

Una Verificación Empírica de Modelos Automatizables de Estimación muy Temprana de Proyectos de Desarrollo de Sistemas de Gestión

Pedro Salvetto ¹, Julio Fernández ¹, Juan Carlos Nogueira ¹, Javier Segovia ²

¹Universidad ORT Uruguay
Cuareim 1451, Código Postal 11100, Montevideo, Uruguay
{salvetto, fernandez_j, nogueira}@ort.edu.uy

² Universidad Politécnica de Madrid, Departamento de Lenguajes Sistemas Informáticos e Ingeniería de Software, Campus de Montegancedo 28660 Boadilla del Monte (Madrid),
Facultad de Informática
fsegovia@fi.upm.es

Resumen En este trabajo estudiamos modelos formales de estimación que no requieren intervención humana. Comparamos los resultados de modelos de estimación basadas en métricas tempranas con los de modelos que consideran, además, métricas disponibles post mortem. Concluimos con un 99% de confianza que no hay diferencias significativas entre ellos, confirmando lo anticipado por DeMarco de que los sistemas de información data-strong pueden estimarse a partir de los datos. Confirmamos la bondad de las métricas propuestas por Calero et al para bases de datos relacionales. Comparando las mediciones finales y las estimaciones tempranas llegamos a un resultado análogo, encontrando que a efectos prácticos no existe diferencia.

Palabras Clave: modelos empíricos, modelos de estimación, verificación empírica, modelos formales

DCXW	Peso de la Complejidad de los datos (escalar)
ETT	Modelo de Estimación Temprana de tiempo (meses)
EPMT	Modelo de Estimación Post Mortem de Tiempo (meses)
ETE	Modelo de Estimación Temprana de Esfuerzo (meses hombre)
EPME	Modelo de Estimación Post Mortem de Esfuerzo (meses hombre)
IF	Número de estructuras de control if en la especificación
FOR	Número de estructuras de control for en la especificación
IFAN	Número de estructuras de control if anidadas en la especificación
FORAN	Número de estructuras de control for anidadas en la especificación
DRT	Profundidad del árbol referencial
RD	Grado de Referencialidad, número claves foráneas
NA	Número de atributos únicos en el esquema de base de datos relacional
NT	Número de tablas en el esquema de bases de datos relacional
EIM	Esfuerzo inicial medio del 10% inicial del proyecto, cantidad promedio De personas asignadas al proyecto en el período
RV	Volatilidad de los requerimientos (porcentual)

Tabla 1 Principales siglas y acrónimos usados en este artículo

1 Introducción

En trabajos anteriores publicamos modelos de estimación muy temprana del tiempo y esfuerzo de desarrollo [7]. Estos modelos se basan en la complejidad del sistema a desarrollar, la volatilidad de los requerimientos, la eficiencia de la organización y la rapidez de ejecución del proyecto de desarrollo.

Su característica más destacable es que producen la estimación sin intervención humana ya que las métricas usadas pueden ser obtenidas automáticamente y no requieren juicio experto [8,6].

La complejidad fue medida desde el esquema de base de datos relacional resultante de la integración automática de las vistas de datos de los usuarios finales del sistema a construir. Para ello usamos las métricas de complejidad de bases de datos definidas por Calero, Piattini, Polo y Ruiz [1]. Conjuntamente con estos modelos definimos una métrica de complejidad esencial de sistemas de información (DCXW ver (1) más adelante).

Luego definimos modelos que se basan en la DCXW, la volatilidad de los requerimientos y la rapidez de ejecución del proyecto. Estos modelos tuvieron un comportamiento similar y en algunos casos mejor que los originales.

Esta métrica de complejidad resulta muy importante porque nos permite 1) interpretar mejor los datos y 2) medir la complejidad de un sistema de información con una cifra independiente de la tecnología y el juicio experto..

Ya hace tiempo, DeMarco [3] propuso que los sistemas que definió como data-strong fueran estimados a partir de la complejidad de los datos. En este trabajo intentamos una verificación empírica. Para responder esta pregunta desarrollamos varios experimentos tendientes a estudiar si en métricas que pueden ser tomadas post

mortem hay una cantidad significativamente mayor de información relativa a la complejidad del sistema a construir que la disponible muy tempranamente en las vistas de datos de los usuarios, desde el punto de vista de la estimación del proyecto.

En las secciones siguientes presentaremos los modelos de estimación temprana, la métrica de estimación temprana de complejidad, identificaremos métricas de complejidad post mortem, presentaremos modelos post mortem y estudiaremos la contribución de las métricas post mortem a la estimación

2 Nuestras hipótesis de trabajo

- 1) Si sólo se desarrolla funcionalidad no redundante y de valor para el negocio o para algún usuario, *la complejidad esencial del sistema a desarrollar queda determinada por la integración de las visiones de datos de los usuarios del sistema y ésta puede medirse a partir del esquema de bases de datos relacionales generado automáticamente a partir de la integración de las visiones de datos de los usuarios.*
- 2) El esquema de base de datos relacional así obtenido contiene información suficiente como para ser usada como entrada de modelos de estimación de tiempo y esfuerzo de desarrollo.
- 3) Aunque no comprendamos en profundidad los complejos procesos internos e interacciones que ocurren durante el proceso de desarrollo, podemos observarlo y construir modelos que estimen aceptablemente su resultado global. En particular, si el desarrollo se realiza usando herramientas que automaticen la generación del código, la integración de módulos y apoyen el trabajo de grupos reducidos con usuarios integrados a los mismos usando una metodología estándar, se estará reduciendo las fuentes de variabilidad y será posible encontrar modelos útiles para la estimación temprana.

3 Los proyectos observados

Los proyectos fueron desarrollados con metodologías ágiles, ciclo de vida evolutivo y herramientas de especificación formal y generación de aplicaciones a partir de su especificación. Todos los sistemas se basaron en las vistas de datos y solicitudes de los usuarios. Se trabajó con equipos de desarrollo reducidos (2 a 5 personas) con usuarios integrados a los mismos. Se desarrollaron utilizando tecnología de bases de datos relacionales. Los sistemas estudiados son de gestión del tipo data-strong [3] donde los algoritmos de alta complejidad o no existen o son excepcionales. En todos los casos, las métricas de entrada de nuestros modelos fueron tomadas en forma automática con la herramienta de que disponemos, a partir de las especificaciones.

La volatilidad de los requerimientos fue estimada por los gerentes de proyecto y el tiempo y esfuerzo requeridos para el desarrollo de cada proyecto fueron

proporcionados por los gerentes de proyecto. Se estudio una muestra de 20 proyectos. El detalle de la información relevada puede encontrarse en la tabla 5.

La metodología de desarrollo y la herramienta de especificación formal¹

La metodología asociada a la herramienta se apoya en dos premisas. Primero en una organización mediana o grande, *nadie tiene una visión global de los datos y los procesos*. Por tanto, se requiere integrar las visiones de diferentes usuarios. Segundo, *los requerimientos y las estructuras de bases de datos sufren cambios continuos e inevitables a lo largo del tiempo*. La herramienta almacena información en una “base de conocimiento” (KB)². Estos datos son de gran utilidad para capturar métricas automáticamente.

La herramienta usada para el desarrollo tiene la capacidad de automatizar 1) el trabajo con los usuarios en el ámbito de sus vistas de datos no normalizadas de la realidad; 2) la integración de estas vistas de datos de usuario; 3) la generación del esquema relacional que representa la integración de las vistas de datos de usuario; 4) la determinación del impacto de un cambio en el esquema relacional; 5) la generación de los programas necesarios para migrar los datos cuando se produce un cambio en las vistas de datos de usuario que impacta en el esquema relacional; 6) la generación del código en diversos lenguajes y plataformas a partir de la especificación (independizándonos por tanto de la tecnología); 7) la integración de varias especificaciones en una sola, detectando las posibles inconsistencias (que obviamente tendrán que levantar los desarrolladores); 8) la publicación de la especificación posibilitando la automatización de la obtención de métricas [8, 6, 5] y 9) el manejo de un nombre único para cada atributo.

Se analizaron proyectos desarrollados con una misma metodología de trabajo estable y equipos de desarrollo de similares características, por lo que la eficiencia es constante dentro de la muestra.

4 Los modelos y la métrica de complejidad esencial de Sistemas de Información

4.1 Medición de la complejidad

La complejidad se midió a partir del esquema de base de datos relacional generado automáticamente por la herramienta al integrar las vistas de datos de los usuarios finales. Para ello se usaron las métricas de complejidad de bases de datos relacionales propuestas por Calero, Piattini, Polo y Ruiz [1]. Estas son: profundidad del árbol referencial (DRT), grado de referencialidad (número de claves foráneas) (RD) y número de atributos únicos (NA). Nosotros agregamos el número de tablas (NT). La automatización del conteo del número de atributos únicos [3, 4] fue posible porque la

¹ GeneXus desarrollada por Artech www.artech.com

² En este artículo, siguiendo la terminología estándar de Genexus, KB (Knowledge Base) refiere a un conjunto de objetos que contiene la especificación del sistema.

herramienta impone una metodología donde *si dos atributos de cualquier vista de datos de usuario se llaman igual representan lo mismo*.

4.2 Los modelos construidos

La volatilidad de los requerimientos (RV) se midió porcentualmente considerando tanto las muertes como los nacimientos de requerimientos y el Esfuerzo Medio Inicial (EMI) se midió como esfuerzo por unidad de tiempo, considerando el primer 10% del tiempo de ejecución del proyecto. El tiempo se midió en meses y el esfuerzo en meses hombre.

Como se explico en trabajos anteriores [7,8] la métrica DCXW contempla el impacto sobre el tiempo y esfuerzo de las métricas DRT, RD, NA y NT

$$DCXW = DRT^{-2,4} RD^{-0,3} NA^{1,8} NT^{-0,15} \quad (1)$$

Los modelos de estimación temprana relacionan ETT y ETE con DCXW, EMI y RV.

Los modelos de estimación post mortem toman en cuenta la complejidad de las estructuras de control presentes en la especificación incorporando las variables IF, IFAN, FOR y FORAN; que representan respectivamente el número de if, if anidados, for y for anidados presentes en la especificación.

Tiempo en Meses	$ETT = 0,02 DCXW^{1,08} EMI^{0,7} RV^{0,15}$
Esfuerzo en Meses Hombre	$ETE = 0,01 DCXW^{1,02} EMI^{0,29} RV^{0,5}$

Tabla 2 Modelos de estimación temprana de tiempo y esfuerzo

Tiempo en Meses	$EPMT = 0,02 DCXW^{1,07} EMI^{-0,71} RV^{0,18} IF^{0,03} IFAN^{-0,06} FOR^{0,004} FORAN^{0,03}$
Esfuerzo en Meses Hombre	$EPME = 0,01 DCXW^{0,98} EMI^{0,35} RV^{0,51} IF^{-0,05} IFAN^{0,05} FOR^{0,002} FORAN^{0,01}$

Tabla 3 Modelos de estimación Post mortem de tiempo y esfuerzo

4.3 Comparación de los Modelos Tempranos vs modelos Post Mortem

4.3.1 Comparación de acuerdo a los criterios de Conte et al

La tabla 4 muestra la capacidad de predicción de estos modelos según los criterios de Conte et al [2] y además medidas de tendencia central y de dispersión de los

valores absolutos de los errores relativos de la estimación para la muestra usada en el experimento (20 proyectos). La capacidad de predicción a un nivel k (PRED (k)) se define como la proporción de veces que se comete un error relativo que en valor absoluto es menor o igual a k y el MRE se define como el promedio de los valores absolutos de los errores relativos. En base a estos dos elementos, se asigna una calificación a cada modelo.

Como esperábamos, se observa que los modelos post mortem estiman mejor que los tempranos. Sin embargo se observa que la introducción, de las variables post mortem no genera mejoras importantes. Por ejemplo si consideramos el esfuerzo, el PRED (20) aumenta de 0,95 a 1 por una única observación que mejora su error relativo. El PRED (k), para nuestro caso no es muy significativo ya que por lo reducido de la muestra una observación representa un 5%.

MODELOS DE ESTIMACIÓN															
ERRORES RELATIVOS EN VALOR ABSOLUTO															
PRED(k)								MIN	MAX	MRE	DESV EST	R2	CAL		
RE:	5%	10%	15%	20%	25%	30%	35%								
ESFUERZO															
ETE	Frecuencia	6	3	6	4	0	1	0,2%	25,9%	10,1%	7,0%	0,9742	Ex		
	% Acumulado	0,30	0,45	0,75	0,95	0,95	1,00								
EPME	Frecuencia	8	4	5	3			0,3%	18,5%	8,3%	5,7%	0,9864	Ex		
	% Acumulado	0,40	0,60	0,85	1,00										
TIEMPO															
ETT	Frecuencia	6	4	4	1	2	2	1	1	0,2%	33,6%	12,2%	10,0%	0,9509	B
	% Acumulado	0,30	0,50	0,70	0,75	0,85	0,95	1,00	1,00						
EPMT	Frecuencia	6	11	2	0	1			0,1%	23,7%	6,4%	5,4%	0,9886	Ex	
	% Acumulado	0,30	0,85	0,95	0,95	1,00									

Tabla 4 Valor predictivo de modelos de estimación de esfuerzo y tiempo

4.4 Análisis Estadístico³

Para estudiar con mayor formalidad el efecto de las variables post mortem en la estimación consideraremos al proyecto de desarrollo como un experimento que medimos al comienzo (estimación temprana) y luego de aplicarle un tratamiento (ejecución del proyecto). Luego de la ejecución del proyecto disponemos de mayor información, lo que nos permite realizar la estimación post mortem con más variables. Finalmente estudiaremos cual es el efecto que tiene el tratamiento sobre los errores relativos de estimación. Ya hemos visto que los errores relativos mejoran con el tratamiento. Deseamos responder la siguiente pregunta:

¿Con que nivel de confianza puede afirmarse que las diferencias entre las mediciones finales, las estimaciones tempranas y las estimaciones post mortem se debe al azar?

³ El análisis de este apartado se realizó siguiendo la metodología sugerida por Juristo y Moreno [4] usando SPSS 11.0 para Windows

La tabla 5 muestra el esfuerzo y tiempo observados, los valores estimados por los modelos, y sus errores absolutos y relativos.

Estadísticos Descriptivos

Inicialmente realizamos un estudio estadístico descriptivo cuyos resultados se presentan en la tabla 6.

Independencia de los errores relativos respecto de la magnitud del esfuerzo y tiempo.

En la fig. 1 se grafican los errores relativos vs la magnitud de la variable estimada (esfuerzo o tiempo). Se observa que es mayor para los proyectos que insumieron menos tiempo y esfuerzo (más pequeños). Sin embargo no se aprecia ningún patrón

N O	ESFUERZO	ETE	EPME	ERRORES				TIEMPO	ETT	EPMT	ERRORES			
				ABSOLUTOS		RELATIVOS					ABSOLUTOS		RELATIVOS	
				EAETE	EAEPME	ERETE	EREPME				EAETT	EAEPMT	ERETT	EREPMT
1	70	68,658	73,442	-1,342	3,442	-0,019	0,049	25	25,222	22,828	0,222	-2,172	0,009	-0,087
2	14	15,723	13,079	1,723	-0,921	0,123	-0,066	6	3,983	4,663	-2,017	-1,337	-0,336	-0,223
3	12	11,947	11,526	-0,053	-0,474	-0,004	-0,040	4	5,154	5,112	1,154	1,112	0,289	0,278
4	26	28,313	29,333	2,313	3,333	0,089	0,128	7	8,488	8,416	1,488	1,416	0,213	0,202
5	5	5,906	5,613	0,906	0,613	0,181	0,123	4	3,669	3,518	-0,331	-0,482	-0,083	-0,120
6	17	15,258	17,044	-1,742	0,044	-0,102	0,003	3	3,341	3,093	0,341	0,093	0,114	0,031
7	15	17,668	17,211	2,668	2,211	0,178	0,147	3	2,580	2,618	-0,420	-0,382	-0,140	-0,127
8	42	31,129	34,218	-10,871	-7,782	-0,259	-0,185	7	5,469	5,343	-1,531	-1,657	-0,219	-0,237
9	4	4,167	4,199	0,167	0,199	0,042	0,050	2	1,955	1,937	-0,045	-0,063	-0,023	-0,031
10	20	23,429	19,308	3,429	-0,692	0,171	-0,036	3	3,039	3,282	0,039	0,282	0,013	0,094
11	7	7,980	8,122	0,980	1,122	0,140	0,160	2	2,524	2,472	0,524	0,472	0,262	0,236
12	67	59,048	67,349	-7,952	0,349	-0,119	0,005	10	11,084	10,337	1,084	0,337	0,108	0,034
13	7	7,012	6,761	0,012	-0,239	0,002	-0,034	2	1,941	2,108	-0,059	0,108	-0,030	0,054
14	7	5,836	6,049	-1,164	-0,951	-0,166	-0,136	4	3,623	3,286	-0,377	-0,714	-0,094	-0,179
15	61	65,353	70,511	4,353	9,511	0,071	0,156	11	12,339	11,359	1,339	0,359	0,122	0,033
16	10	9,599	9,023	-0,401	-0,977	-0,040	-0,098	2	2,029	2,243	0,029	0,243	0,015	0,122
17	61	55,495	57,808	-5,505	-3,192	-0,090	-0,052	11	13,062	11,462	2,062	0,462	0,189	0,042
18	6	5,370	5,348	-0,630	-0,652	-0,105	-0,109	2	2,194	2,062	0,194	0,062	0,097	0,031
19	7	6,957	7,470	-0,043	0,470	-0,006	0,067	2	2,183	2,002	0,183	0,002	0,092	0,001
20	31	34,204	30,548	3,204	-0,452	0,103	-0,015	31	31,052	32,246	0,052	1,246	0,002	0,040
MEDIA	24,450	23,953	24,698	-0,497	0,248	0,009	0,006	7,050	7,248	7,019	0,198	-0,031	0,030	0,010
DESV EST	22,830	21,606	23,658	3,782	3,195	0,124	0,102	7,837	8,019	7,840	0,958	0,911	0,157	0,142

Tabla 5 Observaciones y Valores Estimados de Tiempo y Esfuerzo

Descriptive Statistics													
	N	Rang	Min	Max	Sum	Mean	Std.	Std.	Varia	Skewness	Kurtosis	Std.	Std.
	Stati stic	Stati stic	Stati stic	Stati stic	Stati stic	Stati stic	Error	Stati stic	Stati stic	Stati stic	Error	Stati stic	Error
ESFUERZO	20	66.00	4.00	70.00	489	24.45	5.10	22.83	521	1,101	.512	-.310	.992
EST TEMP ESF	20	64.49	4.17	68.66	479	23.95	4.83	21.61	467	1,104	.512	-.143	.992
EST POST MORTEM ESF	20	69.24	4.20	73.44	494	24.70	5.29	23.66	560	1,208	.512	.041	.992
TIEMPO	20	29.00	2.00	31.00	141	7.05	1.75	7.84	61,4	2,261	.512	4,856	.992
EST TEMP TIEMPO	20	29.11	1.94	31.05	145	7.25	1.79	8.02	64,3	2,098	.512	4,028	.992
EST POST MORTEM TIEMPO	20	30.31	1.94	32.25	140	7.02	1.75	7.84	61,5	2,329	.512	5,502	.992
ERR ABS ETE	20	15.22	-10.9	4.35	-9.95	-.50	.85	3.78	14,3	-1,47	.512	2,279	.992
ERR ABS EPME	20	17.29	-7.78	9.51	4.96	.25	.71	3.19	10,2	.536	.512	4,696	.992
ERR REL ETE	20	.44	-.26	.18	.19	.01	.03	.12	.015	-.352	.512	-.558	.992
ERR REL EPME	20	.35	-.19	.16	.12	.01	.02	.10	.010	.022	.512	-.871	.992
ERR ABS ETT	20	4.10	-2.02	2.08	3.95	.20	.21	.96	.917	-.313	.512	.914	.992
ERR ABS EPMT	20	3.59	-2.17	1.42	-.61	-.03	.20	.91	.829	-.758	.512	.674	.992
ERR REL ETT	20	.62	-.34	.29	.60	.03	.04	.16	.025	-.463	.512	.278	.992
ERR REL EPMT	20	.51	-.24	.28	.19	.01	.03	.14	.020	.006	.512	-.339	.992
Valid N (listwise)	20												

Tabla 6 Estadísticos descriptivos de esfuerzo y modelos de estimación

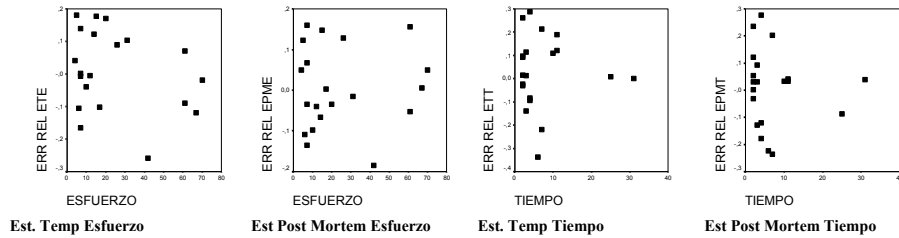


Fig. 1 Errores relativos de la estimación vs variable estimada

Normalidad de los errores relativos

En la fig 2 y tabla 7 se aprecia la normalidad de los errores relativos de las estimaciones de esfuerzo y tiempo. Los errores relativos de la estimación post mortem de tiempo, tienen una distribución que se aparta algo de la normalidad. Esto se observa gráficamente y en la menor significación de la prueba de Kolmogorov-Smirnov (tabla 7).

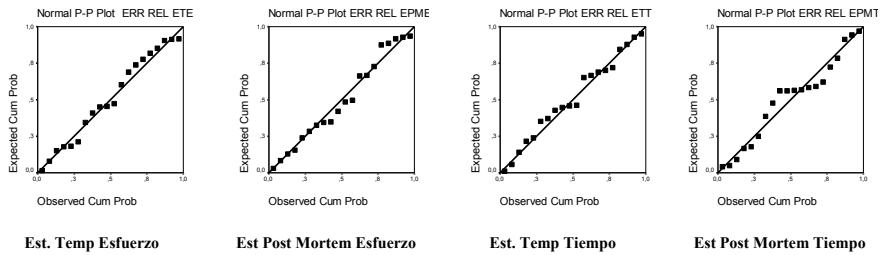


Fig 2 Normalidad de los Errores relativos de Estimación

One-Sample Kolmogorov-Smirnov Test

			ERR REL ETE	ERR REL EPME	ERR REL ETT	ERR REL EPMT
N			20	20	20	20
Normal Parameters	a,b	Mean	,0095	,0060	,0300	,0097
		Std. Deviation	,1241	,1019	,1573	,1420
Most Extreme Differences		Absolute	,0910	,1239	,1026	,1597
		Positive	,0891	,1034	,0886	,1279
		Negative	-,0910	-,1239	-,1026	-,1597
Kolmogorov-Smirnov Z		,4070	,5541	,4587	,7141	
Asymp. Sig. (2-tailed)		,9964	,9186	,9845	,6876	

a. Test distribution is Normal.

b. Calculated from data.

Tabla 7 Prueba de Normalidad de Kolmogorov-Smirnov para los errores relativos de estimación

Igualdad de las medias La alta significación que se observa en la tabla 8 hace pensar en la igualdad de las medias entre los grupos.

Robust Tests of Equality of Means					
		Statistic ^a	df1	df2	Sig.
ESFUERZO	Welch	,006	2	37,946	,994
	Brown-Forsythe	,006	2	56,691	,994
TIEMPO	Welch	,005	2	37,996	,995
	Brown-Forsythe	,005	2	56,973	,995

a. Asymptotically F distributed.

Tabla 8 Prueba de Igualdad de las Medias de Esfuerzo y Tiempo

Análisis de Varianza

En la tabla 9 se presenta el análisis de varianza considerando los tres grupos formados por las mediciones reales, las estimaciones tempranas y las estimaciones post mortem, tanto para tiempo como para esfuerzo. Se puede observar que la mayoría de la varianza se explica dentro de los grupos y no entre los grupos, lo que aunado a la muy alta significación, los estadísticos descriptivos de la tabla 6 y que la prueba de Levene de homogeneidad de la varianza dio una significación de 0,917 para esfuerzo y 0,958 para tiempo nos hace pensar que es probable que las diferencias existentes entre los grupos no se expliquen por el tratamiento y se deban al azar.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
ESFUERZO	Between Groups	5,766	2	2,883	,006	,994
	Within Groups	29406,991	57	515,912		
	Total	29412,757	59			
TIEMPO	Between Groups	,614	2	,307	,005	,995
	Within Groups	3556,672	57	62,398		
	Total	3557,286	59			

Tabla 9 ANOVA Esfuerzo Y Tiempo

Prueba T para muestras emparejadas

En la tabla 10 se presenta la prueba de muestras emparejadas entre esfuerzo y tiempo y sus estimaciones obtenidas de los modelos de estimación temprana y post mortem. Considerando a) que $t_{0,05,19}=2,861$, b) que ninguno de los valores de t es mayor en valor absoluto a este valor, c) la alta significación y d) que el intervalo en que con un 99% de confianza se encuentra la diferencia contiene al 0 para todos los

pares estudiados no podemos concluir que haya diferencias significativas entre: 1) esfuerzo y estimación temprana de esfuerzo (par 1), 2) esfuerzo y estimación post mortem de esfuerzo (par 2), 3) estimación post mortem de esfuerzo y estimación temprana de esfuerzo, 4) tiempo y estimación temprana de tiempo (par 4), 5) estimación post mortem de tiempo y tiempo (par 5) y 6) estimación post mortem de tiempo y estimación temprana de tiempo (par 6). En consecuencia, no puede descartarse el azar para explicar las diferencias entre las mediciones, estimaciones tempranas y post mortem con dicho nivel de confianza.

Paired Samples Test		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	99% Confidence				
					Lower	Upper			
Pair 1	ESFUERZO - EST TEMP ESF	0,50	3,78	0,85	-1,92	2,92	0,588	19	0,563
Pair 2	ESFUERZO - EST POST MORTEM ESF	-0,25	3,19	0,71	-2,29	1,80	-0,347	19	0,732
Pair 3	EST POST MORTEM ESF - EST TEMP ESF	0,75	2,95	0,66	-1,14	2,63	1,132	19	0,272
Pair 4	EST TEMP TIEMPO - TIEMPO	0,20	0,96	0,21	-0,42	0,81	0,923	19	0,368
Pair 5	EST POST MORTEM TIEMPO - TIEMPO	-0,03	0,91	0,20	-0,61	0,55	-0,151	19	0,882
Pair 6	EST POST MORTEM TIEMPO - EST TEMP TIEMPO	-0,23	0,77	0,17	-0,72	0,26	-1,334	19	0,198

Tabla 10 Prueba de diferencias de muestras emparejadas

Pruebas no paramétricas

Se aplicó la prueba de Wilcoxon de igualdad entre las distribuciones de los signos de las diferencias (tabla 11). La alta significación nos dice que las variables comparadas no difieren.

Test Statistics ^d		
	Z	Asymp. Sig. (2-tailed)
EST TEMP ESF - ESFUERZO	-.037 ^a	,970
EST POST MORTEM ESF - ESFUERZO	,000 ^b	1,000
EST TEMP ESF - EST POST MORTEM ESF	-.784 ^c	,433
ERR REL ETE - ERR REL EPME	-.336 ^c	,737
TIEMPO - EST TEMP TIEMPO	-1,045 ^c	,296
TIEMPO - EST POST MORTEM TIEMPO	-.336 ^c	,737
EST TEMP TIEMPO - EST POST MORTEM TIEMPO	-1,419 ^a	,156
ERR REL ETT - ERR REL EPMT	-1,195 ^a	,232

a. Based on negative ranks.
b. The sum of negative ranks equals the sum of positive ranks.
c. Based on positive ranks.
d. Wilcoxon Signed Ranks Test

Tabla 11 Prueba de Wilcoxon de igualdad entre distribuciones de los signos de las diferencias

5 Conclusiones

Las pruebas anteriores nos llevan a concluir que la diferencia entre las mediciones finales de tiempo y esfuerzo y los valores estimados tempranamente y las estimadas a partir de la información disponible post mortem no es estadísticamente significativa para la muestra estudiada. Esto sugiere que probablemente DeMarco tenía razón cuando decía que los sistemas de información data-strong pueden ser estimados a partir de los datos. Lo anterior, junto con la precisión y consistencia demostrada por los modelos según los criterios de Conte et al [2], permite, además, concluir que las métricas propuestas por Calero, Piattini Polo y Ruiz [1] tienen, además de las propiedades formales por ellos establecidas, el soporte empírico que los mismos autores advertían en su publicación que era necesario. La métrica DCXW establece la relación entre las variables independientes que miden la complejidad de una base de datos relacional y explicita la relación anticipada por DeMarco. Asimismo, podemos concluir que las estimaciones tempranas son suficientemente próximas a los valores finales de tiempo y esfuerzo a efectos prácticos (tabla 4). Obviamente, la validez de estas conclusiones se encuentra limitada por la amplitud de la muestra.

6 Líneas de trabajo futuras

La DCXW es una métrica importante para apoyar la gestión de cambios de un proyecto de desarrollo de software [7, 8] y permite predecir el esfuerzo y tiempo requeridos para completar el proyecto. Sin embargo, queda abierto el problema de definir una métrica adecuada para medir el avance del proyecto, elemento importante para apoyar la gestión. Es importante, además, ampliar la muestra y extraer conclusiones con un mayor grado de confianza.

7 Reconocimientos

Este trabajo no hubiera sido posible sin la colaboración de Karina Santo, José Luís Chalar, Gustavo Carriquiry y Claudia Araujo de Artech Consulting, Enrique Latorres y Jose Luís Subelzú del Departamento de Informática del Ministerio de Transporte y Obras Públicas de Uruguay, Juan Andrés Leiras del Departamento de Informática de Sanidad Policial, Óscar Camargo de la Universidad del Trabajo de Uruguay y la Universidad ORT Uruguay y Gonzalo Pérez y Joaquín González de CONEX Consulting. Representaron un invaluable aporte las conversaciones mantenidas con Ernestina Menasalvas tanto en UPM como en ORT. Esta investigación ha sido apoyada por estudiantes de grado y post grado del Laboratorio de Investigación en

Sistemas de Información y la Cátedra de Teoría de Universidad ORT Uruguay. Fueron muy importantes los aportes y constructivos comentarios realizados por Luis Olsina durante sus visitas a ORT como corrector externo de los proyectos finales de grado que participaron de esta investigación. Los comentarios de los revisores anónimos contribuyeron a mejorar la calidad de este trabajo. Los viajes a UPM del primer autor son financiados por el Programa de Desarrollo Tecnológico del BID.

8 Referencias

- 1 Calero, Coral, Piattini, Mario, Polo, Macario, Ruiz, Francisco. Grupo ALARCOS, Departamento de Informática, Universidad de Castilla La Mancha. Métricas para la evaluación de Complejidad de Bases de Datos Relacionales. Computación y Sistemas Vol. 3, N° 4, pp 264-273, 2000, CIC – IPN. ISSN 1405-5546
- 2 Conte, Samuel D., H. E. Dunsmore, and Vincent Y. Shen. 1986. Software engineering metrics and models. Menlo Park, CA: Benjamin/Cummings
- 3 DeMarco, T. Controlling Software Projects. New York: Yourdon, 1982.
- 4 Juristo, Natalia and Moreno Ana M. Basics of Software Engineering Experimentation. Kluwer Academic Publishers 2001.
- 5 Latorres, Enrique, Salvetto, Pedro, Larre Borges Uruguay, Nogueira Juan C, Una herramienta de apoyo a la gestión del proceso de desarrollo de software. CACIC 2003 6-10 octubre 2003 La Plata, Argentina
- 6 Salvetto, Pedro, Nogueira Juan C Size Estimation for Management Information Systems Based on Early Metrics :An Automatic Metric Tool Based in Formal Specifications. Proceedings of the International Conference on Computer Science, Software Engineering, Information Technology, e-Business and Applications (CSITeA'03), June 5-7, 2003 Rio de Janeiro, Brazil in Cooperation with the International Society for Computers and Their Applications (ISCA), USA Winona State University (WSU), USA Universidad Nacional de San Luis (UNSL), Argentina Net of National Universities with Computer Science Careers (RedUNCI), Argentina. Pags 72-77. ISBN 0-9742059-0-7.
- 7 Salvetto, Pedro, Nogueira Juan C, Segovia, Javier Modelos Automatizables de Estimación muy Temprana del Tiempo y Esfuerzo de Desarrollo de Software de Gestión XXX Conferencia Latinoamericana de Informática (CLEI2004), 27/9/04 – 1/10/04, Arequipa, Perú.
- 8 Salvetto, Pedro, Nogueira Juan C, Segovia, Javier. Gestión de Cambios Apoyada por Modelos Formales de Estimación de Tiempo y Esfuerzo. 4ª Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento IIISIC'04, Madrid, España, 3-5 Noviembre 2004.