas Table II shows the distances d_K between all pairs of *FIN*s among *H89*, *H92*, *H96* and *H99*.

Table III shows experimental results using four different prediction methods. The results in Table III are shown in the chronological order of their publication. More specifically, a minimum sugar prediction error around 6% has been published in [42] by Bayesian Combined Predictor (BCP) whose local predictors include linear regressors, artificial neural networks, polynomial predictors, etc.; the results in Table III by both a first-principles model (15%) and intelligent-clustering techniques (5%) have been published in [27]; finally, the result by *FINkNN* classifier (2%) is from [41]. The first three methods in Table III replace a population of measurements by the corresponding average. The last two methods in Table III achieve

prediction “by classification” as outlined above; more specifically the method “Intelligent clustering techniques” (line 3 in Table III) uses the distances between the averages of populations of measurements, whereas method “Classifier *FINkNN*” uses the distances between *FIN*s.

Classifier *FINkNN* in Table III used *FIN*s computed from populations of selected input variables. More specifically, input variables Relative Humidity, and Roots Weight were selected for the *Larisa* agricultural district, input variables Daily Precipitation, Sodim (Na), and Average Root Weight for *Platy*, furthermore input variables Daily Precipitation, Average Root Weight, and Roots Weight were selected for the *Serres* agricultural district as detailed in [41]. Therefore, it seems reasonable to expect that different input variables would be the most effective for predicting sugar production in other countries.

The experimental goal was to achieve prediction of sugar production by classification in one of the classes “good”, “medium”, or “poor”; in particular, the goal in this work was to predict the sugar production level in September based on data available at the end of July. For all agricultural districts a “leave-one-out” series of eleven experiments was carried out such that one year among years 1989 to 1999 was left out, in turn, for testing whereas the remaining ten years were used for training as detailed in [41]. Note that the computational model for “prediction by classification” outlined in this work is highly transferable and, provided that sufficient data are available, can be trained for any region or country. The error rate of 2% reported in Table III for classifier *FINkNN* is the average over both three agricultural districts and eleven years per agricultural district.

The improved classification result of 2.0% by classifier *FINkNN* (line 4 of Table III) has been attributed to the employment of *FIN*s. More specifically it is argued in [41] that using the average value of a population of measurements in a prediction model could be misleading. For instance, two different daily precipitation patterns in a month may be characterized by identical average values, nevertheless their effect on the annual sugar production level might be drastically different; in contrast, the computation of metric d_K between *FIN*s is sensitive to both higher order statistics and other features of a distribution of samples/measurements, e.g. the *skewness* of a distribution, etc. A *FIN*, computed from a population of measurements, has been treated as a single datum; in conclusion a whole population of measurements are jointly considered instead of considering a single, “best” in a sense, measurement the latter is a typical practice in the probabilistic data processing framework.

VI. DISCUSSION AND CONCLUSION

This work has presented novel mathematical tools developed for improving prediction of sugar production for Hellenic Sugar Industry (HSI), Greece. Previous prediction models replace a population of measurements by the corresponding average then a prediction model is used [9], [18], [27], [28], [34], [42], [53], [54]. In the context of this work a population of measurements was represented by a *FIN* (Fuzzy

Interval Number) producing improved prediction results "by classification". Detailed experimental results have been shown comparatively elsewhere [41]. Moreover, the prediction problem in the HSI domain has been also detailed in [27] where a software package has been developed and can be updated on a yearly basis. This work focuses on an analytic study of *FIN*s in the context of mathematical lattice theory.

A *FIN* was described as a set of generalized intervals either, positive, trivial or negative generalized intervals. Based on lattice theory both a metric distance d_K and a lattice inclusion measure function σ_K were introduced analytically in the set F of *FIN*s. A *FIN* has been presented here as a number not adhered to a specific interpretation. It was explained how a positive *FIN* can be interpreted also as a convex fuzzy set.

A simple form of the aforementioned metric d_K has been presented and applied in the literature in a fuzzy set-theoretic context based on *a*–cuts [5], [6]. The enhanced metric d_K has introduced novel (lattice theoretic) perspectives with useful practical implications. For instance, the universe of discourse has been separated from the mathematical field R of real numbers using a underlying positive valuation function. There followed $\aleph_1$ different metric distances between *FIN*s, whereas there are "only" $\aleph_0$ different metric distances between convex fuzzy sets, where $\aleph_0$ and $\aleph_1$ are the cardinalities of the sets of integers and real numbers, respectively, with $\aleph_1 = 2^{\aleph_0} > \aleph_0$ [48]. It was discussed how generalized intervals based on an underlying positive valuation function, can imply a much larger (finite) number of metrics than *a*-cuts can imply in practical applications.

Algorithm CALFIN was presented for constructing a *FIN* from a population of samples/measurements. Moreover, a statistical interpretation of a *FIN* $F(h)$ was proposed such that interval $support(F(h))$ was interpreted as an interval of confidence at *level-h*.

The mathematical tools presented in this work can be useful in a number of applications. For instance a (metric) distance function, such as distances d_K, is typically required in fuzzy-, and interval regression analysis [2], [11], [13], [16], [50], [51], [56].

Furthermore, an employment of the lattice inclusion measure σ_K in neuro-fuzzy models [21], [55] could be useful for analysis and design (Note that the effectiveness of the lattice inclusion measure function σ_K, as an activation function in neural-computing, has already been demonstrated in applications [25], [38], [40]). Moreover note that the logistic function, which is often used as an activation function by the neurons of a neural network, has been presented in example 8 here as a "saturated" underlying positive valuation function. Alternative non-linear underlying positive valuation functions could be employed by the neurons of a neural network to improve the effectiveness of various neural-fuzzy techniques in applications.

Fuzzy sets, in both fuzzy regression and neuro-fuzzy modelling techniques, are frequently employed for approximating a real function $f: \mathsf{R}^m \rightarrow \mathsf{R}^n$. The introduction of potentially $\aleph_1$ different metric distances in the lattice of convex fuzzy sets could improve the design of Fuzzy Inference Systems (*FIS*) [14], [31], [43], [46], [52] using standard function approximation tools [8].

APPENDIX

A. Proof of Proposition 11

(a) Let $F_1 \in \mathsf{F}$. Then $F_1(h) \leq_{\mathsf{M}^h} F_1(h)$ for all h in (0,1], hence $F_1 \leq_{\mathsf{F}} F_1$.

(b) Let $F_1, F_2 \in \mathsf{F}$. First, $F_1 \leq_{\mathsf{F}} F_2$ is equivalent to $F_1(h) \leq_{\mathsf{M}^h} F_2(h)$ for all h in (0,1]. Second, $F_2 \leq_{\mathsf{F}} F_1$ is equivalent to $F_2(h) \leq_{\mathsf{M}^h} F_1(h)$ for all h in (0,1]. Since M^h is a lattice, $F_1(h) \leq_{\mathsf{M}^h} F_2(h)$ and $F_2(h) \leq_{\mathsf{M}^h} F_1(h)$ jointly imply $F_1(h) = F_2(h)$ for all h in (0,1]. In conclusion, $F_1 = F_2$.

(c) Let $F_1, F_2, F_3 \in \mathsf{F}$. Then $F_1 \leq_{\mathsf{F}} F_2$ is equivalent to $F_1(h) \leq_{\mathsf{M}^h} F_2(h)$, and $F_2 \leq_{\mathsf{F}} F_3$ is equivalent to $F_2(h) \leq_{\mathsf{M}^h} F_3(h)$ for h in (0,1]. Since M^h is a lattice, $F_1(h) \leq_{\mathsf{M}^h} F_2(h)$ and $F_2(h) \leq_{\mathsf{M}^h} F_3(h)$ jointly imply $F_1(h) \leq_{\mathsf{M}^h} F_3(h)$ for all h in (0,1]. In conclusion, $F_1 \leq_{\mathsf{F}} F_3$.

Thus, the proposition has been proven.

B. Proof of Proposition 12

Function $d_{\mathrm{h}}(F_1(h), F_2(h))$, $h \in (0,1]$ has been introduced as a metric distance in the lattice M^h of generalized intervals. To prove that $d_{\mathrm{K}}(F_1,F_2) = c\int_0^1 d_{\mathrm{h}}(F_1(h),F_2(h))dh$ defines a metric distance in the lattice F of *FIN*s, laws (MD1)-(MD4) (see in footnote 4) need to be satisfied. For simplicity, and without loss of generality, it is assumed in the remaining of this proof that c=1.

(MD1) $d_{\mathrm{K}}(F_1,F_2) = \int_0^1 d_{\mathrm{h}}(F_1(h),F_2(h))dh$, where $d_{\mathrm{h}}(F_1(h),F_2(h)) \geq 0$. Therefore $d_{\mathrm{h}}(F_1, F_2) \geq 0$.

(MD2) On the one hand, $d_{\mathrm{K}}(F_1,F_1) = \int_0^1 d_{\mathrm{h}}(F_1(h),F_1(h))dh = \int_0^1 0dh = 0$.

On the other hand, let $d_{\mathrm{K}}(F_1,F_2)=0$, that is let $\int_0^1 d_{\mathrm{h}}(F_1(h),F_2(h))dh = 0$. It is known that $d_{\mathrm{h}}(F_1(h),F_2(h)) \geq 0$, hence $d_{\mathrm{h}}(F_1(h),F_2(h))$ is equal to zero *almost everywhere*. Note that the term *almost everywhere* is used here in a "measure theoretic sense", that is $d_{\mathrm{h}}(F_1(h),F_2(h))$ is larger than 0 on a set of measure zero [15]. It follows that $F_1(h) = F_2(h)$ almost everywhere on (0,1]. Since, according to definition 9, both functions F_1 and F_2 are continuous, it follows $F_1(h)=F_2(h)$, $h \in (0,1]$; hence $F_1=F_2$.

(MD3) $d_{\mathrm{K}}(F_1,F_2) = \int_0^1 d_{\mathrm{h}}(F_1(h),F_2(h))dh = \int_0^1 d_{\mathrm{h}}(F_2(h),F_1(h))dh = d_{\mathrm{K}}(F_2,F_1)$.

(MD4) $d_K(F_1,F_3)= \int_0^1 d_h(F_1(h),F_3(h))dh \le \int_0^1 [d_h(F_1(h),F_2(h))+d_h(F_2(h),F_3(h))]dh = \int_0^1 d_h(F_1(h),F_2(h))dh$

$+ \int_0^1 d_h(F_2(h),F_3(h))dh = d_K(F_1,F_2) + d_K(F_2,F_3).$

Thus, the proposition has been proven.

C. Proof of Proposition 13

It is shown below that function $\sigma_K(F_1,F_2)= \int_0^1 k(F_1(h),F_2(h))dh$ satisfies conditions (IM1)-(IM3) of definition 4.

(IM1) $\sigma_K(F_1,F_1)= \int_0^1 k(F_1(h),F_1(h))dh = \int_0^1 1dh = 1.$

(IM2) Let $F_1 <_F F_2$. Since both functions F_1 and F_2 are continuous it follows $F_1(h) <_{M^h} F_2(h)$ hence $k(F_2(h),F_1(h)) < 1$ on a non-zero set. In conclusion, $\sigma_K(F_2,F_1)= \int_0^1 k(F_2(h),F_1(h))dh < 1.$

(IM3) Let $F_1 \le_F F_2$. It follows $F_1(h) \le_{M^h} F_2(h)$ hence $k(X(h),F_1(h)) \le k(X(h),F_2(h)) \Rightarrow \int_0^1 k(X(h),F_1(h))dh \le$

$\int_0^1 k(X(h),F_2(h))dh \Rightarrow \sigma_K(X,F_1) \le \sigma_K(X,F_2)$ for $X \in F_{nn}=F\backslash F_-$.

Thus, the proposition has been proven.

REFERENCES

[1] G. Alefeld, and J. Herzberger, *Introduction to Interval Computations*. New York, NY: Educational Press, Inc., 1983.

[2] H. Bandemer, and S. Gottwald, *Fuzzy Sets, Fuzzy Logic, Fuzzy Methods with Applications*. Chichester, England: John Wiley & Sons, 1995.

[3] G. Birkhoff, *Lattice Theory*. Providence, RI: American Mathematical Society, Colloquium Publications, vol. 25, 1967.

[4] S.B. Chae, *Lebesgue Integration*. New York, N.Y.: Springer-Verlag, 2nd, 1995.

[5] V. Chatzis, and I. Pitas, "Mean and median of fuzzy numbers", *Proc. IEEE Workshop Nonlinear Signal Image Processing*, Neos Marmaras, Greece, 1995, pp. 297-300.

[6] _____, "Fuzzy Scalar and Vector Median Filters based on Fuzzy Distances", *IEEE Trans. Image Processing*, vol. 8, no. 5, pp. 731-734, 1999.

[7] _____, "A generalized fuzzy mathematical morphology and its application in robust 2-D and 3-D object representation", *IEEE Trans. Image Processing*, vol. 9, no. 10, pp. 1798-1810, 2000.

[8] E.W. Cheney, *Approximation Theory*. New York, N.Y.: Chelsea Publishing Company, 2nd edition, 1982.

[9] B.M. Church, and A. Gnanasakthy, "Estimating sugar production from preharvest samples", *British Sugar Beet Review*, vol. 53, no. 3, pp. 9-11, 1983.

[10] B.A. Davey, and H.A. Priestley, *Introduction to Lattices and Order*. Cambridge, GB: Cambridge University Press, 1990.

[11] P. Diamond, "Fuzzy least squares", *Inf. Sci.*, vol. 46, pp. 141-157, 1988.

[12] P. Diamond, and P. Kloeden, *Metric Spaces of Fuzzy Sets*. Singapore: World Scientific, 1994.

[13] P. Diamond, and R. Körner, "Extended fuzzy linear models and least squares estimates", *Computers Math. Applic.*, vol. 33, no. 9, pp. 15-32, 1997.

[14] S. Dick, and A Kandel, Comment on "Combinatorial rule explosion eliminated by a fuzzy rule configuration", *IEEE Trans. Fuzzy Systems*, vol. 7, pp. 475-477, 1999.

[15] J.L. Doob, *Measure Theory*. New York, NY: Springer-Verlag, 1994.

[16] D. Dubois, and H. Prade, "Operations on fuzzy numbers", *Int. J. Systems Sci.*, vol. 9, no. 6, pp. 613-626, 1978.

[17] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern Recognition*. New York, NY: John Wiley & Sons, 2nd edition, 2001.

[18] R.P. Freckleton, A.R. Watkinson, D.J. Webb, and T.H. Thomas, "Yield of sugar-beet in relation to weather and nutrients", *Agricultural and Forest Meteorology*, vol. 93, no. 1, pp. 39-51, 1999.

[19] G. Grätzer, *Lattice Theory*. San Francisco, CA: W.H. Freeman, 1971.

[20] K. Ito (editor), *Encyclopedic Dictionary of Mathematics*, second edition, the Mathematical Society of Japan, English translation Cambridge, MA: The MIT Press, 1987.

[21] J.-S. R. Jang, and C.T. Sun, "Neuro-fuzzy modeling and control", *Proceedings of the IEEE*, vol. 83, no. 3, pp. 378-406, 1995.

[22] V.G. Kaburlasos, "Novel fuzzy system modeling for automatic control applications", *Proceedings of the 4th Intl. Conference on Technology & Automation*, Thessaloniki, Greece, 2002, pp. 268-275.

[23] V.G. Kaburlasos, "Improved Fuzzy Lattice Neurocomputing (FLN) for Semantic Neural Computing", *Proceedings of the 2003 International Joint Conference on Neural Networks* (*IJCNN'2003*), Portland, OR, 20-24 July 2003, (*to be published*).

[24] V.G. Kaburlasos, and V. Petridis, "Regression on heterogeneous fuzzy data", *Proceedings of the 7th European Congress on Intelligent Techniques and Soft Computing* (*EUFIT'99*), Aachen, Germany, 1999, session CC2.

[25] _____, "Fuzzy Lattice Neurocomputing (FLN) models", *Neural Networks*, vol. 13, no. 10, pp. 1145-1170, 2000.

[26] _____, "Improved prediction of industrial yield based on tools from a normed linear space of Fuzzy Interval Numbers (*FIN*s)", *Proceedings 11th Mediterranean Conference on Control and Automation* (*MED'03*), Rhodes, Greece, 18-20 June 2003, session FM1-B.

[27] V.G. Kaburlasos, V. Spais, V. Petridis, L. Petrou, S. Kazarlis, N. Maslaris, and A. Kallinakis, "Intelligent clustering techniques for prediction of sugar production", *Mathematics and Computers in Simulation*, vol. 60, no. 3-5, pp. 159-168, 2002.

[28] A. Kallinakis, "Criteria and factors of sugar beet quality" (in Greek), Technical Report, Hellenic Sugar Industry S.A., Direction of Agricultural Services, Serres 1991.

[29] A. Kaufmann, and M.M. Gupta. *Introduction to Fuzzy Arithmetic- Theory and Applications*. New York, NY: Van Nostrand Reinhold, 1985.

[30] G.J. Klir, and T.A. Folger. *Fuzzy Sets, Uncertainty, and Information*. Prentice-Hall, Englewood Cliffs, New Jersey, 1988.

[31] L.T. Kóczy, and K. Hirota, "Size reduction by interpolation in fuzzy rule bases", *IEEE Trans. Systems, Man, and Cybernetics - B*, vol. 27, pp. 14-25, 1997.

[32] H. Lee-Kwang, and J.-H. Lee, "A method for ranking fuzzy numbers and its application to decision-making", *IEEE Trans. Fuzzy Systems*, vol. 7, pp. 677-685, 1999.

[33] Q. Liang, and J.M. Mendel, "Interval type-2 fuzzy logic systems: Theory and Design", *IEEE Trans. Fuzzy Systems*, vol. 8, no. 5, pp. 535-550, 2000.

[34] N. Maslaris, P. Christodoulou, and G. Zountsas, "Quantitative effect of root/leaf growth rate on root yield and quality parameters in sugar beet", *Zuckerindustrie*, vol. 122, no. 3, pp. 208-211, 1997.

[35] R. Mattolini, and P. Nesi, "An interval logic for real-time system specification", *IEEE Trans. Software Engineering*, vol. 27, no. 3, 2001.

[36] R.E. Moore, *Methods and Applications of Interval Analysis*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1979.

[37] W. Pedrycz, J.C. Bezdek, R.J. Hathaway, and G.W. Rogers, "Two nonparametric models for fusing heterogeneous fuzzy data", *IEEE Trans. Fuzzy Systems*, vol. 6, no. 3, pp. 411-425, 1998.

[38] V. Petridis, and V.G. Kaburlasos, "Fuzzy Lattice Neural Network (FLNN): A hybrid model for learning", *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 877-890, 1998.

[39] _____, "Learning in the framework of fuzzy lattices", *IEEE Trans. on Fuzzy Systems*, vol. 7, no. 4, pp. 422-440, 1999. *Errata in IEEE Trans. on Fuzzy Systems*, vol. 8, no. 2, pp. 236, 2000.

[40] _____, "Clustering and classification in structured data domains using Fuzzy Lattice Neurocomputing (FLN)", *IEEE Trans. on Knowledge and Data Engineering*, vol. 13, no. 2, pp. 245-260, 2001.

[41] _____, "*FINkNN*: A Fuzzy Interval Number k-Nearest Neighbor Classifier for prediction of sugar production from populations of samples", *Journal of Machine Learning Research*, vol. 4(Apr), pp. 17-37, 2003 (can be downloaded from http://www.jmlr.org).

[42] V. Petridis, A. Kehagias, L. Petrou, A. Bakirtzis, S. Kiartzis, H. Panagiotou, and N. Maslaris, "A Bayesian multiple models combination method for time series prediction", *Journal of Intelligent and Robotic Systems*, vol. 31, no. 1-3, pp. 69-89, 2001.

[43] H. Pomares, I. Rojas, J. Ortega, J. Gonzalez, and A. Prieto, "A systematic approach to a self-generating fuzzy rule-table for function approximation", *IEEE Trans. Systems, Man, and Cybernetics - B*, vol. 30, no. 3, pp. 431-447, 2000.

[44] A.T. Popov, "Convexity indicators based on fuzzy morphology", *Pattern Recognition Letters*, vol. 18, pp. 259-267, 1997.

[45] A.L. Ralescu, and D.A. Ralescu, "Probability and fuzziness", *Information Sciences*, vol. 34, pp. 85-92, 1984.

[46] M. Setnes, "Supervised fuzzy clustering for rule extraction", *IEEE Trans. Fuzzy Systems*, vol 8, no. 4, pp. 416-424, 2000.

[47] D. Sinha, and E.R. Dougherty, "A general axiomatic theory of intrinsically fuzzy mathematical morphologies", *IEEE Trans. Fuzzy Systems*, vol. 3, no. 4, pp. 389-403, 1995.

[48] R.R. Stoll, *Set Theory and Logic*. New York, NY: Dover Publications, Inc., 1979.

[49] N. Takagi, and K. Nakashima, "Discrete interval truth values logic and its application", *IEEE Trans. Computers*, vol. 49, no. 3, pp. 219-229, 2000.

[50] H. Tanaka, and H. Lee, "Interval regression analysis by quadratic programming approach", *IEEE Trans. Fuzzy Systems*, vol. 6, no, 4, pp. 473-481,1998.

[51] H. Tanaka, S. Uejima, and K. Asai, "Linear regression analysis with fuzzy model", *IEEE Trans. Systems, Man, and Cybernetics*, vol. 12, pp. 903-907, 1982.

[52] D. Tikk, and P. Baranyi, "Comprehensive analysis of a new fuzzy rule interpolation method", *IEEE Trans. Fuzzy Systems*, vol 8, no. 3, pp. 281-296, 2000.

[53] H. Vandendriessche, "Modelling growth and sugar accumulation of sugar beet (beta vulgaris L.) for potential production conditions", Ph.D. dissertation, Biologische Wetenschappen, Katholieke Universiteit Leuven, Belgium 1995.

[54] C.R. Webb, A.R. Werker, and C.A. Gilligan, "Modelling the dynamical components of the sugar-beet crop", *Annals of Botany*, vol. 80, no. 4, pp. 427-436, 1997.

[55] G.-D. Wu, C.-T. Lin, "A recurrent neural fuzzy network for word boundary detection in variable noise-level environments", *IEEE Trans. Systems, Man, and Cybernetics - B*, vol. 31, no. 1, pp. 84-97, 2001.

[56] M.-S. Yang, C.-H. Ko, "On cluster-wise fuzzy regression analysis", *IEEE Trans. Systems, Man, Cybernetics - B*, vol. 27, no. 1, pp. 1-13, 1997.

[57] J.-S. Yao, and F.-T. Lin, "Fuzzy critical path method based on signed distance ranking of fuzzy numbers", *IEEE Trans. Systems, Man, and Cybernetics - A*, vol. 30, no. 1, pp. 76-82, 2000. Erratum in *IEEE Trans. Systems, Man, and Cybernetics - A*, vol. 30, no. 2, p. 230, 2000.

[58] L.A. Zadeh, "Fuzzy sets", *Information and Control* , vol. 8, pp. 338-353, 1965.

[59] _____, "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets and Systems*, vol. 1, pp. 3-28, 1978.

[60] R. Zwick, E. Carlstein, and D.V. Budescu, "Measures of similarity among fuzzy concepts: A comparative analysis", *International Journal of Approximate Reasoning*, vol. 1, pp. 221-242, 1987.

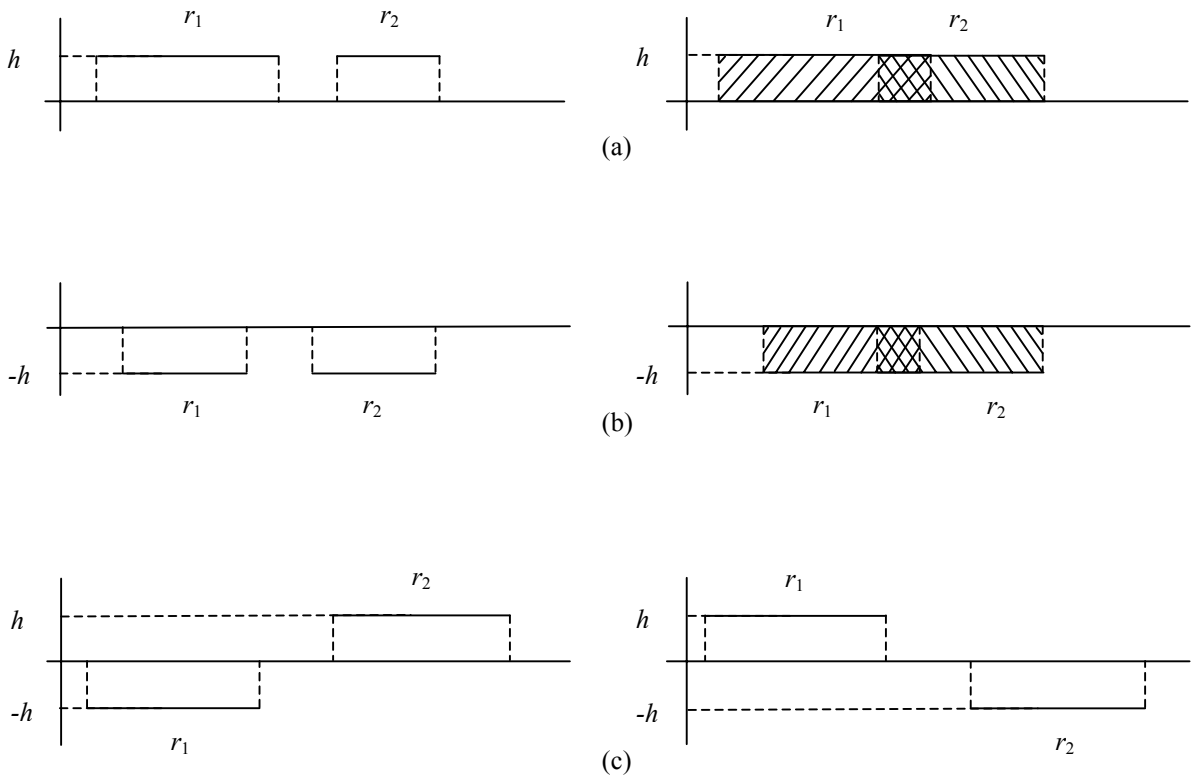


Fig. 1 Incomparable generalized intervals r_1 and r_2 in $\mathbb{M}^h$, symbolically $r_1 \|_{\mathbb{M}^h} r_2$, when

(a) $r_1, r_2 \in \mathbb{M}^h_+$ (positive generalized intervals),

(b) $r_1, r_2 \in \mathbb{M}^h_-$ (negative generalized intervals), and

(c) $r_1 \in \mathbb{M}^h_+$, $r_2 \in \mathbb{M}^h_-$ (one positive and one negative generalized interval).

Different fill-in patterns are shown for partially overlapped generalized intervals.

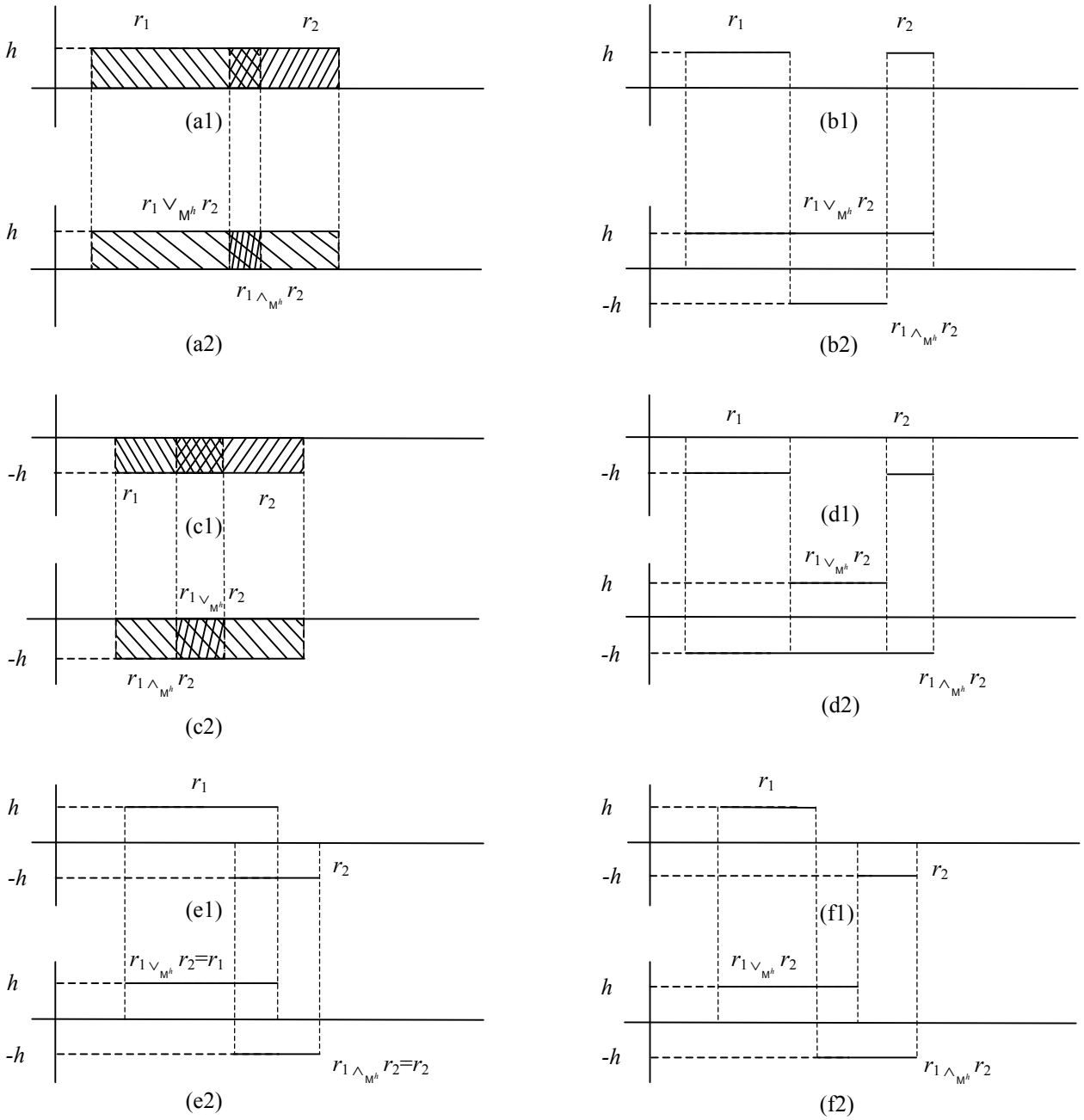


Fig. 2 The *join* ($r_1 \vee_{M^h} r_2$) and *meet* ($r_1 \wedge_{M^h} r_2$) for various pairs r_1 and r_2 of generalized intervals $r_1, r_2 \in \mathbb{M}^h$. Different fill-in patterns are used for partially overlapped generalized intervals.

(a) "Intersecting" positive generalized intervals,

(b) "Non-intersecting" positive generalized intervals,

(c) "Intersecting" negative generalized intervals,

(d) "Non-intersecting" negative generalized intervals,

(e) "Intersecting" positive & negative generalized intervals, and

(f) "Non-intersecting" positive & negative generalized intervals.

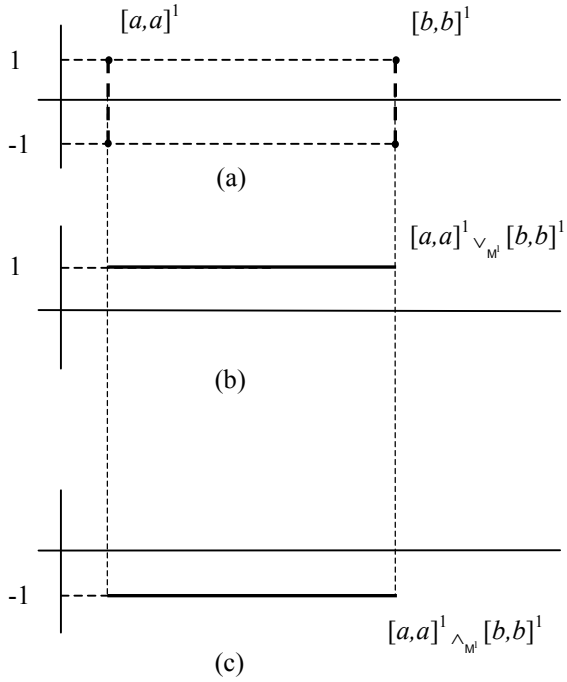
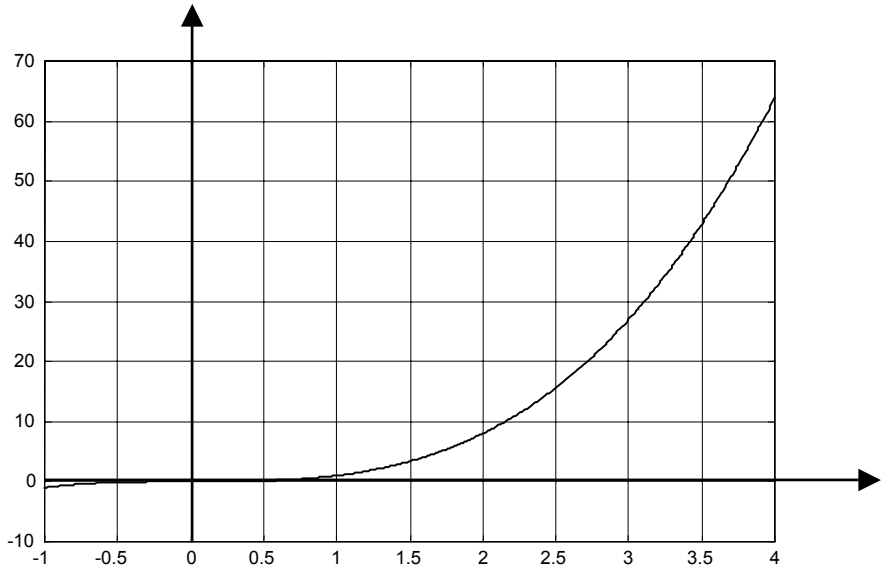


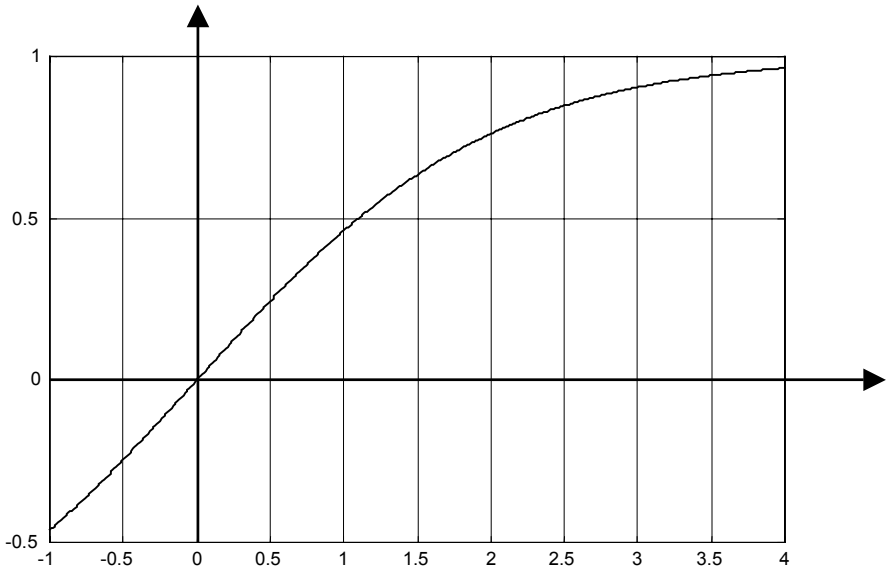
Fig. 3 (a) Trivial generalized intervals $[a,a]^1$ and $[b,b]^1$ in M_0^1,

(b) The lattice join $[a,a]^1 \vee_{M^1} [b,b]^1$ of trivial generalized intervals $[a,a]^1$ and $[b,b]^1$,

(c) The lattice meet $[a,a]^1 \wedge_{M^1} [b,b]^1$ of trivial generalized intervals $[a,a]^1$ and $[b,b]^1$.



(a)



(b)

Fig. 4 (a) A *cubic* underlying positive valuation function $f_1(x)=x^3$, and

(b) A *logistic* underlying positive valuation function $f_2(x)=\frac{2}{1+e^{-x}}-1$.

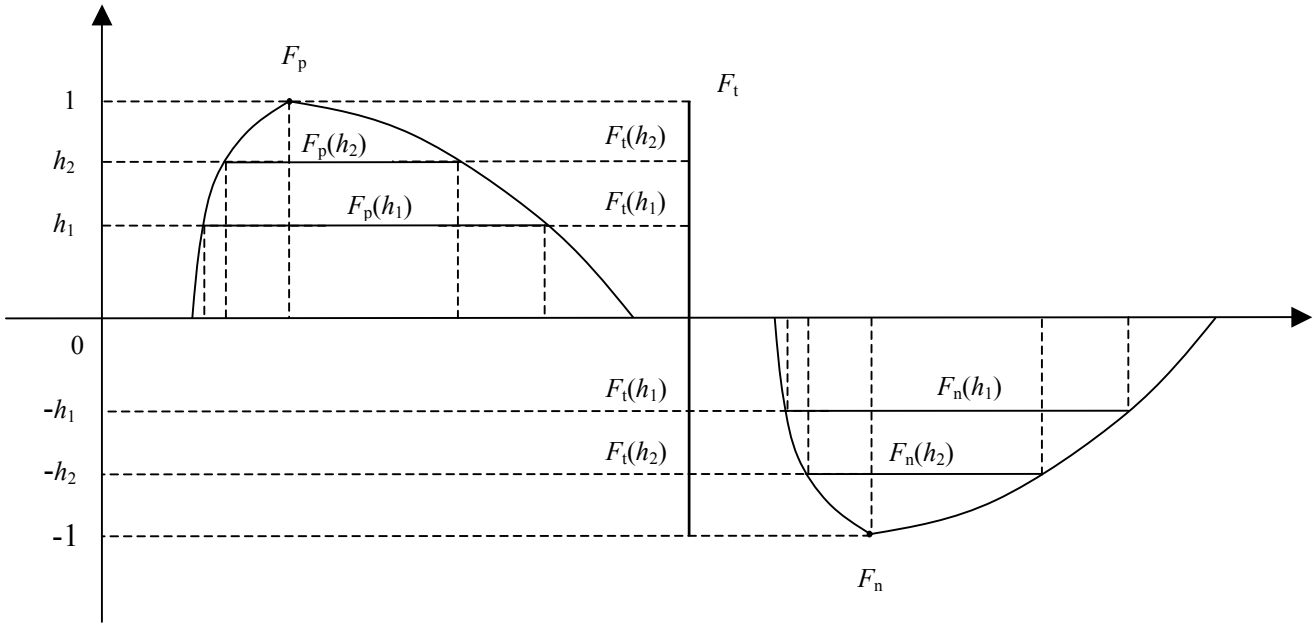


Fig. 5 A positive *FIN* F_p, a trivial *FIN* F_t and a negative *FIN* F_n.

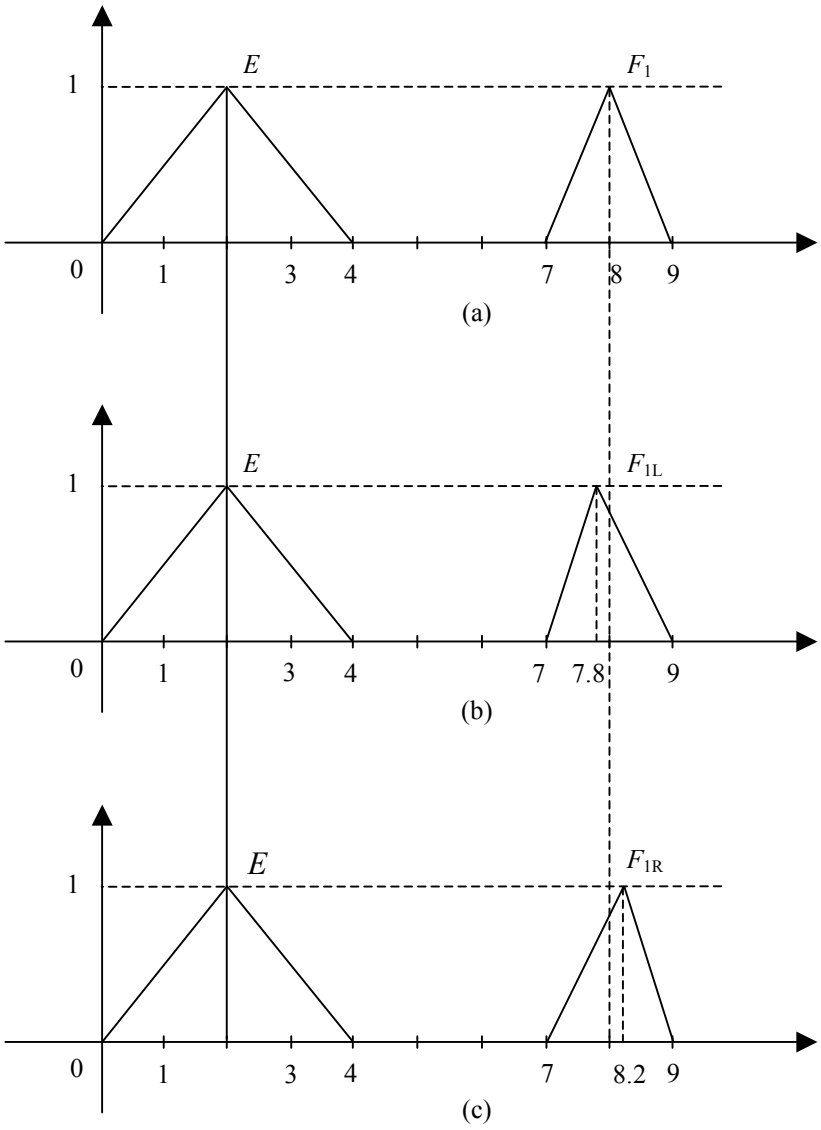
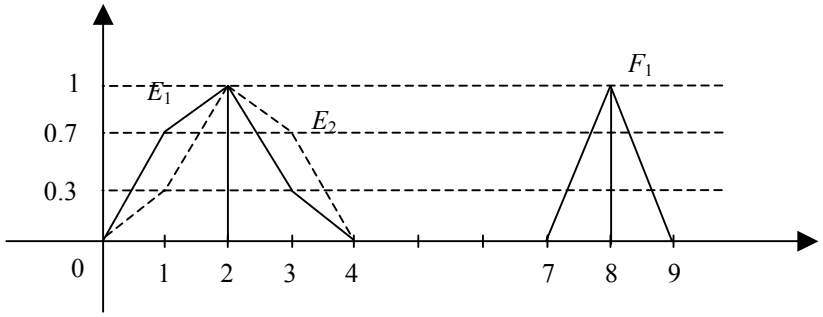


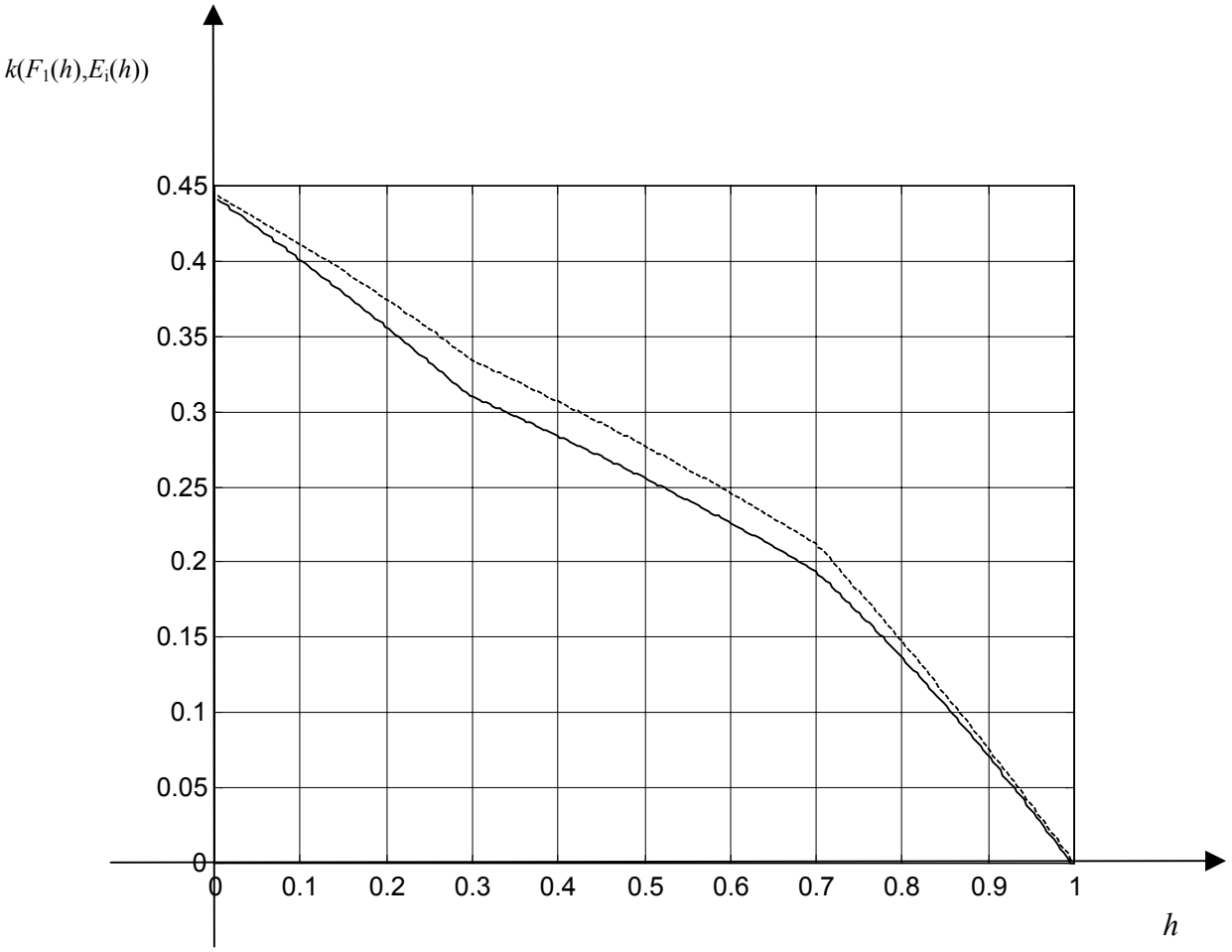
Fig. 6 (a) The degree of inclusion of *FIN* F_1 into *FIN* E is $\sigma_K(F_1 \leq_F E) \approx 0.2512$.

(b) The degree of inclusion of *FIN* F_{1L} into *FIN* E is $\sigma_K(F_{1L} \leq_F E) \approx 0.2537$.

(c) The degree of inclusion of *FIN* F_{1R} into *FIN* E is $\sigma_K(F_{1R} \leq_F E) \approx 0.2488$.



(a)



(b)

Fig. 7 (a) *FIN* F_1 is included more in *FIN* E_2 (dashed membership function) than it is in *FIN* E_1 (solid membership function), in particular $\sigma_K(F_1 \leq_F E_2) \approx 0.2602$ and $\sigma_K(F_1 \leq_F E_1) \approx 0.2465$.

(b) The degrees of inclusion of interval $F_1(h)$ into either intervals $E_1(h)$ (solid line) or $E_2(h)$ (dashed line) versus h for $h \in (0,1]$.

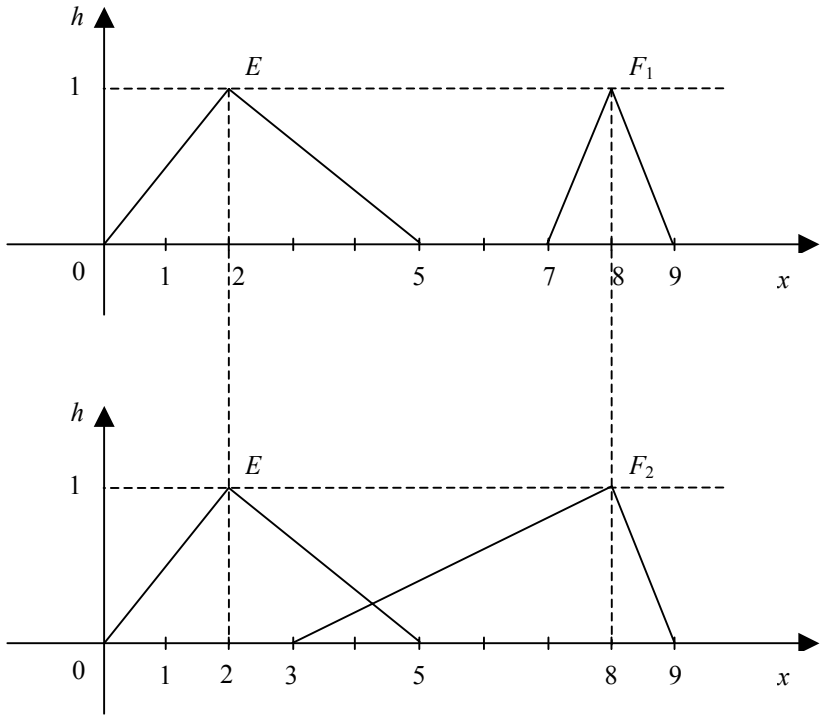


Fig. 8 In the above figure it holds $\sigma_K(F_1 \leq_F E) = 0.3140 = \sigma_K(F_2 \leq_F E)$, nevertheless $d_K(E, F_2) \approx 5.1667 \leq 5.8334 \approx d_K(E, F_1)$ as explained in the text. It follows that metric d_K is more dependable than fuzzy inclusion measure σ_K for quantifying the "relative position" between *FIN*s.

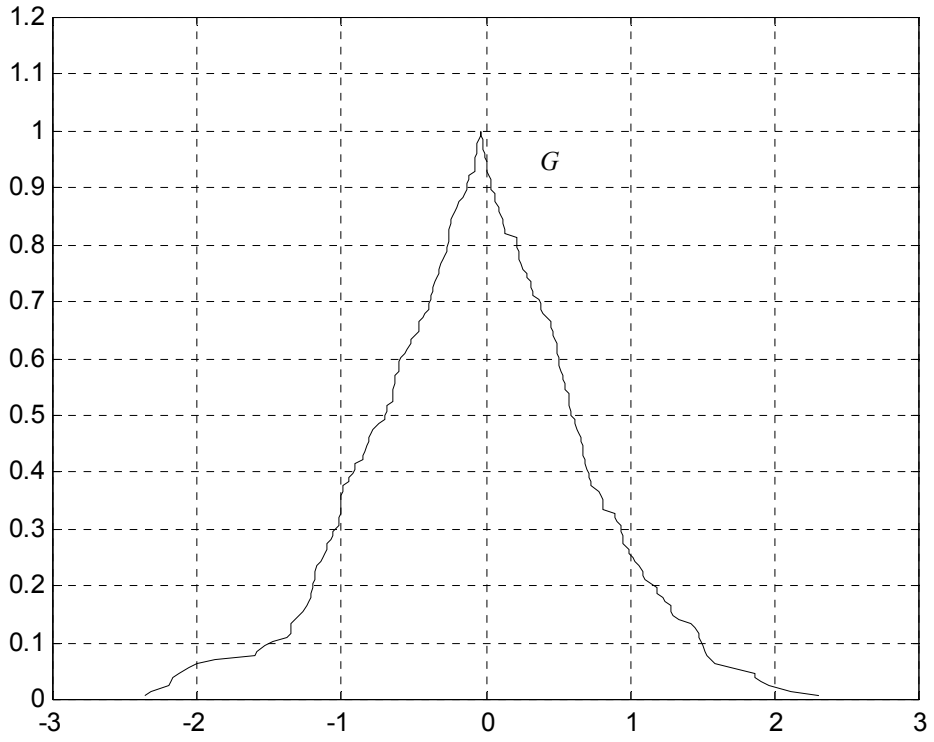


Fig. 9 *FIN G* in this figure was constructed by applying algorithm CALFIN to 222 random numbers generated according to the normal (Gaussian) probability density function N(0,1) with *mean* 0 and *standard deviation* 1.

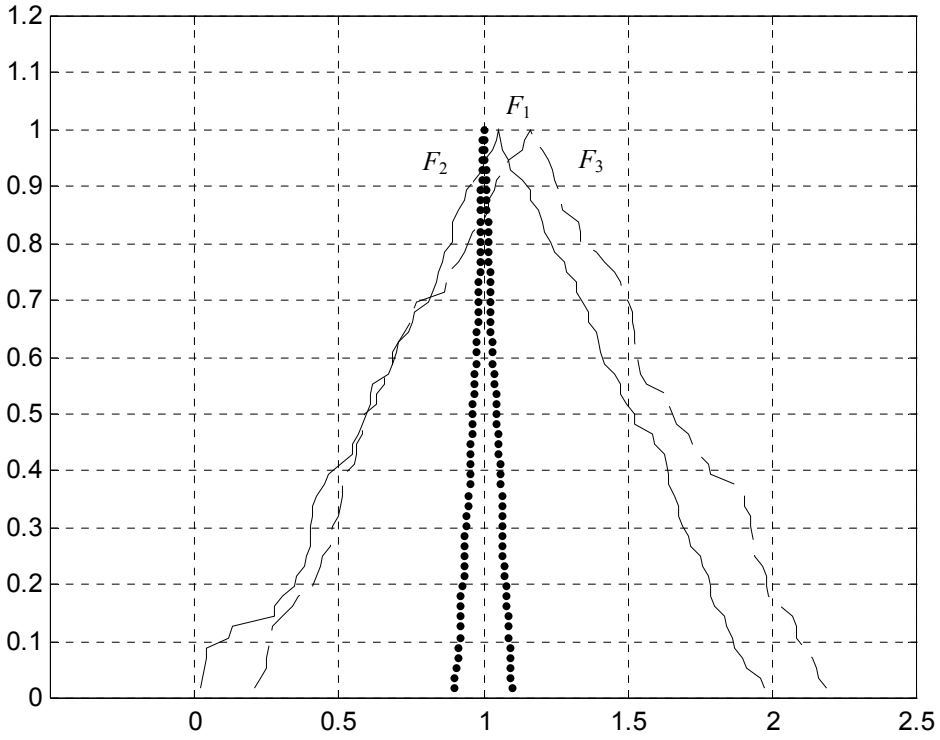
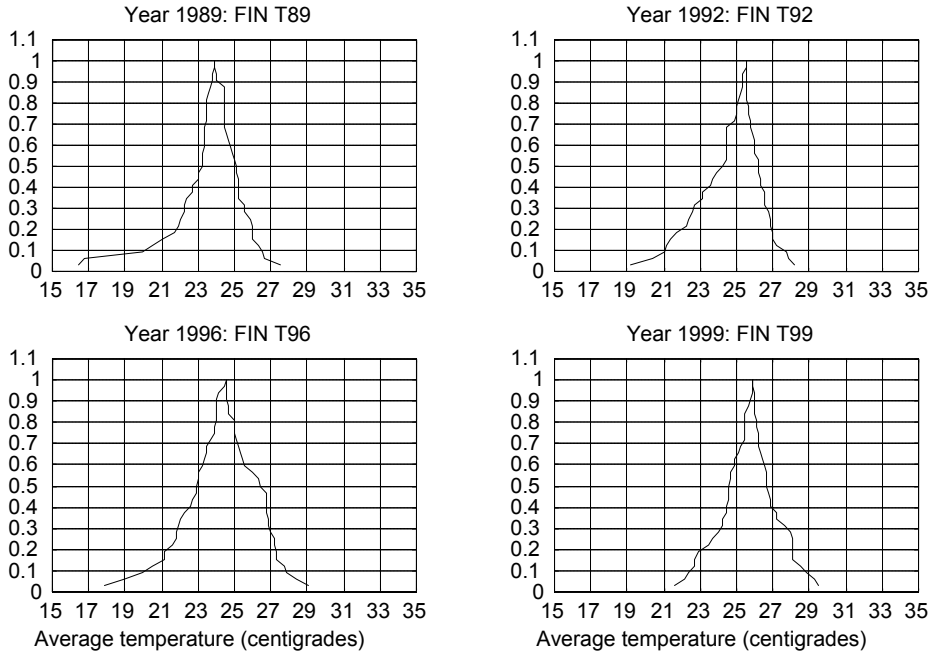
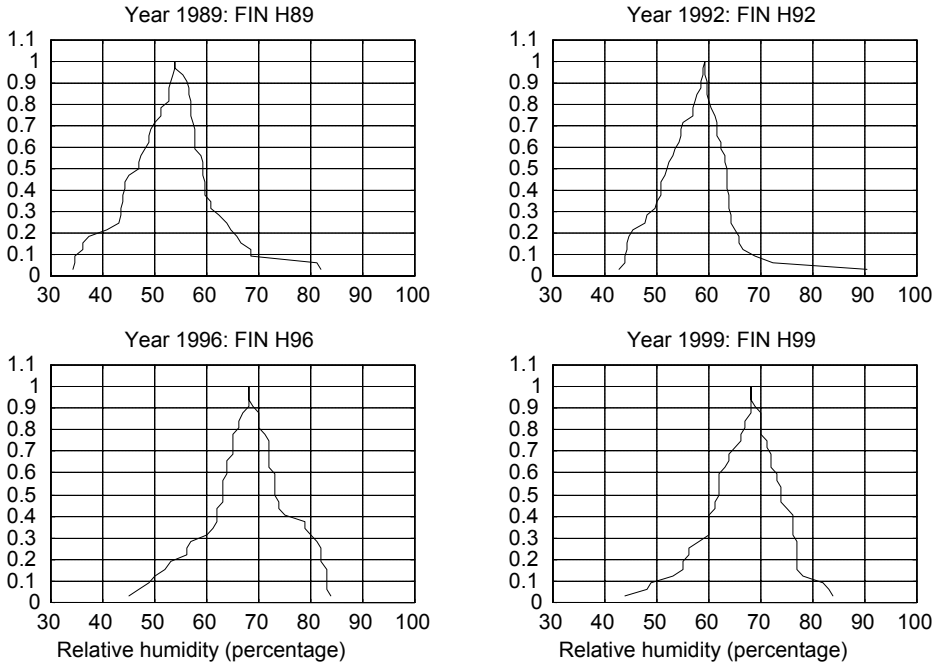


Fig. 10 The three *FIN*s F_1 (solid line), F_2 (dotted line), and F_3 (dashed line) above have been computed from three populations of samples generated randomly according to the uniform probability distribution with ranges [0, 2], [0.9, 1.1] and [0.2, 2.2], respectively. It has been computed $d_K(F_1,F_2)= 0.2843$, $d_K(F_2,F_3)= 0.3480$ and $d_K(F_1,F_3)= 0.1002$.



(a)



(b)

Fig. 11 (a) Four *FIN*s *T89*, *T92*, *T96* and *T99* produced from daily average temperatures for July and August of years 1989, 1992, 1996 and 1999, respectively, at the Thessaloniki area, Greece.

(b) Four *FIN*s *H89*, *H92*, *H96* and *H99* produced from daily relative humidities for July and August of years 1989, 1992, 1996 and 1999, respectively, at the Thessaloniki area, Greece.

TABLE I
DISTANCES d_K BETWEEN *FIN*S *T89*, *T92*, *T96* AND *T99* SHOWN IN FIG.11(a)

	T89 (Year 1989)	*T92* (Year 1992)	*T96* (Year 1996)	*T99* (Year 1999)
T89 (Year 1989)	0	1.2472	0.5821	1.8213
T92 (Year 1992)	1.2472	0	0.8528	0.5741
T96 (Year 1996)	0.5821	0.8528	0	1.3249
T99 (Year 1999)	1.8213	0.5741	1.3249	0

TABLE II
DISTANCES d_K BETWEEN *FIN*S *H89*, *H92*, *H96* AND *H99* SHOWN IN FIG.11(b)

	H89 (Year 1989)	*H92* (Year 1992)	*H96* (Year 1996)	*H99* (Year 1999)
H89 (Year 1989)	0	4.6267	14.7569	14.3932
H92 (Year 1992)	4.6267	0	10.1907	9.8272
H96 (Year 1996)	14.7569	10.1907	0	0.8226
H99 (Year 1999)	14.3932	9.8272	0.8226	0

TABLE III
APPROXIMATE % SUGAR PRODUCTION PREDICTION ERROR FOR VARIOUS METHODS ARRANGED IN THE CHRONOLOGICAL ORDER OF THEIR PUBLICATION

Prediction Method	Approximate % Prediction Error
Bayesian Combined Predictors (BCP)	6.0
First principles model	15.0
Intelligent clustering techniques	5.0
Classifier *FINkNN*	2.0